frontiers
RESEARCH TOPICS

ANTICIPATION AND THE CONTROL OF VOLUNTARY ACTION

Topic Editors
Dorit Wenke and Rico Fischer

frontiers in
PSYCHOLOGY

FRONTIERS COPYRIGHT STATEMENT

Cover image provided by Ibbl sarl, Lausanne CH

ISSN 1664-8714
ISBN 978-2-88919-157-4
DOI 10.3389/978-2-88919-157-4

ABOUT FRONTIERS

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

FRONTIERS JOURNAL SERIES

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing.

All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

DEDICATION TO QUALITY

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.

Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view.

By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

WHAT ARE FRONTIERS RESEARCH TOPICS?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area!

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: researchtopics@frontiersin.org

ANTICIPATION AND THE CONTROL OF VOLUNTARY ACTION

Topic Editors:
Dorit Wenke, Humboldt University of Berlin, Germany
Rico Fischer, Technische Universität Dresden, Germany

A major hallmark in the adaptive control of voluntary action is the ability to anticipate short and long term future events. Anticipation in its various forms is an important prerequisite for (higher order) cognitive abilities such as planning, reasoning and the pursuit of both immediate goals and long-term goals that may even stand in opposition to immediate desires and needs (e.g., to invest in pension funds). Therefore, it is not surprising that diverse and rather independent research lines have evolved, all somehow targeting various anticipatory capacities that are involved in the control of voluntary action and thus, contribute to the uniqueness of human goal-directed behavior.

For example, prediction of the incentive value of action outcomes drives goal-directed instrumental behavior (e.g., Dickinson & Balleine, 2000; Rushworth & Behrens, 2008). Similarly, the Ideo-Motor Principle assumes that actions are selected and activated by the mere anticipation of the sensory experience they produce (e.g., James, 1890; Prinz, 1990). Furthermore, the degree of match between intended, anticipated and actual action effects has been proposed to be a major determinant of motor programming and online action corrections (Jeannerod, 1981), motor learning (e.g., Wolpert, Diedrichsen, & Flanagan, 2011), and the subjective sense of causing and controlling an action and its effects (Sense of Agency; e.g., Abell, Happé, & Frith, 2000).

The role of anticipation in the control of voluntary action, however, goes far beyond the anticipation of immediate action effects and desired goals. For instance, pre-cues and alerting signals are used for advance preparation of what to do (e.g., Meiran, 1996), when to act or expect an event onset (e.g., Callejas, Lupianez, & Tudela, 2004; Los & van der Heuvel, 2001; Nobre & Coull, 2010) and to anticipate conflict (e.g., Correa, Rao, & Nobre, 2009).

Voluntary action is influenced by the anticipation and prediction of mental effort in task processing (e.g., Song & Schwarz, 2008). In addition, the anticipation of long-term future social consequences (e.g., expected aloneness) has been shown to affect cognitive mechanisms involved in logic and reasoning (e.g., Baumeister, Twenge, & Nuss, 2002). Last but not least, learning of statistical contingencies (e.g., conflict frequency) leads to the anticipation and prediction of context-specific executive control requirements (e.g., Crump, Gong, & Milliken, 2006, Dreisbach & Haider, 2006).

The aim of the present special issue is to provide a platform that offers the possibility of cross-fertilization and enhanced visibility among to date rather segregated research lines.

Table of Contents

frontiers in
PSYCHOLOGY

EDITORIAL
published: 14 June 2013
doi: 10.3389/fpsyg.2013.00341

Anticipation and the control of voluntary action

Dorit Wenke[1]* and Rico Fischer[2]*

[1] Department of Psychology, Humboldt-Universität zu Berlin, Berlin, Germany
[2] Department of Psychology, Technische Universität Dresden, Dresden, Germany
*Correspondence: dorit.wenke@hu-berlin.de; fischer@psychologie.tu-dresden.de

Edited by:
Bernhard Hommel, Leiden University, Netherlands

A major hallmark in the adaptive control of voluntary action is the ability to anticipate short and long term future events. Anticipation in its various forms is an important prerequisite for cognitive abilities such as planning, reasoning and the pursuit of both immediate goals and long-term goals (e.g., to invest in pension funds) that sometimes stand in opposition to immediate desires and needs. Therefore, it is not surprising that diverse and rather independent research lines have evolved, all somehow targeting various anticipatory capacities that are involved in the control of voluntary action.

One line of research focuses on anticipating action effects. For example, ideomotor theory assumes that actions are selected and activated by the mere anticipation of the sensory experience they produce (James, 1890/1950). Similarly, prediction of the incentive value of action outcomes has been proposed to drive goal-directed instrumental behavior (Balleine and Dickinson, 1998). Furthermore, the degree of match between intended, anticipated and actual action effects seems to be a major determinant of motor programming and online corrections (Prablanc and Martin, 1992), motor learning (Wolpert et al., 2011), and the subjective sense of causing and controlling actions and their effects (the Sense of Agency; Frith et al., 2000). However, the role of anticipation in the control of voluntary action goes beyond the anticipation of action effects. For instance, pre-cues and alerting signals are used for preparing what to do (Meiran, 1996), when to act or expect an event, (Callejas et al., 2004) and for anticipating conflict (Correa et al., 2009). Similarly, learning of statistical contingencies leads to prediction of context-specific executive control requirements (Crump et al., 2006).

The aim of the present Research Topic has been to provide a platform that offers the possibility of cross-fertilization and enhanced visibility among to date rather segregated research lines concerning the role of anticipation in the control of voluntary action.

Many contributions address the role of anticipating action effects in controlling and understanding actions. Some deal with the role of anticipated value of action outcomes: Watson et al. (2012) provide a review on maladaptive drug seeking behavior from a learning theory perspective. Pezzulo et al. (2013) propose a model in which a single mixed controller balances habitual choice based on cached action values, and mental simulations of action outcomes that underlie goal directed behavior, depending on the usefulness of obtaining new information. Scherbaum et al. (2012) propose a model of temporal discounting—the tendency to choose smaller rewards delivered sooner instead of larger rewards delivered later—that focuses on response threshold and time framing as two factors determining choice behavior in inter-temporal choice.

Further contributions address the role of effect anticipation and feedback evaluation in the control and experience of action: Schilling and Cruse (2012) propose a predictive body model for planning robots' actions. Wang et al. (2012) present data showing that distorted visual movement feedback tends to affect action evaluation more strongly in old than in young adults. Haering and Kiesel (2012) demonstrate that prior causal beliefs influence intentional binding, a temporal illusion often seen as an indirect measure of sense of agency. Hommel and Keizer (2012) show that that object files can contain evaluative information regarding the match (success) viz. mismatch (failure) between predicted and experienced events. Poehlman et al. (2012) argue that supramodal integration through conscious states is primarily related to the skeletal muscle output system where anticipatory processes play a central role.

Thinnes-Elker et al. (2012) discuss different concepts of intention with respect to their implications for brain-machine-interfaces that "decode" brain activity for controlling artificial effectors. Because anticipating the consequences of one's own and others actions is an important aspect of social interactions and sport settings, Weigelt and Memmert (2012) investigated how the implicit processing of the stimulus layout in natural scenes affects the goal-side selection in soccer penalty shooting.

Predictive mechanisms are also involved in our ability to understand other people's actions, and even infants tend to interpret various action components with respect to action goals. In this line, Daum et al. (2012) demonstrate a dissociation between two measures often used to investigate expectations about goal-directed actions in infants, namely post-hoc looking times and predictive gaze. Henrichs et al. (2012) report evidence for an impact of goal salience on infants' goal anticipations of observed reaching actions, as measured by predictive gaze.

Another group of contributions focusses on the role of implicit or explicit cues that are utilized by the cognitive system to adjust cognitive control: Wendt et al. (2012) show that cue-based task preparation during task-switching is modulated by the validity of preceding trial task-cues. Strack et al. (2013) investigated cue-induced preparation, aiming at disentangling anticipatory control adjustments and prevention of upcoming conflict via task recoding. King et al. (2012) applied a model-based analysis and argue that context-specific proportion congruence effects may be accounted for by a prediction error-triggered shift in the decision criterion. Bugg and Crump (2012) provide a review on list-wide, item-specific and context-specific proportion congruence effects.

Reuss et al. (2012) report data suggesting that participants can form expectations of where an event will occur on the basis of non-consciously presented cues. Duthoo et al. (2012) highlight the role of task repetition expectancy in task-switching by varying switch rate contingencies. Fröber and Dreisbach (2012) report data showing that positive affect with low arousal reduced proactive control as indicated by response cueing effects. Umbach et al. (2012) explored how explicit expectations feed into preparatory processes, over and above demand for preparation.

A final group of contributions targets temporal anticipation in the control of voluntary action. Predicting the temporal onset of an event by means of a warning cue allows for temporal orienting and anticipation of an upcoming event. In a brief review, Weinbach and Henik (2012) discuss whether the temporal orienting function of warning cues can be dissociated from cue-based increases of alertness. Finally, de la Rosa et al. (2012) show that temporal preparation guided by regular rhythms is not subject to working memory interference and facilitates performance irrespective of concurrent working memory load.

REFERENCES[1]

Balleine, B. W., and Dickinson, A. (1998). Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology* 37, 407–419.

*Bugg, J. M., and Crump, M. J. C. (2012). In support of a distinction between voluntary and stimulus-driven control: a review of the literature on proportion congruent effects. *Front. Psychol.* 3:367. doi: 10.3389/fpsyg.2012.00367

Callejas, A., Lupianez, J., and Tudela, P. (2004). The three attentional networks: on their independence and interactions. *Brain Cogn.* 54, 225–227. doi: 10.1016/j.bandc.2004.02.012

Correa, A., Rao, A., and Nobre, A. C. (2009). Anticipating conflict facilitates controlled stimulus-response selection. *J. Cogn. Neurosci.* 21, 1461–1472. doi: 10.1162/jocn.2009.21136

Crump, M. J., Gong, Z., and Milliken, B. (2006). The context-specific proportion congruent Stroop effect: location as a contextual cue. *Psychon. Bull. Rev.* 13, 316–321.

*Daum, M. M., Attig, M., Gunawan, R., Prinz, W., and Gredebäck, G. (2012). Actions seen through babies' eyes: a dissociation between looking time and predictive gaze. *Front. Psychol.* 3:370. doi: 10.3389/fpsyg.2012.00370

*de la Rosa, M. D., Sanabria, D., Capizzi, M., and Correa, A. (2012). Temporal preparation driven by rhythms is resistant to working memory interference. *Front. Psychol.* 3:308. doi: 10.3389/fpsyg.2012.00308

*Duthoo, W., De Baene, W., Wühr, P., and Notebaert, W. (2012). When predictions take control: the effect of task predictions on task switching performance. *Front. Psychol.* 3:282. doi: 10.3389/fpsyg.2012.00282

Frith, C. D., Blakemore, S., and Wolpert, D. M. (2000). Explaining the symptoms of schizophrenia: abnormalities in the awareness of action. *Brain Res. Rev.* 31, 357–363.

*Fröber, K., and Dreisbach, G. (2012). How positive affect modulates proactive control: reduced usage of informative cues under positive affect with low arousal. *Front. Psychol.* 3:265. doi: 10.3389/fpsyg.2012.00265

*Haering, C., and Kiesel, A. (2012). Mine is earlier than yours: causal beliefs influence the perceived time of action effects. *Front. Psychol.* 3:393. doi: 10.3389/fpsyg.2012.00393

*Henrichs, I., Elsner, C., Elsner, B., and Gredebäck, G. (2012). Goal salience affects infants' goal-directed gaze shifts. *Front. Psychol.* 3:391. doi: 10.3389/fpsyg.2012.00391

*Hommel, B., and Keizer, A. W. (2012). Binding success and failure: evidence for the spontaneous integration of perceptual features and object evaluations. *Front. Psychol.* 3:581. doi: 10.3389/fpsyg.2012.00581

James, W. (1890/1950). *The Principles of Psychology.* Vol. 2. New York, NY: Dover.

*King, J. A., Donkin, C., Korb, F. M., and Egner, T. (2012). Model-based analysis of context-specific cognitive control. *Front. Psychol.* 3:358. doi: 10.3389/fpsyg.2012.00358

Meiran, N. (1996). Reconfiguration of processing mode prior to task performance. *J. Exp. Psychol. Learn. Mem. Cogn.* 22, 1423–1442.

*Pezzulo, G., Rigoli, F., and Chersi, F. (2013). The mixed instrumental controller: using value of information to combine habitual choice and mental simulation. *Front. Psychol.* 4:92. doi: 10.3389/fpsyg.2013.00092

*Poehlman, T. A., Jantz, T. K., and Morsella, E. (2012). Adaptive skeletal muscle action requires anticipation and "conscious broadcasting". *Front. Psychol.* 3:369. doi: 10.3389/fpsyg.2012.00369

Prablanc, C., and Martin, O. (1992). Automatic control during hand reaching at undetected two-dimensional target displacements. *J. Neurophysiol.* 67, 455–469.

*Reuss, H., Kiesel, A., Kunde, W., and Wühr, P. (2012). A cue from the unconscious – masked symbols prompt spatial anticipation. *Front. Psychol.* 3:397. doi: 10.3389/fpsyg.2012.00397

*Scherbaum, S., Dshemuchadse, M., and Goschke, T. (2012). Building a bridge into the future: dynamic connectionist modeling as an integrative tool for research on intertemporal choice. *Front. Psychol.* 3:514. doi: 10.3389/fpsyg.2012.00514

*Schilling, M., and Cruse, H. (2012). What's next: recruitment of a grounded predictive body model for planning a robot's actions. *Front. Psychol.* 3:383. doi: 10.3389/fpsyg.2012.00383

*Strack, G., Kaufmann, C., Kehrer, S., Brandt, S., and Stürmer, B. (2013). Anticipatory regulation of action control in a Simon task: behavioral, electrophysiological, and fMRI correlates. *Front. Psychol.* 4:47. doi: 10.3389/fpsyg.2013.00047

*Thinnes-Elker, F., Iljina, O., Apostolides, J. K., Kraemer, F., Schulze-Bonhage, A., and Aertsen, A. (2012). Intention concepts and brain-machine interfacing. *Front. Psychol.* 3:455. doi: 10.3389/fpsyg.2012.00455

*Umbach, V. J., Schwager, S., Frensch, P. A., and Gaschler, R. (2012). Does explicit expectation really affect preparation? *Front. Psychol.* 3:378. doi: 10.3389/fpsyg.2012.00378

*Wang, L., Sutter, C., Müsseler, J., Dangel, R. J. Z., and Disselhorst-Klug, C. (2012). Perceiving one's own limb movements with conflicting sensory feedback: the role of mode of movement control and age. *Front. Psychol.* 3:289. doi: 10.3389/fpsyg.2012.00289

*Watson, P., de Wit, S., Hommel, B., and Wiers, R. W. (2012). Motivational mechanisms and outcome expectancies underlying the approach bias toward addictive substances. *Front. Psychol.* 3:440. doi: 10.3389/fpsyg.2012.00440

*Weigelt, M., and Memmert, D. (2012). Goal-side selection in soccer penalty kicking when viewing natural scenes. *Front. Psychol.* 3:312. doi: 10.3389/fpsyg.2012.00312

*Weinbach, N., and Henik, A. (2012). Temporal orienting and alerting – the same or different? *Front. Psychol.* 3:236. doi: 10.3389/fpsyg.2012.00236

*Wendt, M., Luna-Rodriguez, A., Reisenauer, R., Jacobsen, T., and Dreisbach, G. (2012). Sequential modulation of cue use in the task switching paradigm. *Front. Psychol.* 3:287. doi: 10.3389/fpsyg.2012.00287

Wolpert, D. M., Diedrichsen, J., and Flanagan, J. R. (2011). Principles of sensorimotor learning. *Nat. Rev. Neurosci.* 12, 739–751.

Received: 24 May 2013; accepted: 24 May 2013; published online: 14 June 2013.

Citation: Wenke D and Fischer R (2013) Anticipation and the control of voluntary action. Front. Psychol. 4*:341. doi: 10.3389/fpsyg.2013.00341*

This article was submitted to Frontiers in Cognition, a specialty of Frontiers in Psychology.

---

[1] Articles that are part of the research topic are marked with asterisks (∗).

frontiers in
PSYCHOLOGY

REVIEW ARTICLE
published: 22 October 2012
doi: 10.3389/fpsyg.2012.00440

Motivational mechanisms and outcome expectancies underlying the approach bias toward addictive substances

P. Watson[1,2], S. de Wit[2,3], Bernhard Hommel[4,5]* and R. W. Wiers[1,2]

[1] *ADAPT Lab, Department of Developmental Psychology, University of Amsterdam, Amsterdam, Netherlands*
[2] *Cognitive Science Center Amsterdam, University of Amsterdam, Amsterdam, Netherlands*
[3] *Department of Clinical Psychology, University of Amsterdam, Amsterdam, Netherlands*
[4] *Cognitive Psychology Unit, Leiden University, Leiden, Netherlands*
[5] *Leiden Institute for Brain and Cognition, Leiden, Netherlands*

Edited by:
Dorit Wenke, Humboldt University at Berlin, Germany

Reviewed by:
Giovanni Pezzulo, National Research Council of Italy, Italy
Andreas B. Eder, University of Würzburg, Germany

****Correspondence:***
Bernhard Hommel, Cognitive Psychology Unit, Department of Psychology, Leiden University, Room 2B05, Wassenaarseweg 52, 2333 AK Leiden, Netherlands.
e-mail: hommel@fsw.leidenuniv.nl

Human behavior can be paradoxical, in that actions can be initiated that are seemingly incongruent with an individual's explicit desires. This is most commonly observed in drug addiction, where maladaptive behavior (i.e., drug seeking) appears to be compulsive, continuing at great personal cost. Approach biases toward addictive substances have been correlated with actual drug-use in a number of studies, suggesting that this measure can, in some cases, index everyday maladaptive tendencies. At present it is unclear whether this bias to drug cues is a Pavlovian conditioned approach response, a habitual response, the result of a Pavlovian-instrumental transfer process, or a goal-directed action in the sense that expectancy of the rewarding effects of drugs controls approach. We consider this question by combining the theoretical framework of associative learning with the available evidence from approach bias research. Although research investigating the relative contributions of these mechanisms to the approach bias is to date relatively limited, we review existing studies and also outline avenues for future research.

Keywords: approach, dual-process theory, addiction, associative learning, motivation, goal-directed action, habit, Pavlovian-instrumental transfer

INTRODUCTION

To what extent is human behavior under voluntary control? Drug addiction is an extreme example, where drug seeking continues despite negative social or interpersonal consequences. Although many drug users are fully aware of the negative consequences and seek treatment in order to abstain from drug use, risk of relapse remains high. This highlights the paradoxical, destructive characteristic of addiction: that drug-seeking behavior persists despite explicit motivations to the contrary. Understanding the cognitive and motivational mechanisms that maintain such behaviors may allow us to better understand action control in general.

The approach bias is a behavioral inclination to approach rather than avoid certain stimuli. Experimental research into the approach bias has provided evidence for correlations with actual drug use and it is theorized that an approach bias may contribute to problematic drug-related behavior (Stacy and Wiers, 2010). An important question that remains to be addressed, however, is what the approach bias represents and how it relates to other features of drug use such as craving. Specifically, it is not clear whether the approach bias has the characteristics of being a goal-directed behavior, controlled by the expectancy of a rewarding outcome. Alternatively, it may better fit the profile of a Pavlovian conditioned response, or of a persistent, habitual response to drug cues, or it may be driven by Pavlovian-instrumental interactions.

In the present paper, we provide a comprehensive overview of these possible mechanisms that may facilitate the approach bias toward addictive substances. To this end, we relate the approach bias to theories of addiction as well as theoretical concepts of associative learning theory (based on fundamental animal as well as human behavioral research). Furthermore, we critically review the experimental measurement of approach bias and the evidence that it can underlie maladaptive behaviors. The existing literature does not allow firm conclusions to be drawn about the relative contributions of different mechanisms, although we hope that this manuscript will inspire empirical investigations of this issue. To further stimulate such investigations, we will outline several possible avenues for future research at the end of this article.

THEORETICAL BACKGROUND

Addiction has been described as a progressive neurological disorder of learning and memory whereby everyday associative learning processes become pathological (Hyman, 2005; Koob and Volkow, 2010). Teenagers experimenting with alcohol, for example, may discover that it makes them feel sociable and lively. The initial learning of associations between rewarding outcomes and the contexts or behaviors that lead to them, allows for the emergence of goal-directed behavior (e.g., approaching the bar at a party *in order* to feel sociable). For many teenagers, this goal-directed behavior can over time become habitual, such that simply being in a party context is the impetus to move toward the bar, regardless of any consideration of possible (pleasant) outcomes. For some individuals it may ultimately also persist when this behavior has undesirable consequences. Those individuals that continue to consume alcohol despite pervasive negative consequences and sometimes even explicit intentions to abstain may be regarded as *compulsive* drug-users (Koob and Volkow, 2010).

A number of learning and reward processes underlie this behavioral transition from voluntary drug use to clinically recognized drug dependency whereby the consumption is maintained at increasingly high cost. Associative learning theories, arising from the systematic study of animal behavior and neurobiology, offer a coherent framework for defining and dissociating these processes (see Balleine et al., 2008; Balleine and O'Doherty, 2009). Initial empirical investigations in humans support the case for applying the associative framework to human behavior (review: de Wit and Dickinson, 2009). The psychological constructs arising from associative learning theory are also paralleled in some neurocomputational models of decision-making, and we refer the interested reader to recent publications on this topic for an in-depth discussion (Daw et al., 2005; Dayan et al., 2006; Balleine et al., 2008). We shall discuss approach behavior within the associative learning context, because this context provides the most useful concepts to understand and investigate approach and avoidance behavior. At the same time, however, we emphasize that a purely non-propositional, associative framework can be argued to be insufficient as an exhaustive account of human decision-making (see for example Mitchell et al., 2009, but see also associated commentaries).

PAVLOVIAN OR CLASSICAL CONDITIONING

Associations between environmental cues (e.g., beer at a party) and motivationally relevant events (e.g., feeling lively and sociable) are acquired over time and may play an important role in guiding everyday decision-making. The conditions under which such associations emerge have been studied by means of Pavlovian (or classical) conditioning methods that establish a predictive relationship between these (Pavlov, 1927). Conditioned stimuli (CS) are cues that were once irrelevant (e.g., a food bowl) and that through repetitive pairing with a motivationally relevant stimulus (e.g., food) come to elicit conditioned responses. These conditioned responses can be consummatory in nature (e.g., salivation) or preparatory (approach toward the food bowl; Konorski, 1967). Furthermore, it has been shown in humans that Pavlovian conditioning can lead to acquired likes and dislikes of previously neutral objects and places (evaluative conditioning; Hermans et al., 2002; Hofmann et al., 2010). These processes may play an important role in drug-seeking behavior as contexts that were previously paired with the rewarding experience of drug taking become preferred and will elicit conditioned approach responses that may support drug seeking.

PAVLOVIAN VERSUS INSTRUMENTAL LEARNING

While conditioned approach may well contribute to the approach bias, there may also be an instrumental component. Whereas Pavlovian behavior results from the contingencies between stimuli and motivationally relevant events, instrumental behavior arises from the contingency between a response and a motivationally relevant outcome (Skinner, 1938). If the outcome is rewarding, the instrumental agent will acquire the responses that lead to this outcome.

Embedded within any instrumental contingency is also a Pavlovian relationship between the context and the outcome. For example, where a light stimulus (S) may signal that a lever press response (R) will produce a food pellet outcome (O) there is an inherent, parallel S–O relationship between the light and the rewarding food pellet being conditioned. Often a conditioned Pavlovian response will facilitate instrumental behavior (e.g., salivation or approach to the food bowl will facilitate eating behavior). There are occasions, however, when contradicting instrumental and Pavlovian responses (whether preparatory or consummatory) can cause conflict (Sheffield, 1965; Hershberger, 1986). Hershberger (1986), for example, created a "looking glass world" in which a food bowl receded with twice the speed at which hungry chicks ran toward it, and drew near at twice the speed at which the chicks ran away from it. To gain access to the food the chicks had to learn to overcome the Pavlovian bias to approach the food bowl, which acted as a CS for the food it contained. Most chicks continued to run toward the bowl, however, and thereby lost the available food. The approach bias of these chicks was clearly controlled predominantly by Pavlovian conditioning, as sensitivity to the R–O contingency should have allowed them to learn to make the opposite response of running away from the food bowl. On the other hand, it is well-known that animals, as well as humans, are capable of instrumental behavior. In a later section we will review the evidence for an instrumental component of the approach bias in humans as measured in the laboratory.

FROM GOAL-DIRECTED ACTIONS TO HABITUAL RESPONSES

Instrumental conditioning could contribute to the approach bias by giving rise to either goal-directed approach or to habitual approach that is triggered directly by environmental stimuli. Goal-directed actions are performed in order to achieve desirable outcomes (e.g., approaching the bar at a party to feel more lively and sociable) and are thus flexibly modulated by the incentive value of the outcome (Adams and Dickinson, 1981). Over time, however, these appetitive outcomes gradually reinforce S–R associations, that give rise to habitual responding that is directly evoked by the context. In this scenario, the party context triggers approach, rather than consideration of the drinking outcome. Overtraining of an instrumental action is one way to bring about habitual responding (Adams, 1982; Dickinson, 1985; Tricomi et al., 2009). In the early stages of drug use, drug-seeking behavior appears to meet the criteria for goal-direction action. Habitual drug seeking triggered by certain cues and contexts may, however, help to maintain drug-seeking behavior, even when the drug is no longer desired.

In animal studies, the degree to which behavior is goal-directed or habitual is formally assessed by means of the outcome devaluation procedure. In this procedure instrumental training is followed by devaluation of the instrumental outcome (e.g., through satiation on a particular food reward). Subsequently, an extinction test is conducted to assess instrumental responding for the devalued outcome. If behavior is predominantly under goal-directed control, responding for the devalued outcome should be immediately reduced. In contrast to goal-directed behavior, S–R habits are not sensitive to devaluation of outcomes and such behavior will persist.

PAVLOVIAN-TO-INSTRUMENTAL TRANSFER

A popular beer brand logo can prompt thoughts of beer drinking, which may increase the probability that an individual will head to the nearest bar and realize that outcome. This anticipatory effect is formally described as Pavlovian-to-instrumental transfer (PIT). In lab demonstrations of this phenomenon, a common outcome (such as a food reward) functions as both a Pavlovian reward and also an instrumental reward, in separate training phases. This training allows for the separate development of both Pavlovian S–O expectancies and instrumental O–R associations. The interaction effect is then assessed by presenting the Pavlovian cue whilst the subject is given the opportunity to perform the instrumental response for that outcome. Many studies have shown that although the Pavlovian cue was never directly paired with a response, the expectancy elicited by the cue increases the likelihood of instrumental responding for that specific reward (outcome-specific PIT) or in some cases boosts responding generally (general PIT; Estes, 1948; Rescorla and Solomon, 1967; Corbit and Balleine, 2005). PIT effects are now well documented in humans (Hogarth et al., 2007; Bray et al., 2008; Talmi et al., 2008; Hogarth and Chase, 2011; Huys et al., 2011; Nadler et al., 2011; Hogarth, 2012) and both forms of transfer (specific and general) could play a role in instrumental approach behavior. For example, the sight of a beer brand logo may remind one of beer drinking which may activate approach behavior that is previously been instrumental in obtaining beer, via S–O–R associations. General PIT, on the other hand, can only further strengthen a preexisting bias. For example, if there is already an approach bias toward alcohol, then any reward-associated cue (such as a cigarette logo for smokers) may further increase that bias by boosting the dominant approach response. Intriguingly, in animals, PIT effects have been shown to be insensitive to outcome devaluation (Rescorla, 1994; Holland, 2004) suggesting that these could play an important role in addiction relapse. Two recent studies in smokers provided evidence that Pavlovian cues predicting smoking outcomes increase the likelihood of responding for cigarettes and that furthermore, this can occur regardless of the current incentive value of the smoking outcome (Hogarth and Chase, 2011; Hogarth, 2012). In these studies, Pavlovian cues increased responding for cigarettes even after participants had read health warnings about cigarettes (Hogarth and Chase, 2011) or been treated with nicotine replacement therapy (Hogarth, 2012) and, crucially, had decreased responding in the absence of the cues. It seems paradoxical that this behavior, controlled by the anticipation of the outcome it produces, is not modulated by the current incentive value of that outcome, but this effect has been convincingly demonstrated in animals and humans (Rescorla, 1994; Holland, 2004; Hogarth and Chase, 2011; Hogarth, 2012; but see Allman et al., 2010). It appears, therefore, that in outcome-specific Pavlovian-instrumental interactions, the representation of the outcome contains sensory, but not motivationally relevant information (see Delamater and Oakeshott, 2007). The result is that the perceptual characteristics of the outcome prompt the associated response, regardless of the current incentive value of that outcome. Future research should elucidate the exact mechanism that mediates outcome-specific PIT, but on the basis of these outcome-reevaluation studies, we will make a distinction in the remainder of this manuscript between goal-directed action and outcome-specific PIT, with the latter also being mediated by anticipation of the outcome but occurring independently of the incentive goal status of the outcome.

THEORIES OF ADDICTION

The motivation driving destructive drug-seeking behavior is a key component of all major theories of addiction and most provide some explanation for why environmental cues can trigger relapse, even after long periods of sobriety. Some theories make clear predictions about the mechanisms that could facilitate an approach bias toward drug cues and whether such a bias is goal-directed, stimulus-bound habits, or due to PIT anticipatory processes. Relevant to the discussion at hand are incentive sensitization, theories based on the role of expectancy, various dual-process models (including habit theories of addiction), and negative reinforcement models.

INCENTIVE SENSITIZATION

The incentive sensitization model proposes that repeated drug use causes neuroadaptations in mesolimbic dopaminergic systems controlling the incentive values assigned to drug stimuli (Robinson and Berridge, 1993; Berridge, 2007). Over time a pathological incentive value becomes attributed to drug cues and contexts prompting compulsive drug-taking. This incentive sensitization increases even whilst levels of subjective pleasure decrease over the course of addiction (defined as increased "wanting" even in the absence of "liking" a drug). Thus whilst drugs can become disliked and an individual may have explicit motivations to avoid them, cues remain extremely salient and continue to elicit craving and motivate approach behavior (Robinson and Berridge, 1993, 2000, 2001).

EXPECTANCY THEORIES

Some models propose that expectancies of drug outcomes play a crucial role in motivating drug-seeking behavior. Goldman and colleagues argue that drug use is a goal-directed choice, based on the expectation of the hedonic effect of the drug outcome (Goldman et al., 1987; Goldman, 2002). Following a meta-analysis of conditioning studies using tobacco rewards, Hogarth and Duka (2006) found evidence for the role of expectancies in drug-seeking behaviors. Recently, however, this view was extended, given demonstrations that cue-elicited anticipation of a cigarette reward prompted responding for that reward, even when incentive value was low. The authors suggested therefore that parallel goal-directed expectancies and PIT anticipatory processes jointly determine action control (Hogarth and Chase, 2011; Hogarth, 2012).

DUAL-PROCESS MODELS

There are various dual-process models of addiction. Similarly to Tiffany's habit theory of addiction (Tiffany, 1990; Tiffany and Conklin, 2000), associative theories make a distinction between goal-directed and stimulus-bound behaviors (Everitt et al., 2001; Everitt and Robbins, 2005; de Wit and Dickinson, 2009; Hogarth and Chase, 2011; Hogarth, 2012). Another group of dual-process models describe an automatic, appetitive system opposed by an executive control system (Strack and Deutsch, 2004; Wiers et al., 2007; Hofmann et al., 2008, 2009; Stacy and Wiers, 2010). These will be discussed in turn.

Habitual versus goal-directed control in dual-process models

Tiffany's habit theory of addiction (Tiffany, 1990; Tiffany and Conklin, 2000) proposes that over time, drug-taking "rituals" become automatic behavioral schema, prompted by the environment. Whilst there are many unique features within Tiffany's model, this transition from goal-directed to habitual behavior is also captured by associative dual-process models of addiction (Everitt et al., 2001; Everitt and Robbins, 2005). According to this view, the reinforcing effects of drugs lead to strong S–R associations between contextual stimuli and drug-seeking behaviors. Over time, approach behavior toward drugs becomes a habitual response, triggered by environmental cues. This behavioral transition appears to be paralleled by impaired functioning of cortico-striatal networks supporting goal-directed behavior (Porrino et al., 2004; Everitt and Robbins, 2005). In a slightly different vein, Hogarth and Chase (2011) argue that goal-directed and PIT anticipatory processes operate in an additive manner, jointly determining behavior.

Implicit versus explicit control in dual-process models

Another set of dual-process models are centered upon the notion that appetitive behavior can be automatically triggered by a variety of cues and that this behavior needs to be regulated by executive control processes, maintaining goal focus, and motivation to resist use and abuse of drugs (Strack and Deutsch, 2004; Wiers et al., 2007; Hofmann et al., 2008, 2009; Stacy and Wiers, 2010). Individual differences in impulsivity and cognitive control modulate the effectiveness of this regulation (Dawe et al., 2004; Hofmann et al., 2009; Peeters et al., 2012). The impulsive and executive dual processes can be mapped fairly well onto the stimulus-bound and goal-directed distinction of associative models, although it should be noted that the reflective control system is often argued to be propositional in nature, not associative (Strack and Deutsch, 2004).

Negative reinforcement theories

Alleviation of a negative affective state – either withdrawal symptoms or more generally, depression or stress – is a commonly cited cause of relapse (Carey and Correia, 1997; Shiffman and Waters, 2004; Kuntsche et al., 2005). Negative reinforcement theories highlight the role of internal cues (negative affective states) in prompting drug use (Koob and Le Moal, 1997; Baker et al., 2004; Eissenberg, 2004; Ahmed and Koob, 2005). Koob and colleagues propose that addiction is the result of dysfunction in not only the reward system but also the anti-reward system, driving aversive states (Ahmed and Koob, 2005; Koob and Le Moal, 2005). It is beyond the scope of the present paper to provide a detailed overview of negative reinforcement theories, but for the purposes of the current discussion we would like to note that similar motivational mechanisms may underlie drug-seeking based on the rewarding properties of drugs versus avoidance of aversive states (Baker et al., 2004; Eissenberg, 2004). Avoidance behavior in the context of drug use could be either a goal-directed strategy based on expectancies of the alleviating outcome, or a stimulus-response habit reinforced by alleviation of negative states, or the result of Pavlovian-instrumental interactions.

INTERIM CONCLUSIONS

Assessing the characteristics of the motivational mechanisms underlying addictive approach behavior, should provide evidence in favor of different models of addiction. These models overlap with respect to several common predictions. They all provide an explanation for why relapse can be triggered by environmental cues, whether this is due to S–R associations, incentive sensitization, or triggering of goal-directed expectancies. There are nonetheless a number of subtle distinctions. Many implicit/explicit dual-process models propose that the approach bias represents an automatic, positive evaluation of drug cues, which is argued to be distinct from explicit processes. Goal-directed expectancy theories on the other hand would argue that the bias arises due to positive expectancy of the drug outcome. Some unique predictions derive from these models, which can be empirically tested. For example, associative dual-process models predict that approach behavior will eventually be resistant to outcome devaluation as behavior transitions to habitual control. This is in stark contrast to goal-directed expectancy theories, which predict that decreases in outcome value will continue to reduce responding for drug outcomes.

EMPIRICAL DATA

A longstanding idea in psychological science is that a considerable amount of behavior is driven by rapid, evolutionary relevant, affective evaluations of stimuli. These affective evaluations classify all stimuli as either "negative" or "positive," facilitating in the latter case, approach behavior (Bindra, 1974; Dickinson and Dearing, 1979; Chen and Bargh, 1999; Fazio, 2001; Krieglmeyer et al., 2010).

Motivations and affective attitudes are commonly assessed via questionnaires. Unfortunately, however, conclusions from such explicit measures can be difficult as participants may lack insight into the driving forces behind their actions and choices. In the case of addiction, invalid self-reports may result from self-presentational strategies or self-deception in an attempt to maintain a positive self-image. Moreover, introspection has been argued to not be a reliable and objective method of assessing motivational states (Berridge et al., 2009; Schooler and Mauss, 2009; Wood and Neal, 2009; Neal et al., 2012). As behavior becomes more habitual over time, there may not be a corresponding shift in subjective awareness of this fact. In the case of drug relapse, *post hoc* evaluation of one's behavior may lead an individual to conclude that their behavior was motivated by a craving for the drug as opposed to being prompted by the external environment.

To overcome these difficulties and problems, a number of indirect, speeded reaction-time tasks have been developed to assess the valence and strength of affective evaluative associations (and the resulting approach behavior) without the need for explicit reflection on the part of the subject (Fazio, 2001; De Houwer, 2006; De Houwer et al., 2009). We focus here on measures of action tendencies, although it should be noted that varieties of the Implicit Association Task have also been used to assess approach and avoidance associations of a target category such as alcohol (Palfai and Ostafin, 2003; Ostafin and Palfai, 2006).

MEASUREMENTS OF APPROACH BIAS

In order to directly assess approach tendencies, a number of tasks have been developed that measure speed of approach toward (generally pictorial) stimuli. Approach bias is generally measured as the difference in reaction time on trials where participants make an approach movement (such as pulling a joystick) versus an avoid movement (pushing a joystick) to the pictorial stimuli on a computer screen. A number of tasks have been developed, and whilst they all measure approach bias, they are confusingly and interchangeably labeled as either stimulus-response compatibility (SRC) tasks (utilizing either a manikin or a joystick), approach avoidance tasks (AATs), or affective Simon tasks. To avoid confusion, we will use the explicit paradigm labels when describing these tasks in later sections.

THE MANIKIN TASK

The manikin task provides an indirect measure of approach and avoidance behaviors. Approach tendencies are assessed by calculating the difference in reaction times across two blocks of experimental trials. In the first block the participant moves a computerized manikin toward one category of stimuli (e.g., alcohol) and away from other stimuli (e.g., soft drinks). In a subsequent block this assignment reverses. Using this task, participants have been seen to approach positive words faster than they are avoided, with the reverse effect for negative words (De Houwer et al., 2001). In addition, the manikin task has been used to assess approach behaviors in studies focusing on eating disorders (Woud et al., 2011), obesity (Havermans et al., 2011), Pavlovian conditioning of neutral stimuli (Thewissen et al., 2007; van Gucht et al., 2008), as well as approach tendencies toward alcohol (Field et al., 2005b, 2008; Schoenmakers et al., 2008; van Hemel-Ruiter et al., 2011; Barkby et al., 2012), cigarettes (Mogg et al., 2003, 2005; Bradley et al., 2004; Field et al., 2005a), and cannabis (Field et al., 2006; Cousijn et al., 2012).

THE JOYSTICK TASK

The joystick task can be used to measure differences in reaction times when the participant pushes a joystick away from his/her body in response to stimuli as opposed to pulling the joystick toward his/her body on a subsequent block. This task has been used to study phobias and anxiety (see Roefs et al., 2011), lifestyle and fitness goals (Fishbach and Shah, 2006), and food deprivation manipulations (Seibt et al., 2007).

Whilst it was originally suggested that approach and avoidance movements are represented as stored motor patterns, triggered by automatic, affective stimuli evaluations, it has became increasingly clear that motor actions *per se* do not represent either approach or avoidance. The same motor response (such as arm flexion) may represent approach in one situation (moving something toward oneself) but avoidance in another situation (quickly moving hand away from a stimulus to be avoided; Chen and Bargh, 1999). Indeed many studies have now shown that it is an individual's interpretation of the result of the behavior that is important (i.e., is the stimulus moved closer or further away) and as such, neutral body movements can be interpreted as approach and avoidance actions depending on the outcome (Lavender and Hommel, 2007; Seibt et al., 2008; van Dantzig et al., 2008; Krieglmeyer et al., 2010). The zooming joystick task (ZJT) is thus a disambiguated version of the joystick task, designed to avoid misinterpretation and recategorization of pushing and pulling movements. The introduction of a zooming feature ensures that participants experience the illusion of stimuli moving away from them and coming toward them when they push or pull the joystick (this is achieved by reducing or enlarging the size of the picture). This zooming feature reduces the possibility of participants interpreting pulling movements as avoid rather than approach. Using the zooming version of the task (negative) approach tendencies to spiders have been assessed in spider phobia (Rinck and Becker, 2007).

By asking participants to respond to an irrelevant task feature such as orientation or location of the picture on screen instead of the content of the picture, the task may be rendered more implicit (De Houwer, 2003). Using the ZJT with *irrelevant feature instructions*, approach bias has been examined in heavy drinkers (Wiers et al., 2009, 2010), alcoholic patients (Wiers et al., 2011), at-risk adolescents (Peeters et al., 2012), and heavy cannabis users (Cousijn et al., 2011).

RELIABILITY AND VALIDITY

Studies correlating approach bias measures to real-life behavior, as opposed to self-report measurements, are limited in number (due in part to the complexity of such designs). Rinck and Becker (2007) found that an approach bias on the ZJT predicted actual approach behaviors to live spiders, over and above that which was predicted with spider-phobia questionnaires. More importantly for the present discussion, approach bias on an irrelevant feature version of the ZJT was correlated with the amount of alcohol drunk in what was described to participants as an unrelated consumer "taste test" following the task (Wiers et al., 2010).

The split-half reliability of the manikin and joystick tasks is variable but generally good when using task-relevant feature instructions (Rinck and Becker, 2007; Krieglmeyer and Deutsch, 2010; Field et al., 2011). The advantage of instructing participants to respond on the basis of a task irrelevant feature is that it makes the task less susceptible to explicit control on the part of the participant, the practical drawback is that compatibility effects tend to be smaller (Krieglmeyer and Deutsch, 2010; Field et al., 2011). This reduced effect may be due to the fact that attention is not drawn to the affective properties of the stimuli in the irrelevant feature version. Several studies have, nonetheless, demonstrated robust approach (or avoidance) biases using irrelevant feature instructions (De Houwer et al., 2001; Rinck and Becker, 2007; Seibt et al., 2007; Wiers et al., 2009, 2010; Veenstra and de Jong, 2010; Cousijn et al., 2011). The reliability of the irrelevant feature instruction version was found to be poor in one study (Krieglmeyer and Deutsch, 2010) whilst another reported reasonably good reliability (Cousijn et al., 2011).

There are, evidentially, pros and cons to the various approach/avoidance task versions. Two studies, conducted with both the manikin and the standard joystick tasks, unexpectedly failed to find evidence for a correlation between the two approach bias scores (Krieglmeyer and Deutsch, 2010; van Hemel-Ruiter et al., 2011). The reasons for this are not immediately clear although there are major differences in how approach and avoidance are conceptualized within the tasks. Both the standard joystick task (without the zooming feature) and the manikin task are susceptible to recategorization – the manikin can be recategorized as someone other than the self and the approach/avoid movements in the standard joystick task are relatively ambiguous. Future studies should carefully select the task paradigm, depending on the research question, a point that we will return to in a later section "Outstanding Questions and Future Directions."

APPROACH BIASES IN ADDICTION

Using these aforementioned approach/avoidance paradigms, addiction researchers have provided substantial evidence for a relationship between drug-approach bias and drug use. That is, although the approach bias is measured experimentally in a lab, with superficial key press or lever movements, the behavioral tendency to be faster at approaching rather than avoiding drug stimuli, does seem to confer information about drug behavior more generally. Approach tendencies have been demonstrated in heavy (non-clinical) users of alcohol (Schoenmakers et al., 2008; Wiers et al., 2009, 2010), social drinkers (Field et al., 2005b), and cigarette smokers (Field et al., 2005a; Thewissen et al., 2007). Cigarette smokers have been seen to show a greater approach bias than non-smokers (Mogg et al., 2003; Bradley et al., 2004), as do cannabis users versus non-users (Field et al., 2006; Cousijn et al., 2011). Hazardous (non-clinical) drinkers were seen to have a stronger approach bias compared to light drinkers (Field et al., 2008). These results suggest a reliable relationship between drug use and drug-approach bias, particularly when examining healthy participants with moderate levels of dependence.

It should be noted that patterns of results can differ depending on the populations studied. Whilst lighter drinkers showed a weaker approach bias (Field et al., 2008), this pattern was reversed in one study investigating light versus heavy cigarette smokers (Mogg et al., 2005). In addition, in contrast to the aforementioned alcohol studies (with students), three studies involving patients receiving treatment for alcoholism did not find stronger approach tendencies for alcohol pictures compared to soft drink pictures (Wiers et al., 2011; Barkby et al., 2012; Spruyt et al., in press). These studies are small in number, but differing patterns of results in different populations at different stages of addiction can likely tell us something about the role of explicit motivations in approach bias measurements. We will discuss this in further detail, in the later section "Outstanding Questions and Future Directions."

APPROACH TENDENCIES IN ADDICTION AND THE UNDERLYING MOTIVATIONS

The question remains as to which of the cognitive and motivational mechanisms outlined earlier in the "Theoretical Background" section contribute to the approach bias as measured in the lab. Evidence for the contribution of a Pavlovian component to experimental measures of the approach bias comes from studies that have shown that CS-reward learning quickly engenders an approach bias toward these novel CSs. Using different variations of AAT, approach bias has been conditioned toward novel Pavlovian stimuli predicting cigarette outcomes (Thewissen et al., 2007) and chocolate outcomes (van Gucht et al., 2008). A direct association between these stimuli and an instrumental approach response cannot mediate the approach bias to novel CSs, indicating therefore that Pavlovian mechanisms do play a role in the approach bias.

There is evidence, however, suggesting that the approach bias cannot be completely reduced to a purely Pavlovian conditioned response. As discussed previously, Hershberger (1986) showed that under conditions where chicks needed to make a withdrawal response in order to make a food bowl move toward them, they were unable to suppress the urge to approach the food bowl. This behavioral inflexibility provides evidence that the approach behavior of the chicks was predominantly controlled by a Pavlovian mechanism. In contrast to chicks, however, humans are perfectly well able to adapt their approach behavior. To our knowledge, the human equivalent of Hershberger's experiment has not been conducted yet, but a recent study did employ a similar design. In a manikin task, participants were required to make an initial brief avoid movement in order to approach positive words and an initial brief approach movement to avoid negative words. Krieglmeyer et al. (2011) showed that even when the initial movement is avoidance, participants will still react faster if the final outcome is that the manikin approaches positive words. The reverse was true for avoiding negative words such that even if the initial movement was to approach a negative word, participants reacted faster if the final outcome was avoidance. This study suggests that the approach bias is more complex than being a mere Pavlovian approach response, as the final outcome (is the stimulus further away or closer) and not the initial direction of movement (toward or away from the stimuli), seems to influence reaction times.

Retraining studies with the joystick task provide further evidence for instrumental control over approach behavior (Kawakami et al., 2007, 2008; Wiers et al., 2010, 2011). For example, Wiers et al. (2010) presented the vast majority of a set of alcohol pictures in the push rather than pull format and found that retraining reduced the approach bias toward these pictures. The observation that participants can modify the bias following training (avoiding appetitive pictures) suggests that the bias is more than a conditioned response and shows a degree of flexibility that is in line with an instrumental account of the approach bias.

The results discussed above suggest that the approach response is not a purely Pavlovian response, although it is challenging to disentangle the relative contributions of Pavlovian and instrumental mechanisms using these paradigms. The question remains, nonetheless, as to whether the approach bias is flexibly modulated by changes in incentive value of the outcome or whether it is merely triggered by the drug stimuli. A number of studies have found that approach bias measurements increase in line with self-report craving scores, a result that is generally interpreted to suggest that approach behavior is sensitive to the current incentive value of the outcome (Field et al., 2005b, 2008; van Gucht et al., 2008). However, this correlative finding should be interpreted with caution as it does not necessarily imply a causal relationship between craving and the approach bias. Furthermore, other studies failed to find evidence for a relationship between craving and the approach bias (Mogg et al., 2003; Thewissen et al., 2007; Wiers et al., 2010; Cousijn et al., 2011). However, none of those studies really address whether behavior is *immediately* sensitive to a change in the incentive outcome value. So far, such outcome-reevaluation designs have yielded mixed results. Two studies, using the manikin task, manipulated craving by giving participants a placebo drink in one session and a dose of alcohol in another. Approach bias scores to alcohol pictures (Schoenmakers et al., 2008) and smoking pictures (Field et al., 2005a) were then compared between the two sessions (alcohol or placebo). Both studies found that self-reported craving was higher in the alcohol session but there was no difference in the approach bias scores – a null effect that although difficult to interpret, is more in line with the habitual account. These results are in contrast to a study using the standard joystick task, that examined the effects of satiety on the approach bias (Seibt et al., 2007). Participant's responses to images of food were measured either before or after lunch and satiety did appear to reduce the bias in the non-deprived group, suggesting that approach behavior was driven by the current desire for food. Unfortunately, however, this study failed to include a neutral control picture condition, and we can therefore not ascertain whether hunger increased the approach movement toward food pictures specifically, or approach behavior generally. Still, these results suggest that this line of research should be extended further to critically assess motivational modulation of approach.

As discussed previously, some dual-process theories suggest that the approach bias results from an interaction between associative learning processes and explicit cognitive control processes. Barkby et al. (2012) provided correlational evidence for the importance of cognitive control, by testing patients receiving treatment for alcohol addiction on the manikin task. Their critical finding was that approach bias scores on the manikin task correlated with individual differences in explicit approach/avoidance intentions. Further evidence, that behavioral intentions can influence approach behavior, comes from a study using a variant of the ZJT (Sharbanee et al., in press). Rather than calculating approach bias as a difference score between the push and pull reaction times, this study made a distinction between "pull alcohol picture" trials and "push alcohol picture" trials – the former trial type assumed to be congruent with an appetitive tendency and the latter incongruent. Only incongruent trials, therefore, should demand recruitment of executive control processes (to overcome the appetitive tendency and push the alcohol picture away). As expected, results showed that working memory scores modulated reaction times in problem drinkers attempting (unsuccessfully) to control their alcohol consumption, but this effect was only observed on incongruent "push alcohol" trials. This suggests that the approach bias arises due to a complex interaction between the strength of the approach tendency and the ability to inhibit this tendency when required.

We should point out, however, that many other studies suggest that approach tendencies are not always under intentional control. These studies report seemingly "automatic" approach biases that are not in line with instrumental withdrawal intentions: participants scoring higher on a restrained eating scale showed a greater approach bias toward food cues (Veenstra and de Jong, 2010); smokers showed an approach bias toward smoking cues that they reported as unpleasant (Bradley et al., 2008); and appetitive Pavlovian stimuli inhibited instrumental withdrawal in situations where the instrumental withdrawal behavior was rewarded with money (Huys et al., 2011). It appears, therefore, that explicit intentions can sometimes influence approach, but a complete account of the approach bias will also have to encapsulate the role of associative learning processes.

To summarize, the evidence surveyed suggests that both Pavlovian and instrumental mechanisms play a role in facilitating the approach bias, but it is not yet clear how these processes interact or sum to produce this behavioral tendency. Furthermore it remains to be seen whether the approach bias is flexibly modulated by outcome value or has the characteristics of a habitual response to drug cues. Recent research examining instrumental responding for cigarette outcomes has argued that goal-directed and PIT processes operate in parallel, summing in an additive manner (Hogarth and Chase, 2011; Hogarth, 2012). However, the role of PIT in the approach bias remains to be empirically addressed. Furthermore, next to associative mechanisms, behavioral intentions may also modulate the approach bias, with one study suggesting that the approach bias measures some combination of both appetitive and regulatory control processes (Sharbanee et al., in press).

OUTSTANDING QUESTIONS AND FUTURE DIRECTIONS

Approach bias tasks offer a fast and simple manner of measuring approach tendencies to drug-related stimuli and appear to tell us something about drug use, given that a number of studies have correlated approach bias scores with actual drug use (Field et al., 2008; Wiers et al., 2010; Cousijn et al., 2011) and shown group differences between heavier versus lighter/non-users (Mogg et al., 2003; Bradley et al., 2004; Field et al., 2006; Cousijn et al., 2011). Whilst understanding the mechanisms that underlie the approach bias is an important theoretical question, it should be noted that these tasks are not ideally suited to dissociating the various motivation mechanisms introduced in the earlier "Theoretical Background" section. To isolate goal-directed approach from Pavlovian approach for example, requires a task where the instrumental actions are bidirectional (i.e., left and right). In such a task, the relationships between the stimulus and the outcomes are held constant whilst the relationships between the *direction* of action and the outcomes are manipulated. The relative contribution of Pavlovian processes is equal to both actions, and hence controlled for (Dickinson et al., 1996). In addition, it has been observed that outcome devaluation modulates both conditioned Pavlovian responses and goal-directed instrumental responses (Colwill and Rescorla, 1988), and as such we cannot differentiate between these two in an outcome-reevaluation study if the approach bias is the dependent measure. Nonetheless, whether the approach bias is flexibly modulated by outcome reevaluation or is directly triggered by the drug stimuli is an outstanding question. The studies that have employed outcome-reevaluation paradigms have yielded mixed results (Field et al., 2005a; Seibt et al., 2007; Schoenmakers et al., 2008) and this line of research within the context of addictive substances, should be continued.

Specifically, outcome-reevaluation studies conducted with individuals at different levels of dependency, could address the question of whether the approach bias becomes more habitual over the course of addiction. Given the observation that users receiving clinical treatment may not show a very strong approach bias (Wiers et al., 2011; Barkby et al., 2012; Spruyt et al., in press), this method could also be used to assess whether control over approach behavior is regained during (successful) treatment. The work of

Hogarth and colleagues in the field of smoking addiction suggests that there are a number of ways to reevaluate addictive substances, namely health warnings, temporary satiety through consumption, and treatments aimed at alleviating withdrawal symptoms (Hogarth and Chase, 2011; Hogarth, 2012). Another way may be to pair the consumption of an appetitive substance with an aversive flavor (Howard, 2001; van Gucht et al., 2010). In order to conduct an outcome-reevaluation test, an approach bias measurement would first be taken with neutral and category of interest (e.g., smoking) pictures. Then the smoking outcome would be devalued (e.g., through satiety) and the approach bias measurement would be repeated. If the approach bias measurement is not reduced following outcome devaluation, this would suggest that the approach bias is a stimulus-bound response to drug stimuli. Different versions of the approach bias tasks, as reviewed earlier, may be better suited to reevaluation studies given that ideally a repeated measures design is employed and that the second measurement, following outcome devaluation, should be conducted in extinction (without presentation of the outcome). The standard joystick or manikin tasks therefore, with irrelevant feature version, would be preferable in such a paradigm – as participant awareness of the study aims should be reduced as much as possible.

Pavlovian-instrumental interactions are thought to play an important role in addictive behavior (Hogarth and Chase, 2011; Hogarth, 2012). However, whether PIT processes *can* confer a specific or general motivating effect on approach/avoidance movements on these tasks has not been assessed. This could be investigated using, for instance, the manikin task. In an initial instrumental (O–R) learning phase, participants would make an approach movement to earn one specific outcome and avoidance to earn another (e.g., approach is rewarded with beer; avoidance rewarded with wine). This would be followed by Pavlovian (S–O) training where participants would learn the predictive relationship between neutral stimuli and these same outcomes. During the subsequent transfer test, occasional stimulus presentations would be expected to facilitate/speed up the response associated with a common outcome. For example, a stimulus for beer would be expected to facilitate approach, while a stimulus for wine should facilitate avoidance. General PIT, on the other hand, could be assessed by comparing the influence of a stimulus associated with a third outcome (e.g., whiskey) versus one associated with no alcoholic drink.

fMRI research of the approach bias is scarce, but given the wealth of knowledge that exists concerning the neural correlates of the various motivational mechanisms highlighted in this review (see Balleine and O'Doherty, 2009 for a detailed overview), this could be a very fruitful avenue for investigating many of the questions raised thus far. Approach bias to cannabis using a manikin task was recently investigated in an fMRI study with heavy cannabis users (Cousijn et al., 2012). Results suggested that ventromedial prefrontal cortex (vmPFC) was recruited during congruent "approach cannabis" blocks as opposed to incongruent "avoid cannabis" blocks. The vmPFC/orbitofrontal cortex, along with the caudate, have been consistently implicated in goal-directed action (Valentin et al., 2007; de Wit et al., 2009, 2012b). Similar prefrontal regions are argued to encode Pavlovian outcome values (Gottfried et al., 2002, 2003). The results of Cousijn and colleagues suggest, therefore, that the approach bias is driven by mechanisms flexibly modulated by outcome value, as opposed to habits – the latter being mediated not by prefrontal regions, but instead networks involving the posterior putamen and premotor cortex (Tricomi et al., 2009; Ashby et al., 2010; de Wit et al., 2012b). We should point out however that another fMRI study reported contradictory findings – namely that vmPFC was recruited on incongruent and not congruent trials (Roelofs et al., 2009). Important methodological differences could account for these differential results – Roelofs and colleagues used a standard joystick task with affective, non-drug-related facial stimuli (happy/unhappy faces) and the effects disappeared when participants were instructed to approach/avoid on the basis of gender rather than facial expression (irrelevant feature version). Future studies should hopefully be able to resolve these findings.

Research examining the neural correlates of approach tendencies would be best suited to approach bias paradigms with high internal reliability – such as the relevant feature version of the manikin task. A question of interest is whether networks implicated in goal-directed control versus habitual control are recruited, and whether this is different between groups who are at different stages of addiction. In humans, differential regions of the amygdala are thought to mediate general and outcome-specific PIT (Prévost et al., 2012) and their engagement in approach bias tasks could also be examined. fMRI can also be used to assess the role of brain regions such as the anterior cingulate, known to be important in overcoming response conflict and cognitive control more generally (reviews: Botvinick et al., 2004; Ridderinkhof et al., 2004), during approach bias tasks. Although the fMRI studies mentioned earlier examined contrasts of either congruent > incongruent or incongruent > congruent – reporting the results of both contrasts would be beneficial to test hypotheses relating to the interaction between appetitive responses and explicit cognitive control.

Pharmacological manipulations could also be employed to investigate the effect of neurotransmitter depletion on approach tendencies, with dopamine being an obvious candidate. Females submitted to dopamine precursor depletion, for example, were seen to rely more on habitual S–R knowledge at the expense of goal-directed O–R knowledge in a task designed to assess the relative balance of these two systems (de Wit et al., 2012a). A GABA antagonist (Baclofen), also thought to have effects via mediation of dopaminergic systems, reduces craving and consumption in alcohol, and cigarette addiction (Franklin et al., 2009; Gorsane et al., 2012) yet the effect on the approach bias has not been studied. Studies of this type would help with attempts to understand what, exactly, the approach bias represents and how it relates to other measures such as craving.

Finally, the extent to which these manipulations selectively affect approach versus avoidance (rather than the composite approach bias score) is worth investigation. Some studies have started to tease apart the relative contributions of "congruent" approach responses to appetitive stimuli as opposed to "incongruent" avoidance movements away from appetitive stimuli, with interesting insights (Roelofs et al., 2009; Barkby et al., 2012; Cousijn et al., 2012; Sharbanee et al., in press). By looking at these processes separately we can gain a better understanding of what the approach bias actually measures and what role various cognitive and motivational mechanisms play in producing this effect.

CONCLUSION

Human behavior can be paradoxical, in that actions are initiated that are seemingly incongruent with an individual's explicit motivations. This is most commonly observed in addiction, where maladaptive behavior (i.e., drug seeking) appears to be compulsive. Different theoretical approaches attempt to explain this behavior in different ways, with some suggesting that positive expectancies drive such behavior and others arguing that environmental stimuli can trigger behaviors incongruent with current goals and behavioral intentions. A number of studies have observed correlations between problematic drug use and approach bias scores (Field et al., 2008; Cousijn et al., 2011) suggesting that approach bias measurements can index everyday behaviors. Understanding the cognitive and motivational mechanisms that drive such an approach bias may therefore provide insight into both adaptive and maladaptive action control.

Determining the mechanisms underlying approach may have clinical implications. Cognitive therapy may be useful if expectancies and cognitive control are important determinants of approach. On the other hand, alternative approaches such as exposure response prevention therapy or counter-conditioning, that target the behavior directly, may be more appropriate if approach is stimulus-bound (e.g., van Gucht et al., 2010). Retraining the approach bias using the ZJT has been shown to be effective in reducing approach behavior toward alcohol cues (Wiers et al., 2010), and in a clinical population such retraining of the approach bias leads to a significantly smaller relapse rate at one-year follow-up compared to individuals receiving sham training (Wiers et al., 2011). A better understanding of the motivational mechanisms that underlie the approach bias, will also provide a better understanding of what exactly is being trained in this novel paradigm and how it can be better improved as a viable treatment.

Integration of the literature on approach bias, motivation, and associative learning provides a clear framework with which to identify and disentangle the relative contributions of various cognitive and motivational mechanisms underlying such maladaptive behavior. Whilst the literature surveyed suggests that both Pavlovian and instrumental mechanisms contribute to the approach bias, it remains to be elucidated exactly how they interact and sum to produce approach behavior. Hopefully further research assessing these questions will be forthcoming, within the limits that are inherent to such a task paradigm. Understanding the mechanisms that underlie an approach bias will provide a better understanding of the complex interplay of automatic processes, outcome expectancies, and behavioral intentions underlying human action control. This is not only theoretically important but ultimately has implications for clinical treatment.

ACKNOWLEDGMENTS

This PhD studentship is made possible through funding received from the Netherlands Organization for Scientific Research (NWO; 433-09-243).

REFERENCES

Adams, C. D. (1982). Variations in the sensitivity of instrumental responding to reinforcer devaluation. *Q. J. Exp. Psychol. Sect. B Comp. Physiol. Psychol.* 34, 77–98.

Adams, C. D., and Dickinson, A. (1981). Instrumental responding following reinforcer devaluation. *Q. J. Exp. Psychol. Sect. B Comp. Physiol. Psychol.* 33, 109–121.

Ahmed, S. H., and Koob, G. F. (2005). Transition to drug addiction: a negative reinforcement model based on an allostatic decrease in reward function. *Psychopharmacology (Berl.)* 180, 473–490.

Allman, M. J., DeLeon, I. G., Cataldo, M. F., Holland, P. C., and Johnson, A. W. (2010). Learning processes affecting human decision making: an assessment of reinforcer-selective Pavlovian-to-instrumental transfer following reinforcer devaluation. *J. Exp. Psychol. Anim. Behav. Process.* 36, 402–408.

Ashby, F. G., Turner, B. O., and Horvitz, J. C. (2010). Cortical and basal ganglia contributions to habit learning and automaticity. *Trends Cogn. Sci. (Regul. Ed.)* 14, 208–215.

Baker, T. B., Piper, M. E., McCarthy, D. E., Majeskie, M. R., and Fiore, M. C. (2004). Addiction motivation reformulated: an affective processing model of negative reinforcement. *Psychol. Rev.* 111, 33–51.

Balleine, B. W., Daw, N. D., and O'Doherty, J. P. (2008). "Multiple forms of value learning and the function of dopamine," in *Neuroeconomics: Decision Making and the Brain*, ed. P. W. Glimcher (Waltham, MA: Academic Press), 367–385.

Balleine, B. W., and O'Doherty, J. P. (2009). Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology* 35, 48–69.

Barkby, H., Dickson, J. M., Roper, L., and Field, M. (2012). To approach or avoid alcohol? Automatic and self-reported motivational tendencies in alcohol dependence. *Alcohol. Clin. Exp. Res.* 36, 361–368.

Berridge, K. (2007). The debate over dopamine's role in reward: the case for incentive salience. *Psychopharmacology (Berl.)* 191, 391–431.

Berridge, K. C., Robinson, T. E., and Aldridge, J. W. (2009). Dissecting components of reward: "liking," "wanting," and learning. *Curr. Opin. Pharmacol.* 9, 65–73.

Bindra, D. (1974). A motivational view of learning, performance, and behavior modification. *Psychol. Rev.* 81, 199–213.

Botvinick, M. M., Cohen, J. D., and Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: an update. *Trends Cogn. Sci.* 8, 539–546.

Bradley, B. P., Field, M., Healy, H., and Mogg, K. (2008). Do the affective properties of smoking-related cues influence attentional and approach biases in cigarette smokers? *J. Psychopharmacol.* 22, 737–745.

Bradley, B., Field, M., Mogg, K., and De Houwer, J. (2004). Attentional and evaluative biases for smoking cues in nicotine dependence: component processes of biases in visual orienting. *Behav. Pharmacol.* 15, 29.

Bray, S., Rangel, A., Shimojo, S., Balleine, B., and O'Doherty, J. P. (2008). The neural mechanisms underlying the influence of Pavlovian cues on human decision making. *J. Neurosci.* 28, 5861–5866.

Carey, K. B., and Correia, C. J. (1997). Drinking motives predict alcohol-related problems in college students. *J. Stud. Alcohol* 58, 100–105.

Chen, M., and Bargh, J. A. (1999). Consequences of automatic evaluation: immediate behavioral predispositions to approach or avoid the stimulus. *Pers. Soc. Psychol. Bull.* 25, 215–224.

Colwill, R. M., and Rescorla, R. A. (1988). Associations between the discriminative stimulus and the reinforcer in instrumental learning. *J. Exp. Psychol. Anim. Behav. Process.* 14, 155.

Corbit, L. H., and Balleine, B. W. (2005). Double dissociation of basolateral and central amygdala lesions on the general and outcome-specific forms of Pavlovian-instrumental transfer. *J. Neurosci.* 25, 962–970.

Cousijn, J., Goudriaan, A. E., Ridderinkhof, K. R., van den Brink, W., Veltman, D. J., and Wiers, R. W. (2012). Approach-bias predicts development of cannabis problem severity in heavy cannabis users: results from a prospective FMRI study. *PloS ONE* 7, e42394. doi:10.1371/journal.pone.0042394

Cousijn, J., Goudriaan, A. E., and Wiers, R. W. (2011). Reaching out towards cannabis: approach-bias in heavy cannabis users predicts changes in cannabis use. *Addiction* 106, 1667–1674.

Daw, N. D., Niv, Y., and Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat. Neurosci.* 8, 1704–1711.

Dawe, S., Gullo, M. J., and Loxton, N. J. (2004). Reward drive and rash impulsiveness as dimensions of impulsivity: implications for substance misuse. *Addict. Behav.* 29, 1389–1405.

Dayan, P., Niv, Y., Seymour, B., and Daw, N. D. (2006). The misbehavior of value and the discipline of the will. *Neural. Netw.* 19, 1153–1160.

De Houwer, J. (2003). "A structural analysis of indirect measures of attitudes," in *The Psychology of Evaluation: Affective Processes in Cognition and Emotion*, eds J. Musch and K. C. Klauer (Mahwah, NJ: Lawrence Erlbaum), 219–244.

De Houwer, J. (2006). "What are implicit measures and why are we using them," in *The Handbook of Implicit Cognition and Addiction*, eds R. W. Wiers and A. W. Stacy (Thousand Oaks, CA: Sage Publishers), 11–28.

De Houwer, J., Crombez, G., Baeyens, F., and Hermans, D. (2001). On the generality of the affective Simon effect. *Cogn. Emot.* 15, 189–206.

De Houwer, J., Teige-Mocigemba, S., Spruyt, A., and Moors, A. (2009). Implicit measures: a normative analysis and review. *Psychol. Bull.* 135, 347.

de Wit, S., Corlett, P. R., Aitken, M. R., Dickinson, A., and Fletcher, P. C. (2009). Differential engagement of the ventromedial prefrontal cortex by goal-directed and habitual behavior toward food pictures in humans. *J. Neurosci.* 29, 11330–11338.

de Wit, S., and Dickinson, A. (2009). Associative theories of goal-directed behaviour: a case for animal–human translational models. *Psychol. Res.* 73, 463–476.

de Wit, S., Standing, H. R., Devito, E. E., Robinson, O. J., Ridderinkhof, K. R., Robbins, T. W., et al. (2012a). Reliance on habits at the expense of goal-directed control following dopamine precursor depletion. *Psychopharmacology (Berl.)* 219, 621–631.

de Wit, S., Watson, P., Harsay, H. A., Cohen, M. X., van de Vijver, I., and Ridderinkhof, K. R. (2012b). Corticostriatal connectivity underlies individual differences in the balance between habitual and goal-directed action control. *J. Neurosci.* 32, 12066–12075.

Delamater, A. R., and Oakeshott, S. (2007). Learning about multiple attributes of reward in Pavlovian conditioning. *Ann. N. Y. Acad. Sci.* 1104, 1–20.

Dickinson, A. (1985). Actions and habits: the development of behavioural autonomy. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 308, 67–78.

Dickinson, A., Campos, J., Varga, Z. I., and Balleine, B. (1996). Bidirectional instrumental conditioning. *Q. J. Exp. Psychol. Sect. B Comp. Physiol. Psychol.* 49, 289–306.

Dickinson, A., and Dearing, M. F. (1979). "Appetitive-aversive interactions and inhibitory processes," in *Mechanisms of Learning and Motivation*, eds A. Dickinson and R. A. Boakes (Hillsdale, NJ: Erlbaum), 203–231.

Eissenberg, T. (2004). Measuring the emergence of tobacco dependence: the contribution of negative reinforcement models. *Addiction* 99(Suppl 1), 5–29.

Estes, W. K. (1948). Discriminative conditioning. II. Effects of a Pavlovian conditioned stimulus upon a subsequently established operant response. *J. Exp. Psychol.* 38, 173–177.

Everitt, B. J., Dickinson, A., and Robbins, T. W. (2001). The neuropsychological basis of addictive behaviour. *Brain Res. Rev.* 36, 129–138.

Everitt, B. J., and Robbins, T. W. (2005). Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. *Nat. Neurosci.* 8, 1481–1489.

Fazio, R. H. (2001). On the automatic activation of associated evaluations: an overview. *Cogn. Emot.* 15, 115–141.

Field, M., Caren, R., Fernie, G., and De Houwer, J. (2011). Alcohol approach tendencies in heavy drinkers: comparison of effects in a relevant stimulus-response compatibility task and an approach/avoidance Simon task. *Psychol. Addict. Behav.* 25, 697–701.

Field, M., Eastwood, B., Bradley, B. P., and Mogg, K. (2006). Selective processing of cannabis cues in regular cannabis users. *Drug Alcohol Depend.* 85, 75–82.

Field, M., Kiernan, A., Eastwood, B., and Child, R. (2008). Rapid approach responses to alcohol cues in heavy drinkers. *J. Behav. Ther. Exp. Psychiatry* 39, 209–218.

Field, M., Mogg, K., and Bradley, B. P. (2005a). Alcohol increases cognitive biases for smoking cues in smokers. *Psychopharmacology (Berl.)* 180, 63–72.

Field, M., Mogg, K., and Bradley, B. P. (2005b). Craving and cognitive biases for alcohol cues in social drinkers. *Alcohol* 40, 504–510.

Fishbach, A., and Shah, J. Y. (2006). Self-control in action: implicit dispositions toward goals and away from temptations. *J. Pers. Soc. Psychol.* 90, 820–832.

Franklin, T. R., Harper, D., Kampman, K., Kildea-McCrea, S., Jens, W., Lynch, K. G., et al. (2009). The GABA B agonist baclofen reduces cigarette consumption in a preliminary double-blind placebo-controlled smoking reduction study. *Drug Alcohol Depend.* 103, 30–36.

Goldman, M. S. (2002). Expectancy and risk for alcoholism: the unfortunate exploitation of a fundamental characteristic of neurobehavioral adaptation. *Alcohol. Clin. Exp. Res.* 26, 737–746.

Goldman, M. S., Brown, S. A., and Christiansen, B. A. (1987). "Expectancy theory: thinking about drinking," in *Psychological Theories of Drinking and Alcoholism*, eds H. T. Blane and K. E. Leonard (New York: Guilford Press), 181–226.

Gorsane, M. A., Kebir, O., Hache, G., Blecha, L., Aubin, H. J., Reynaud, M., et al. (2012). Is baclofen a revolutionary medication in alcohol addiction management? Review and recent updates. *Subst. Abuse*. 33, 336–349.

Gottfried, J. A., O'Doherty, J., and Dolan, R. J. (2002). Appetitive and aversive olfactory learning in humans studied using event-related functional magnetic resonance imaging. *J. Neurosci.* 22, 10829–10837.

Gottfried, J. A., O'Doherty, J., and Dolan, R. J. (2003). Encoding predictive reward value in human amygdala and orbitofrontal cortex. *Science* 301, 1104–1107.

Havermans, R. C., Giesen, J. C. A. H., Houben, K., and Jansen, A. (2011). Weight, gender, and snack appeal. *Eat. Behav.* 12, 126–130.

Hermans, D., Vansteenwegen, D., Crombez, G., Baeyens, F., and Eelen, P. (2002). Expectancy-learning and evaluative learning in human classical conditioning: affective priming as an indirect and unobtrusive measure of conditioned stimulus valence. *Behav. Res. Ther.* 40, 217–234.

Hershberger, W. (1986). An approach through the looking-glass. *Learn. Behav.* 14, 443–451.

Hofmann, W., De Houwer, J., Perugini, M., Baeyens, F., and Crombez, G. (2010). Evaluative conditioning in humans: a meta-analysis. *Psychol. Bull.* 136, 390–421.

Hofmann, W., Friese, M., and Strack, F. (2009). Impulse and self-control From a dual-systems perspective. *Perspect. Psychol. Sci.* 4, 162–176.

Hofmann, W., Gschwendner, T., Friese, M., Wiers, R. W., and Schmitt, M. (2008). Working memory capacity and self-regulatory behavior: toward an individual differences perspective on behavior determination by automatic versus controlled processes. *J. Pers. Soc. Psychol.* 95, 962–977.

Hogarth, L. (2012). Goal-directed and transfer-cue-elicited drug-seeking are dissociated by pharmacotherapy: evidence for independent additive controllers. *J. Exp. Psychol. Anim. Behav. Process.* 38, 266–278.

Hogarth, L., and Chase, H. W. (2011). Parallel goal-directed and habitual control of human drug-seeking: implications for dependence vulnerability. *J. Exp. Psychol. Anim. Behav. Process.* 37, 261–276.

Hogarth, L., Dickinson, A., Wright, A., Kouvaraki, M., and Duka, T. (2007). The role of drug expectancy in the control of human drug seeking. *J. Exp. Psychol. Anim. Behav. Process.* 33, 484–496.

Hogarth, L., and Duka, T. (2006). Human nicotine conditioning requires explicit contingency knowledge: is addictive behaviour cognitively mediated? *Psychopharmacology (Berl.)* 184, 553–566.

Holland, P. C. (2004). Relations between Pavlovian-instrumental transfer and reinforcer devaluation. *J. Exp. Psychol. Anim. Behav. Process.* 30, 104–117.

Howard, M. O. (2001). Pharmacological aversion treatment of alcohol dependence. I. Production and prediction of conditioned alcohol aversion. *Am. J. Drug Alcohol Abuse* 27, 561–585.

Huys, Q. J. M., Cools, R., Gölzer, M., Friedel, E., Heinz, A., Dolan, R. J., et al. (2011). Disentangling the roles of approach, activation and valence in instrumental and pavlovian responding. *PLoS Comput. Biol.* 7, e1002028. doi:10.1371/journal.pcbi.1002028

Hyman, S. E. (2005). Addiction: a disease of learning and memory. *Am. J. Psychiatry* 162, 1414–1422.

Kawakami, K., Phills, C. E., Steele, J. R., and Dovidio, J. F. (2007). (Close) distance makes the heart grow fonder: improving implicit racial attitudes and interracial interactions through approach behaviors. *J. Pers. Soc. Psychol.* 92, 957–971.

Kawakami, K., Steele, J. R., Cifa, C., Phills, C. E., and Dovidio, J. F. (2008). Approaching math increases math = me and math = pleasant. *J. Exp. Soc. Psychol.* 44, 818–825.

Konorski, J. (1967). *Integrative Activity of the Brain*. Chicago: University of Chicago Press.

Koob, G. F., and Le Moal, M. (1997). Drug abuse: hedonic homeostatic dysregulation. *Science* 278, 52–58.

Koob, G. F., and Le Moal, M. (2005). Plasticity of reward neurocircuitry and the "dark side" of drug addiction. *Nat. Neurosci.* 8, 1442–1444.

Koob, G. F., and Volkow, N. D. (2010). Neurocircuitry of addiction. *Neuropsychopharmacology* 35, 217–238.

Krieglmeyer, R., De Houwer, J., and Deutsch, R. (2011). How farsighted are behavioral tendencies of approach and avoidance? The effect of stimulus valence on immediate vs. ultimate distance change. *J. Exp. Soc. Psychol.* 47, 622–627.

Krieglmeyer, R., and Deutsch, R. (2010). Comparing measures of approach–avoidance behaviour: the manikin task vs. two versions of the joystick task. *Cogn. Emot.* 24, 810–828.

Krieglmeyer, R., Deutsch, R., De Houwer, J., and De Raedt, R. (2010). Being moved: valence activates approach-avoidance behavior independently of evaluation and approach-avoidance intentions. *Psychol. Sci.* 21, 607–613.

Kuntsche, E., Knibbe, R., Gmel, G., and Engels, R. (2005). Why do young people drink? A review of drinking motives. *Clin. Psychol. Rev.* 25, 841–861.

Lavender, T., and Hommel, B. (2007). Affect and action: towards an event-coding account. *Cogn. Emot.* 21, 1270–1296.

Mitchell, C. J., De Houwer, J., and Lovibond, P. F. (2009). The propositional nature of human associative learning. *Behav. Brain Sci.* 32, 183–198; discussion 198–246.

Mogg, K., Bradley, B. P., Field, M., and De Houwer, J. (2003). Eye movements to smoking-related pictures in smokers: relationship between attentional biases and implicit and explicit measures of stimulus valence. *Addiction* 98, 825–836.

Mogg, K., Field, M., and Bradley, B. P. (2005). Attentional and approach biases for smoking cues in smokers: an investigation of competing theoretical views of addiction. *Psychopharmacology* 180, 333–341.

Nadler, N., Delgado, M. R., and Delamater, A. R. (2011). Pavlovian to instrumental transfer of control in a human learning task. *Emotion* 11, 1112–1123.

Neal, D. T., Wood, W., Labrecque, J. S., and Lally, P. (2012). How do habits guide behavior? Perceived and actual triggers of habits in daily life. *J. Exp. Soc. Psychol.* 48, 492–498.

Ostafin, B. D., and Palfai, T. P. (2006). Compelled to consume: the implicit association test and automatic alcohol motivation. *Psychol. Addict. Behav.* 20, 322–327.

Palfai, T. P., and Ostafin, B. D. (2003). Alcohol-related motivational tendencies in hazardous drinkers: assessing implicit response tendencies using the modified-IAT. *Behav. Res. Ther.* 41, 1149–1162.

Pavlov, I. P. (1927). *Conditioned Reflexes*. London: Oxford University Press.

Peeters, M., Wiers, R. W., Monshouwer, K., van de Schoot, R., Janssen, T., and Vollebergh, W. A. M. (2012). Automatic processes in at-risk adolescents: the role of alcohol-approach tendencies and response inhibition in drinking behavior. *Addiction* 107, 1939–1946.

Porrino, L. J., Lyons, D., Smith, H. R., Daunais, J. B., and Nader, M. A. (2004). Cocaine self-administration produces a progressive involvement of limbic, association, and sensorimotor striatal domains. *J. Neurosci.* 24, 3554–3562.

Prévost, C., Liljeholm, M., Tyszka, J. M., and O'Doherty, J. P. (2012). Neural correlates of specific and general Pavlovian-to-instrumental transfer within human amygdalar subregions: a high-resolution fMRI study. *J. Neurosci.* 32, 8383–8390.

Rescorla, R. (1994). Transfer of instrumental control mediated by a devalued outcome. *Learn. Behav.* 22, 27–33.

Rescorla, R. A., and Solomon, R. L. (1967). Two-process learning theory: relationships between Pavlovian conditioning and instrumental learning. *Psychol. Rev.* 74, 151–182.

Ridderinkhof, K. R., Ullsperger, M., Crone, E. A., and Nieuwenhuis, S. (2004). The role of the medial frontal cortex in cognitive control. *Science* 306, 443–447.

Rinck, M., and Becker, E. S. (2007). Approach and avoidance in fear of spiders. *J. Behav. Ther. Exp. Psychiatry* 38, 105–120.

Robinson, T. E., and Berridge, K. C. (1993). The neural basis of drug craving: an incentive-sensitization theory of addiction. *Brain Res. Rev.* 18, 247–291.

Robinson, T. E., and Berridge, K. C. (2000). The psychology and neurobiology of addiction: an incentive-sensitization view. *Addiction* 95(Suppl 2), S91–117.

Robinson, T. E., and Berridge, K. C. (2001). Incentive-sensitization and addiction. *Addiction* 96, 103–114.

Roefs, A., Huijding, J., Smulders, F. T. Y., MacLeod, C. M., de Jong, P. J., Wiers, R. W., et al. (2011). Implicit measures of association in psychopathology research. *Psychol. Bull.* 137, 149–193.

Roelofs, K., Minelli, A., Mars, R. B., van Peer, J., and Toni, I. (2009). On the neural control of social emotional behavior. *Soc. Cogn. Affect. Neurosci.* 4, 50–58.

Schoenmakers, T., Wiers, R. W., and Field, M. (2008). Effects of a low dose of alcohol on cognitive biases and craving in heavy drinkers. *Psychopharmacology* 197, 169–178.

Schooler, J. W., and Mauss, I. B. (2009). "To be happy and to know it: the experience and meta-awareness of pleasure," in *Pleasures of the Brain*, eds M. L. Kringelbach and K. C. Berridge (New York: Oxford University Press), 244–254.

Seibt, B., Häfner, M., and Deutsch, R. (2007). Prepared to eat: how immediate affective and motivational responses to food cues are influenced by food deprivation. *Eur. J. Soc. Psychol.* 37, 359–379.

Seibt, B., Neumann, R., Nussinson, R., and Strack, F. (2008). Movement direction or change in distance? Self- and object-related approach–avoidance motions. *J. Exp. Soc. Psychol.* 44, 713–720.

Sharbanee, J. M., Stritzke, W. G., Wiers, R. W., Young, P., Rinck, M., and MacLeod, C. M. (in press). The interaction of approach-alcohol action tendencies and working memory capacity predicts the inability to regulate drinking behaviour. *Psychol. Addict. Behav*.

Sheffield, F. D. (1965). "Relation between classical and instrumental conditioning," in *Classical Conditioning*, ed. W. F. Prokasy (New York: Appleton-Century-Crofts), 302–322.

Shiffman, S., and Waters, A. J. (2004). Negative affect and smoking lapses: a prospective analysis. *J. Consult. Clin. Psychol.* 72, 192–201.

Skinner, B. F. (1938). *The Behavior of Organisms*. New York: Appleton-Century-Crofts.

Spruyt, A., De Houwer, J., Tibboel, H., Verschuere, B., Crombez, G., Verbanck, P., et al. (in press). On the predictive validity of automatically activated approach/avoidance tendencies in abstaining alcohol-dependent patients. *Drug Alcohol Depend*.

Stacy, A. W., and Wiers, R. W. (2010). Implicit cognition and addiction: a tool for explaining paradoxical behavior. *Annu. Rev. Clin. Psychol.* 6, 551–575.

Strack, F., and Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Pers. Soc. Psychol. Rev.* 8, 220–247.

Talmi, D., Seymour, B., Dayan, P., and Dolan, R. J. (2008). Human Pavlovian-instrumental transfer. *J. Neurosci.* 28, 360–368.

Thewissen, R., Havermans, R., Geschwind, N., van den Hout, M., and Jansen, A. (2007). Pavlovian conditioning of an approach bias in low-dependent smokers. *Psychopharmacology* 194, 33–39.

Tiffany, S. T. (1990). A cognitive model of drug urges and drug-use behavior: role of automatic and nonautomatic processes. *Psychol. Rev.* 97, 147–168.

Tiffany, S. T., and Conklin, C. A. (2000). A cognitive processing model of alcohol craving and compulsive alcohol use. *Addiction* 95, 145–153.

Tricomi, E., Balleine, B. W., and O'Doherty, J. P. (2009). A specific role for posterior dorsolateral striatum in human habit learning. *Eur. J. Neurosci.* 29, 2225–2232.

Valentin, V. V., Dickinson, A., and O'Doherty, J. P. (2007). Determining the neural substrates of goal-directed learning in the human brain. *J. Neurosci.* 27, 4019–4026.

van Dantzig, S., Pecher, D., and Zwaan, R. A. (2008). Approach and avoidance as action effects. *Q. J. Exp. Psychol.* 61, 1298–1306.

van Gucht, D., Baeyens, F., Vansteenwegen, D., Hermans, D., and Beckers, T. (2010). Counterconditioning reduces cue-induced craving and actual cue-elicited consumption. *Emotion* 10, 688–695.

van Gucht, D., Vansteenwegen, D., Van den Bergh, O., and Beckers, T. (2008). Conditioned craving cues elicit an automatic approach tendency. *Behav. Res. Ther.* 46, 1160–1169.

van Hemel-Ruiter, M. E., de Jong, P. J., and Wiers, R. W. (2011). Appetitive and regulatory processes in young adolescent drinkers. *Addict. Behav.* 36, 18–26.

Veenstra, E. M., and de Jong, P. J. (2010). Restrained eaters show enhanced automatic approach tendencies towards food. *Appetite* 55, 30–36.

Wiers, R. W., Bartholow, B. D., van Den Wildenberg, E., Thush, C., Engels, R. C. M. E., Sher, K. J., et al. (2007). Automatic and controlled processes and the development of addictive behaviors in adolescents: a review and a model. *Pharmacol. Biochem. Behav.* 86, 263–283.

Wiers, R. W., Eberl, C., Rinck, M., Becker, E. S., and Lindenmeyer, J. (2011). Retraining automatic action tendencies changes alcoholic patients' approach bias for alcohol and improves treatment outcome. *Psychol. Sci.* 22, 490.

Wiers, R. W., Rinck, M., Dictus, M., and Van den Wildenberg, E. (2009). Relatively strong automatic appetitive action-tendencies in male carriers of the OPRM1 G-allele. *Genes Brain. Behav.* 8, 101–106.

Wiers, R. W., Rinck, M., Kordts, R., Houben, K., and Strack, F. (2010). Retraining automatic action-tendencies to approach alcohol in hazardous drinkers. *Addiction* 105, 279–287.

Wood, W., and Neal, D. T. (2009). The habitual consumer. *J. Consum. Psychol.* 19, 579–592.

Woud, M. L., Anschutz, D. J., Van Strien, T., and Becker, E. S. (2011). Measuring thinspiration and fear of fat indirectly. A matter of approach and avoidance. *Appetite* 56, 451–455.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 16 June 2012; accepted: 03 October 2012; published online: 22 October 2012.

*Citation: Watson P, de Wit S, Hommel B and Wiers RW (2012) Motivational mechanisms and outcome expectancies underlying the approach bias toward addictive substances. Front. Psychology **3**:440. doi: 10.3389/fpsyg.2012.00440*

This article was submitted to Frontiers in Cognition, a specialty of Frontiers in Psychology.

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 04 March 2013
doi: 10.3389/fpsyg.2013.00092

The Mixed Instrumental Controller: using Value of Information to combine habitual choice and mental simulation

Giovanni Pezzulo[1,2]*, **Francesco Rigoli**[2] **and Fabian Chersi**[2]

[1] *Istituto di Linguistica Computazionale, "Antonio Zampolli," Consiglio Nazionale delle Ricerche, Pisa, Italy*
[2] *Istituto di Scienze e Tecnologie della Cognizione, Consiglio Nazionale delle Ricerche, Roma, Italy*

Edited by:
Dorit Wenke, Humboldt University at Berlin, Germany

Reviewed by:
Bernard W. Balleine, University of Sydney, Australia
Sanne De Wit, University of Amsterdam, Netherlands
Nicolas W. Schuck, Max Planck Institute for Human Development, Germany

***Correspondence:**
Giovanni Pezzulo, Istituto di Scienze e Tecnologie della Cognizione, Consiglio Nazionale delle Ricerche, Via S. Martino della Battaglia, 44 – 00185 Roma, Italy.
e-mail: giovanni.pezzulo@cnr.it

Instrumental behavior depends on both goal-directed and habitual mechanisms of choice. Normative views cast these mechanisms in terms of model-free and model-based methods of reinforcement learning, respectively. An influential proposal hypothesizes that model-free and model-based mechanisms coexist and compete in the brain according to their relative uncertainty. In this paper we propose a novel view in which a single Mixed Instrumental Controller produces both goal-directed and habitual behavior by flexibly balancing and combining model-based and model-free computations. The Mixed Instrumental Controller performs a cost-benefits analysis to decide whether to chose an action immediately based on the available "cached" value of actions (linked to model-free mechanisms) or to improve value estimation by mentally simulating the expected outcome values (linked to model-based mechanisms). Since mental simulation entails cognitive effort and increases the reward delay, it is activated only when the associated "Value of Information" exceeds its costs. The model proposes a method to compute the Value of Information, based on the uncertainty of action values and on the distance of alternative cached action values. Overall, the model by default chooses on the basis of lighter model-free estimates, and integrates them with costly model-based predictions only when useful. Mental simulation uses a sampling method to produce reward expectancies, which are used to update the cached value of one or more actions; in turn, this updated value is used for the choice. The key predictions of the model are tested in different settings of a double T-maze scenario. Results are discussed in relation with neurobiological evidence on the hippocampus – ventral striatum circuit in rodents, which has been linked to goal-directed spatial navigation.

Keywords: model-based reinforcement learning, hippocampus, ventral striatum, goal-directed decision-making, exploration-exploitation, value of information, forward sweeps

1. INTRODUCTION

Goal-directed decision-making describes choice as depending on the evaluation of action-outcome contingencies (Balleine and Dickinson, 1998). Consider the case of a thirsty rat facing a T-maze with water in its left end. When behavior is controlled by goal-directed mechanisms of choice, the rat goes left because it predicts a water outcome (expectancy), and wants to reach it (goal state). Goal-directed mechanisms are considered to be very flexible as they rapidly readapt choice after changed conditions (e.g., devaluation of stimuli previously associated with high value). In contrast, habitual choice mechanisms rely on fixed stimulus-response reactions arising after extensive training. Consider again the case of the rat in the T-maze. If it has been rewarded a sufficient number of times for going left, it will tend to choose left again even if there is no reward. Compared to goal-directed mechanisms, habitual mechanisms are less flexible (e.g., they readapt very slowly after devaluation) but also faster and less demanding.

Normative views of animal behavior cast habitual and goal-directed mechanisms of choice in terms of model-free and model-based methods of reinforcement learning (RL), respectively (Daw et al., 2005). Model-free methods use "cached" action values to choose actions (i.e., aggregated values that can be recalled quickly). A long tradition of experimental and theoretical work in neuroscience uses model-free methods of RL, and in particular temporal-difference (TD) methods (Schultz et al., 1997), Q learning (Watkins and Dayan, 1992), and actor-critic architectures (Houk et al., 1995), to explain essential aspects of decision circuits such as dopamine bursts and the functioning of the basal ganglia.

Model-based methods use instead internal forward models to mentally simulate future action possibilities and their associated values. Model-based mechanisms are well known in the reinforcement learning literature (Sutton and Barto, 1981, 1998) and are nowadays increasingly studied in neuroscience and neuroeconomics in relation to perceptual, value-based, and economic choices (Pezzulo et al., 2007; Glimcher et al., 2009; Daw, 2012; Pezzulo and Rigoli, 2011; O'Doherty, 2012; Solway and Botvinick, 2012). Here we focus on goal-directed spatial navigation, which has been linked to the hippocampus – ventral striatum circuit in the rodent brain. It has been reported that rats navigating in mazes stop at decision points and turn the head in one of the possible directions, then to the other. When they turn their heads, place cells in the hippocampus "sweep forward" in the corresponding branch of the maze, as if the rat had really moved in that direction (Johnson and Redish, 2007). In correspondence of forward sweeps, ventral striatum activation is observed as well (van der Meer and Redish, 2009). Based on such evidence, it has been proposed that the hippocampus – ventral striatum circuit implements a *mental simulation* mechanism that realizes goal-directed choice, with the hippocampus linked to forward modeling and the ventral striatum linked to the evaluation of covert expectations of rewards constructed by the hippocampus (van der Meer and Redish, 2009, 2010, 2011; Battaglia et al., 2011; Pennartz et al., 2011; Chersi and Pezzulo, 2012; Erdem and Hasselmo, 2012; Penner and Mizumori, 2012; van der Meer et al., 2012). This view links well with the idea of a "vicarious trial and error" mechanism in rats (Tolman, 1948).

Habitual and goal-directed mechanisms of choice coexist and interact in the brain (Balleine and Dickinson, 1998). However, the proximal mechanisms that are responsible for their interactions are incompletely known. An influential theory proposes a continuous competition between habitual and goal-directed mechanisms of choice (implemented as two separate controllers) regulated by their relative *uncertainty* (Daw et al., 2005; Niv et al., 2006; Dayan, 2009). This theory captures the key role of uncertainty in the arbitration of goal-directed and habitual mechanisms of choice, and can reproduce (among the other things) the effects of *habitization*, or the gradual passage from goal-directed to habitual mechanisms after sufficient learning (Balleine and Dickinson, 1998). Mechanistically, this is due to the fact that the initial uncertainty of the habitual controller (compared to the goal-directed one) is higher (as it learns less efficiently from experience) but becomes lower after sufficient learning. This theory assumes that the model-free and model-based controllers are actively engaged in every decision (although ultimately only one of them is selected) and therefore it cannot explain why the hippocampal forward sweeps, putatively associated with model-based computations, vanish with habitization (van der Meer and Redish, 2009). Furthermore, this theory does not consider that model-based computations might have *costs*, linked to the cognitive effort due to planning (Gershman and Daw, 2011) and to the temporal discounting of rewards due to the time required for planning (Shadmehr, 2010).

We propose that a single instrumental process of decision-making produces both goal-directed and habitual behavior by flexibly combining aspects of model-based and model-free computations. We call this system a *Mixed Instrumental Controller (MIC)*. At decision points, the MIC performs a cost-benefits analysis, comparing the advantage of mental simulation (in terms of improving reward information) with its costs. More specifically, the MIC calculates the *Value of Information* (*VoI*; Howard, 1966) of mental simulation on the basis of uncertainty and of how much the alternative "cached" action values differ against each other. Then, the Value of Information is compared against the cost of mental simulation (in terms of cognitive effort and time). As a consequence of this, goal-directed mechanisms (mental simulations) are activated only when necessary, in line with evidence on rats' forward sweeps. In sum, the MIC combines model-based and model-free computations and does not lend itself to a complete separation of goal-directed and habitual controllers (in the strict sense devised in Daw et al., 2005); hence the label "mixed."

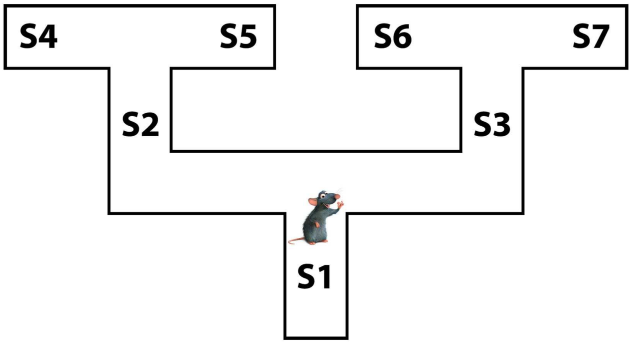

FIGURE 1 | The rat navigation scenario used in the simulations: a double T-maze.

In the rest of the article, we introduce the proposed *Mixed Instrumental Controller* model and test it in a simulated rat navigation scenario, in which decisions (going right or left) correspond to the selection of a branch in a double T-maze; see **Figure 1**. Rewards can be allocated at any of the seven points indicated as *S*1–*S*7. This scenario permits studying how selection of habitual vs. goal-directed processes at decision points changes as a function of learning, and to link elements of the model to neurobiological findings in rodents.

2. METHODS: THE MIXED INSTRUMENTAL CONTROLLER MODEL

Figure 2 illustrates the algorithm followed by the mixed instrumental controller model. This algorithm can be separated in four sub-processes, called meta-choice (between cached values and mental simulation), mental simulation, choice, and learning. Below, we describe each sub-process in details.

2.1. META-CHOICE BETWEEN CACHED VALUES AND MENTAL SIMULATION

At decision points (S1, S2, and S3), the agent (a simulated rat) has to decide whether to turn right or left. The agent has stored a prior estimate of each action value (*Q value*, see Watkins and Dayan, 1992), together with an estimate of each Q value uncertainty. Based on this information, at decision points, the agent first chooses whether to mentally explore the action consequences, in order to improve the action value estimates, or to simply rely on prior Q value estimates. This process can be viewed as a meta-choice between habitual (corresponding to "cached" Q values) and goal-directed processes (corresponding to mental simulation). At every decision point, this meta-choice is performed separately for each action (going left and right). In other words, the system might mentally simulate only the more uncertain action(s), not necessarily all.

This meta-choice amounts to computing the *Value of Information* (*VoI*; Howard, 1966) obtained with a mental simulation related to a given action *Act1* (e.g., going left at a decision point when left or right actions are possible). As solving an optimal solution to this problem is generally intractable in non-stationary

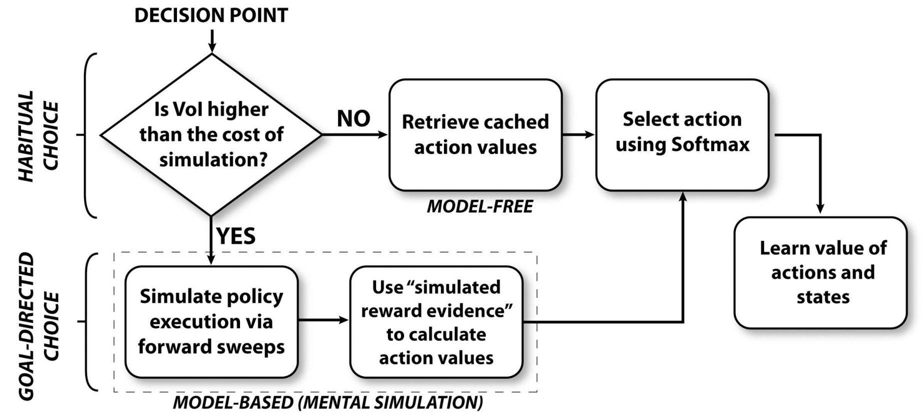

FIGURE 2 | Overview of the mixed instrumental controller (MIC).

environments, to determine VoI_{Act1} we adopt a simpler method described in equation (1):

$$VoI_{Act1} = \frac{C_{Act1}}{|Q_{Act1} - Q_{Act2}| + \epsilon} \quad (1)$$

This equation indicates that, for each action, our model considers two elements: (1) the difference between the Q_{Act1} value and the Q_{Act2} value of the alternative action (plus an ϵ to ensure that the sum is non-zero); (2) the uncertainty (C_{Act1}) relative to Q_{Act1}. The ratio between the two elements represents the estimated VoI_{Act1} obtained with mental simulation. This value is compared with the *cost* of mental simulation, which can be thought to be connected to the cognitive effort due to search (Gershman and Daw, 2011) and the temporal discounting of rewards due to the passing of time (Shadmehr, 2010). This cost is implemented here as a fixed threshold γ.

2.2. MENTAL SIMULATION

When VoI_{Act1} is smaller than the threshold γ, the agent relies on the cached Q_{Act1} value estimates for choice. On the contrary, when VoI_{Act1} is bigger than the threshold, forward sweeps are performed to simulate the effects of possible action executions. These simulated effects are then considered as pseudo-observations and are used to improve the estimation of Q_{Act1}.

Figure 3 shows the graphical model (Dynamic Bayesian Network; Murphy, 2002) used for mental simulation (see Botvinick and An, 2008; Dindo et al., 2011; Pezzulo and Rigoli, 2011; Solway and Botvinick, 2012 for related models). Nodes represent random variables including policies (π), actions (A), belief states (S), rewards (R), pseudo-observations (O) along with their temporal index *t*. Arrows connecting nodes indicate conditional probabilities among corresponding variables. Mental simulation consists in "clamping" current state and policy nodes (in other words, in considering these nodes as observed), and compute the conditional aggregated "value," which depends on the rewards gained at every time steps. The clamped policy at the first time step corresponds to the simulated action, while the policies clamped at following time steps are randomly chosen with equal probability. For instance, at S1 the agent could simulate the "going left" action by clamping the policy of going left at the first time step, and clamping a random policy (e.g., going right) at the second time step.

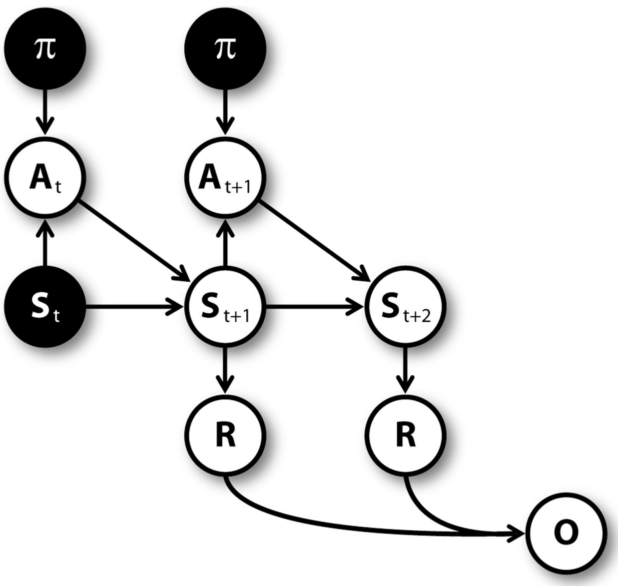

FIGURE 3 | Graphical model for mental simulations, unrolled for three time steps. Filled nodes are "clamped" (i.e., considered as observed) during mental simulation (see **Table 1**).

Mental simulations are repeated for several times, and every time the computed value is stored. The number of simulations is proportional to uncertainty (C_{Act1}); the proportion is regulated by a parameter λ. In addition, the number of simulated time steps for every simulation depends on uncertainty as well. Specifically, when uncertainty is higher than a threshold ζ, the agent simulates a sequence of actions (i.e., a whole path in the T-maze) and uses rewards to compute its aggregated value. Alternatively, the agent simulates a shorter path (whose length is regulated by a parameter η) and retrieves the Q value of one of the actions associated to

Table 1 | Nodes of the graphical model of Figure 3.

Node	Explanation	Values
π	Policies	{S1 → left, S2 → left, S3 → left} . . . {S1 → right, S2 → right, S3 → right}
A	Actions	Left, right (or equivalently: Act1, Act2)
R	Rewards	[0 . . . n]
S	Belief states	S1, S2, S3, S4, S5, S6, S7
O	Pseudo-observations	[0 . . . n]

the last simulated state. This Q value incorporates the cumulative expected value from that state on, rather than only the value of the state (i.e., it is a *return* and not a *reward* in reinforcement learning, see Sutton and Barto, 1998). Values relative to future states are discounted with a factor δ.

Once all mental simulations have been executed, the computed values are considered as pseudo-observations ($O_1, O_2, \ldots, O_n$, one for each simulation) and are used to improve the estimate of Q_{Act1}. The stored value is used as a prior(Q_{Act1}^{Prior}) and the pseudo-observations are used to compute a posterior value ($Q_{Act1}^{Posterior}$). This computation is described by equation (2) (assuming that the distribution variance of the Q_{Act1} value is known and is equal to 1, see Bishop, 2006):

$$Q_{Act1}^{Posterior} = \frac{Q_{Act1}^{Prior} + C_{Act1} \cdot \sum_{i=1}^{N} O_i}{1 + C_{Act1} \cdot N} \tag{2}$$

where C_{Act1} is the uncertainty, namely the prior variance on the mean of the Q_{Act1}^{Prior} value distribution, O_i is the pseudo-observation *i*, and N is the number of pseudo-observations.

2.3. CHOICE

At every decision point, a choice between actions is made by considering the value of the different possible actions (Q_{Act1} and Q_{Act2}). Note that this value can be either the cached Q value (if mental simulation was not used) or the posterior Q value calculated with equation (2) (if mental simulation was used). The choice is made according to the following softmax equation:

$$\begin{aligned} P\left(Action = Act1 | Q_{Act1}, Q_{Act2}\right) \\ = \frac{exp\left(\beta \cdot Q_{Act1}\right)}{exp\left(\beta \cdot Q_{Act1}\right) + exp\left(\beta \cdot Q_{Act2}\right)} \end{aligned} \tag{3}$$

where Q_{Act1} and Q_{Act2} are the Q values relative to the two possible actions (say going left or right at a decision point), and β is the inverse temperature parameter.

2.4. LEARNING

The MIC has two forms of learning.

2.4.1. *On-line learning of C and Q values*

Once the agent executes an action, he moves toward a new position and, in some cases, collects a reward. On the base of this novel experience, the agent learns. First, the Q_{Act1} value corresponding to the executed action is updated. The obtained reward, which is summed up to the Q value corresponding to the best action associated to the new position, is considered as an observation O. This observation is used to estimate the Q value at the following trial using the generative model represented by the graphical model shown in **Figure 4**. At every trial *x*, the prior $Q_{Act1,x}$ value and uncertainty $C_{Act1,x}$ are used by a particle filtering algorithm to compute the $Q_{Act1,x+1}$ value and the uncertainty $C_{Act1,x+1}$ at trial $x + 1$. The prior $Q_{Act1,x}$ value considered here is the "cached" Q_{Act1} value that is available before mental simulations (if any) were made.

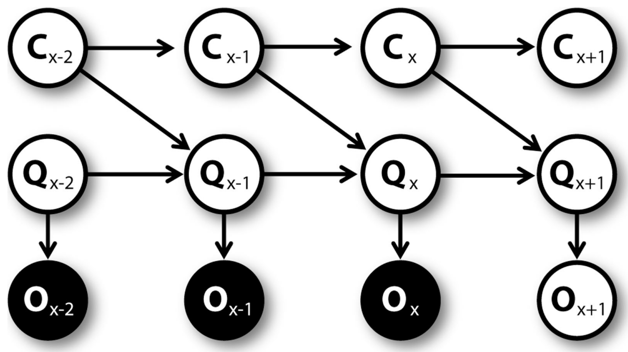

FIGURE 4 | Graphical model for learning *C* and *Q* values, unrolled.

The specific particle filtering algorithm is the following: for n = 1 to N, random vectors $[C_{Act1,n} Q_{Act1,n}]$ are sampled from the prior Gaussian distributions of uncertainty $N \sim (C_{Act1,x}, k)$ (where k is a known parameter) and of Q value $N \sim (Q_{Act1,x}, C_{Act1,n})$. Then, the sampled vectors are weighted proportionally to $P(O_{Act1,x}/Q_{Act1,n})$. After this, N vector samples are drawn from the previous vector set, each with a probability proportional to its weight. Finally, the posterior uncertainty is computed as $C_{Act1,x+1} = \sum C_{Act1,n}/N$ and the posterior Q value as $Q_{Act1,x+1} = \sum Q_{Act1,n}/N$.

2.4.2. *Value learning*

The model uses a model-based method to learn state values (i.e., the rewards *R* in the graphical model shown in **Figure 3**). Every time a reward is encountered in a state *s*, the mean of the expected reward conditional to that state $R(s)_{x+1}$ is updated according to equation (4):

$$R(s)_{t+1} = R(s)_t + \alpha\left(R_{observed} - R\left(s\right)_t\right) \tag{4}$$

where α is a learning rate.

3. RESULTS OF THE SIMULATIONS

We tested the MIC model in five simulated experiments. In the simulations, an artificial agent faced a double T-maze (see **Figure 1**) and, for several trials, had to choose twice to go either right or left. The simulations tested two key predictions of the model. First, we expected that the MIC model was able to learn the correct policy based on available rewards. Second, we expected that the MIC model executed forward sweeps only in certain circumstances, namely when the *VoI* was high. Specifically, we expected

Table 2 | Parameters and constants used in all the simulations.

Label	Explanation	Value
α	Learning rate for the model-based value representations	0.2
β	Inverse temperature parameter of the softmax function	0.4
γ	Threshold for mental simulation	0.5
–	Discount factor	1
ϵ	Small number used in the VoI to avoid division by zero	0.0001
ζ	Threshold relative to uncertainty for shortening the mental simulation	3
η	Length of the simulation when uncertainty is lower than ζ	1
–	Starting reward values for the model-based representations	1
–	Initial value of uncertainty in the simulations	4
κ	Uncertainty variance	1
λ	Number of forward sweeps during mental simulation	C × 3
–	Prior Q values at the first trial	1

to observe forward sweeps at the beginning of learning in all simulations. In addition, forward sweeps were expected to gradually decrease and disappear in simulations where variances were small and/or alternative Q values were not close to each other (simulations 1, 2, 4), contrary to simulation 5 where variances were high and alternative Q values were close to each other. Finally, we expected forward sweeps to reappear following unexpected changes in reward (simulation 3), and to decrease and disappear again as learning proceeded. In all the simulations, we assumed that the agent already knew the transition function, namely the conditional probabilities of outcomes given previous states and actions in the graphical model shown in **Figure 3**. The parameters and constants used in all the simulations are shown in **Table 2**.

3.1. SIMULATION 1: SIMPLE AND STABLE ENVIRONMENT WITH LOW VARIANCE

In the first simulation, a reward having a mean of 5 ($r = 5$) was placed at S7 (i.e., top right), while other positions had zero mean reward. Reward variance was relatively small for all positions, namely 0.2. The aim of this experiment was studying the gradual transition from goal-directed to habitual mechanisms of choice as a function of learning. Indeed, in stable environments, a given sequence of actions (in this case, right-right) is always reinforced and, after a certain amount of learning, can be selected by using habitual mechanisms, without the effort entailed by mental simulations. We hypothesized, as experience increased, a decrease in number and length of mental simulations (corresponding to goal-directed control), leading to relying on prior Q estimates (corresponding to habitual control).

Figure 5 describes the experimental results. **Figure 5A** shows the probability of choosing left turns at S1, S2, and S3. It shows a rapid decrease of preference for left turns at S1 and S3, as it was expected given that reward could be collected with two right turns. Turning right or left at S2 was equiprobable as neither S4 nor S5 were rewarded. **Figure 5B** shows the value of uncertainty along trials for going right at S1, which diminished rapidly. **Figure 5C** shows the number of samples used for the mental simulation for going right at S1, which is proportional to uncertainty. A value of zero indicates that the mental simulation is not used at all. Our results show that, during learning, mental simulations decreased in number, suggesting a gradual shift from goal-directed to habitual control. Moreover, **Figure 5D** indicates that, along learning, the length of forward sweeps decreased as well. The mechanisms tested in the present simulation can explain why learning in stable and simple environments produces habitization, which parallels a reduction (in number and length) of hippocampal forward sweeps and covert expectation of reward in ventral striatum (van der Meer and Redish, 2009). The development of habits entails also a "shift" of activation in dorsolateral striatum from actual reward locations to decision points and then to starting points (Jog et al., 1999). In our framework, this corresponds to the states in which the agent is highly confident of acquiring reward (i.e., at S7 before learning, at S3, and successively at S1 after learning).

3.2. SIMULATION 2: COMPLEX AND STABLE ENVIRONMENT WITH LOW VARIANCE

In the second simulation, multiple rewards were placed in the maze: S2 ($r = 2$), S4($r = 1$), S7 ($r = 5$). Like in the previous simulation, reward variances were relatively small (0.2). The goal of this simulation was to test whether the agent was able to shift from goal-directed to habitual control in a more complex environment.

Figure 6 describes the results. **Figure 6A** indicates that the agent was able to learn the correct policy. **Figure 6B** shows a decrease in uncertainty along learning for the action "going right" at S1. **Figures 6C,D** indicate that both the number and length of forward sweeps diminished along learning. The results of this simulation show that the MIC model can choose adaptively even in environments that have multiple rewards. In addition, due to the low reward variance, the model habituated (i.e., diminished forward sweeps) almost as fast as simulation 1. Compared to simulation 1, the choice of actions was more variable, matching the amount of rewards at different branches of the T-maze. This is due to the use of a softmax rule, which selects actions in proportion to their Q values rather than always selecting the action having the highest Q value.

3.3. SIMULATION 3: NON-STATIONARY ENVIRONMENT

In the third simulation, a single reward ($r = 5$) was initially placed at S7, and then moved to S4 after 50 trials. Reward variances were relatively small (0.2). The aim of this simulation was studying how the model re-adapts to novel contingencies. In other words, the agent had to learn an action sequence (right twice) and, after contingencies had changed, to re-learn a novel action sequence (left twice).

Figure 7 describes the results. **Figure 7A** shows that the policy was updated correctly in correspondence with the introduction of novel contingencies (Balleine and Dickinson, 1998). **Figure 7B** indicates that uncertainty decreased from trial 1 to 50, but, at this point, it increased again because previous contingencies had

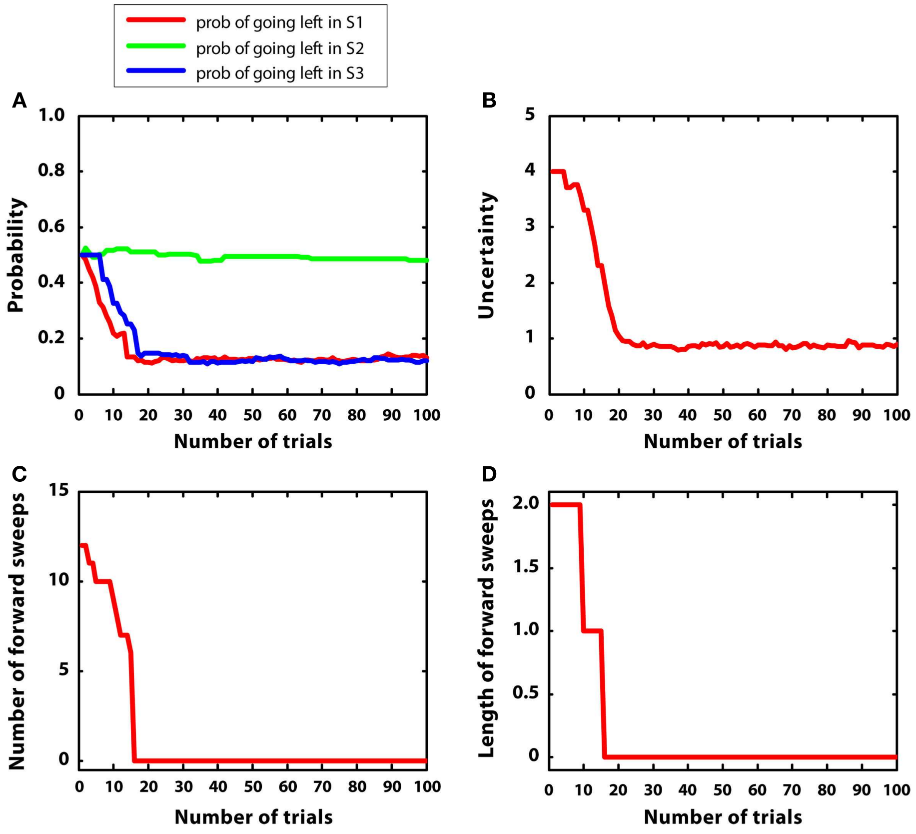

FIGURE 5 | Results of simulation 1, simple and stable environment with low variance, for 100 trials. (A) (top left) Shows the probability of going left at S1, S2, and S3 during the trials. **(B)** (top right) Plots the uncertainty for going right at S1. **(C)** (bottom left) Shows the number of forward sweeps (associated to mental simulation) used for the choice at S1; zero means that mental simulation is not used. **(D)** (bottom right) Shows the length of forward sweeps used for the choice at S1.

changed. This pattern was mirrored by the number and length of forward sweeps, shown in **Figures 7C,D**. These results show that the habitual system takes control in stationary environments but, after surprising outcomes are encountered, goal-directed mechanisms (corresponding to mental simulations) are activated again, due to a rapid uncertainty increase. This pattern of results suggests a specific prediction done by the MIC model in relation to the mechanisms regulating forward sweeps in rats, which requires empirical testing.

3.4. SIMULATION 4: SIMPLE ENVIRONMENT WITH HIGH VARIANCE

In this simulation, mean rewards were as in simulation 1 (i.e., $r = 5$ at S7), but with 5-times larger variances ($var = 1$). We hypothesized that uncertainty was bigger and less stable in this condition, compared to simulation 1. **Figure 8** describes the results. **Figure 8A** shows that the agent learned the correct policy (although beliefs were less stable compared to simulation 1). **Figure 8B** confirms that uncertainty was bigger and less stable than simulation 1. **Figures 8C,D** shows that, at the beginning of learning, mental simulations were activated for more trials compared to simulation 1. This is consistent with the idea that forward sweeps in the hippocampus are not only a function of experience (i.e., the more experience, the less forward sweeps) but also a function of environmental uncertainty (Gupta et al., 2010). However, with a certain amount of learning, in this simulation the habitual system took control as in simulation 1, and forward sweeps were no more activated. The reason was that, although variance was high, the environment was "simple." In other words, the difference between alternative Q values was big and the animal was quite confident about the best choice to take. This pattern of results represents a specific prediction of the MIC model, which requires empirical testing.

3.5. SIMULATION 5: COMPLEX ENVIRONMENT WITH HIGH VARIANCE

In the last simulation, mean rewards were like simulation 2 (i.e., $r = 2$ in S2; $r = 1$ in S4, $r = 5$ in S7); however, in this case, reward variances were bigger, namely they were equal to 1. The goal of this simulation was to observe the artificial agent in a complex environment with high variance. **Figure 9** describes the results. **Figure 9A** confirms that the agent was able to learn the correct policy, although beliefs were more noisy than in simulation 2. **Figure 9B** shows that uncertainty was bigger and less stable than in simulation 2. This led to activate mental simulations along the whole learning period (see **Figures 9C,D**) although to a larger extent at the beginning. The use of mental simulations along the whole learning period is caused by two factors. First, high

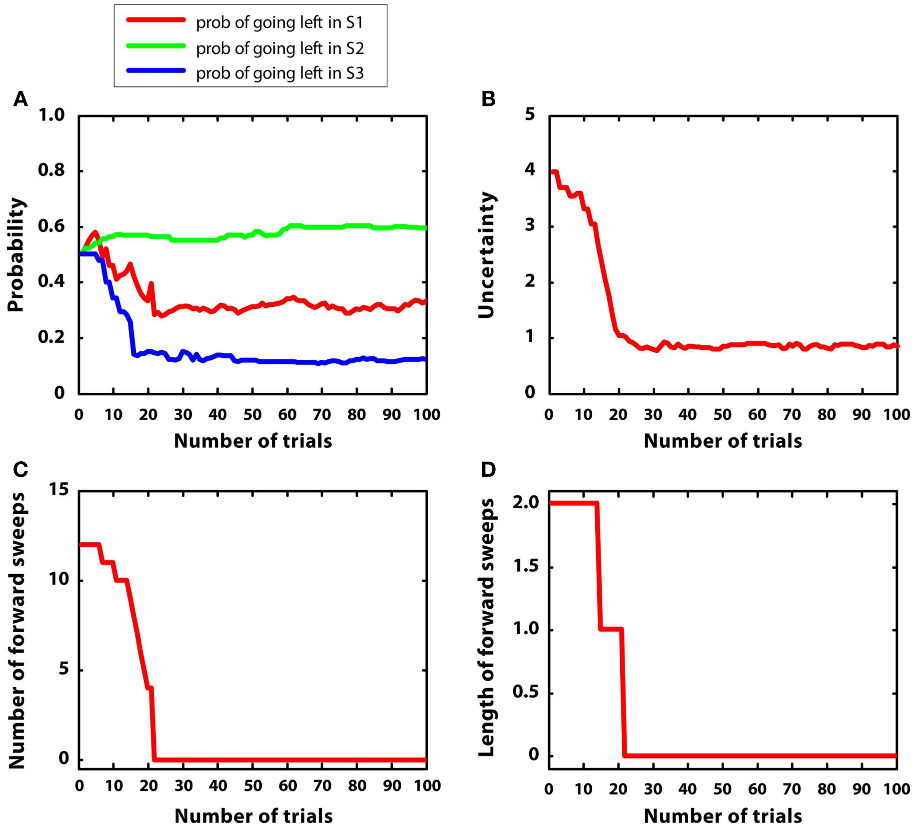

FIGURE 6 | Results of simulation 2, complex and stable environment with low variance. (A–D): see **Figure 5**.

reward variance increased uncertainty. Second, in this simulation, the environment was complex, namely different paths were not much different to each other in terms of total reward. Indeed, going left at S1 led to r = 3, whereas going right led to r = 5, which are relatively close to each other. These results suggest that in complex and uncertain environments the forward sweeps could persist for a longer time, and the passage from goal-directed to habitual strategies could be incomplete.

4. GENERAL DISCUSSION

The Mixed Instrumental Controller (MIC) is an integrative model describing how model-based (mental simulation) and model-free mechanisms (Q learning) could interact in both cooperative and competitive ways, producing a continuum of habitual and goal-directed strategies of choice.

In the Mixed Instrumental Controller, model-free mechanisms are used by default and supported by model-based computations when the Value of Information of the latter surpasses its costs; this is typically true when uncertainty is high and alternative cached action values are close to each other. Furthermore, the relative contribution of model-based mechanisms can vary: the less the uncertainty, the fewer the samples used to implement the forward sweeps. In sum, the MIC permits to flexibly balance model-free and model-based methods depending on environmental circumstances.

To decide when mental simulation is necessary, the Mixed Instrumental Controller solves a "dilemma" that is similar to the well known *exploration-exploitation dilemma*, except that in this context the exploration is "mental" and not overt. Specifically, the *mental exploration* consists in performing mental simulations to access expectancies and associated reward predictions, and ultimately to better estimate action values. The *exploitation* consists in choosing an action on the basis of the already available ("cached") estimate of action values, rather than performing mental simulation. The dilemma can be solved by comparing the Value of Information that can be retrieved using mental simulation with the cost of the simulation. Computing an optimal solution to this problem is generally intractable in non-stationary environments, and it is still unclear if and how the brain does so (Aston-Jones and Cohen, 2005; Daw et al., 2006; Pezzulo and Couyoumdjian, 2006; Behrens et al., 2007; Cohen et al., 2007; Bromberg-Martin and Hikosaka, 2011; Niv and Chan, 2011). The MIC model implements an approximate solution to this problem that considers accuracy of choice (i.e., probability of acquiring higher reward) and uses a fixed cost of acquiring information (in terms of cognitive effort and time); the former factor favoring mental exploration, and the latter exploitation. Overt exploration is not explicitly modeled in the MIC, but it results from the adoption of a softmax function for the choice.

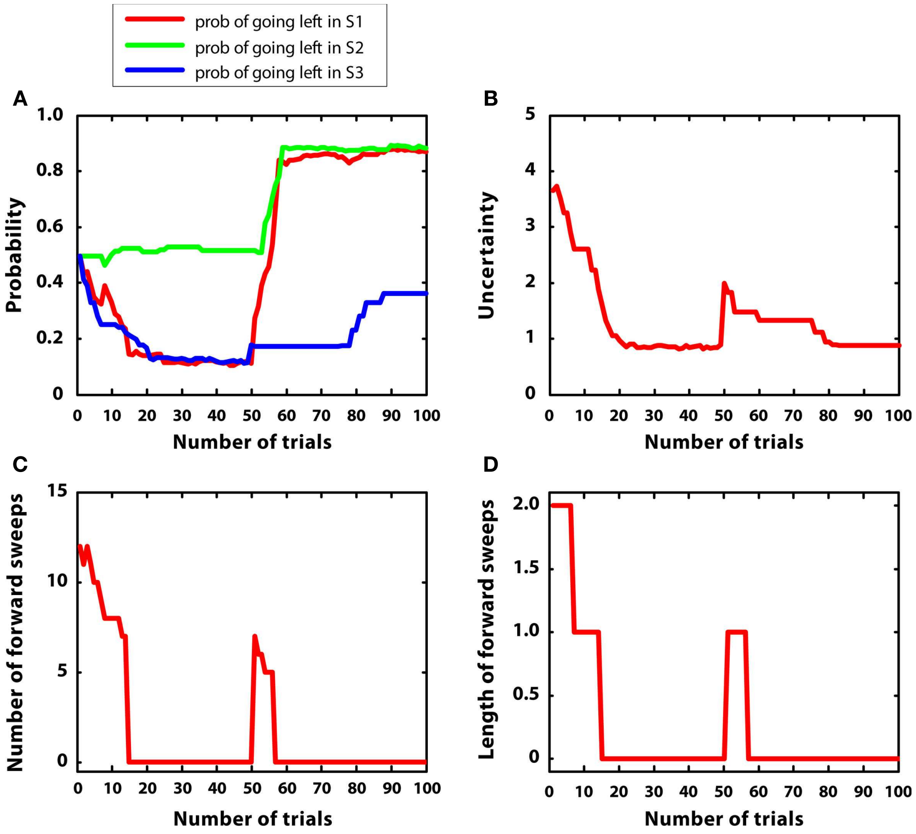

FIGURE 7 | Results of simulation 3, non-stationary environment. (A–D): see **Figure 5**.

Our simulations in environments having different characteristics (stable or volatile, low or high variance) show that there are multiple factors that can cause the Value of Information to be higher, and most notably the variance and the difference in value between the competing alternatives. Generally, mental simulations at decision points diminish after sufficient learning, in line with evidence showing that in this condition habitization replaces goal-directed mechanisms of choice (Jog et al., 1999). However, if variance is high or if the values of the alternatives are too close, the system is slower in developing habits. Different from alternative models, in the MIC the habitization is accompanied by a reduced use of model-based computations; this mechanism can explain why hippocampal-striatal forward sweeps, possibly encoding covert simulations at decision points, vanish after sufficient experience (van der Meer and Redish, 2009).

When environmental contingencies change, mental simulations are used anew, consistent with evidence of a passage from habitual to goal-directed strategies after outcome devaluation (unless it occurs after "overtraining"). When contingencies change, the goal-directed system can immediately change behavior. Furthermore, changed environmental conditions increase the VoI and speeds up the updating of C and Q values; see **Figure 4**. However, the reaction to outcome devaluation can be slower (or impaired) when actions are over-trained (Dickinson, 1985) because the (non-active) goal-directed system cannot instruct an immediate change of strategy and updating C and Q values takes longer.

It is worth noting that although the MIC model is sensible to the volatility of the environment, this element is not explicitly modeled (but, see below for a possible extension of the model). Finally, our results in the more complex environments (with high uncertainty and variance) are in keeping with evidence that forward sweeps are not a simple replay of previous experience but are modulated by task uncertainty (Gupta et al., 2010). Moreover, the MIC model makes the further prediction that the difference between alternative cached action values has a role as well in influencing forward sweeps.

5. CONCLUSION

We proposed that essential aspects of goal-directed and habitual control can be captured within a single instrumental process of decision-making, the Mixed Instrumental Controller (MIC), which flexibly balances and combines model-based and model-free computations. We linked the functioning of the MIC model to a neural circuit formed by the hippocampus and the ventral striatum, which has been shown to be active during goal-directed navigation and the choice between spatially defined goals.

The MIC model elaborates on a previous influential model (Daw et al., 2005; Niv et al., 2006; Dayan, 2009) which emphasizes that goal-directed and habitual mechanisms of choice are linked

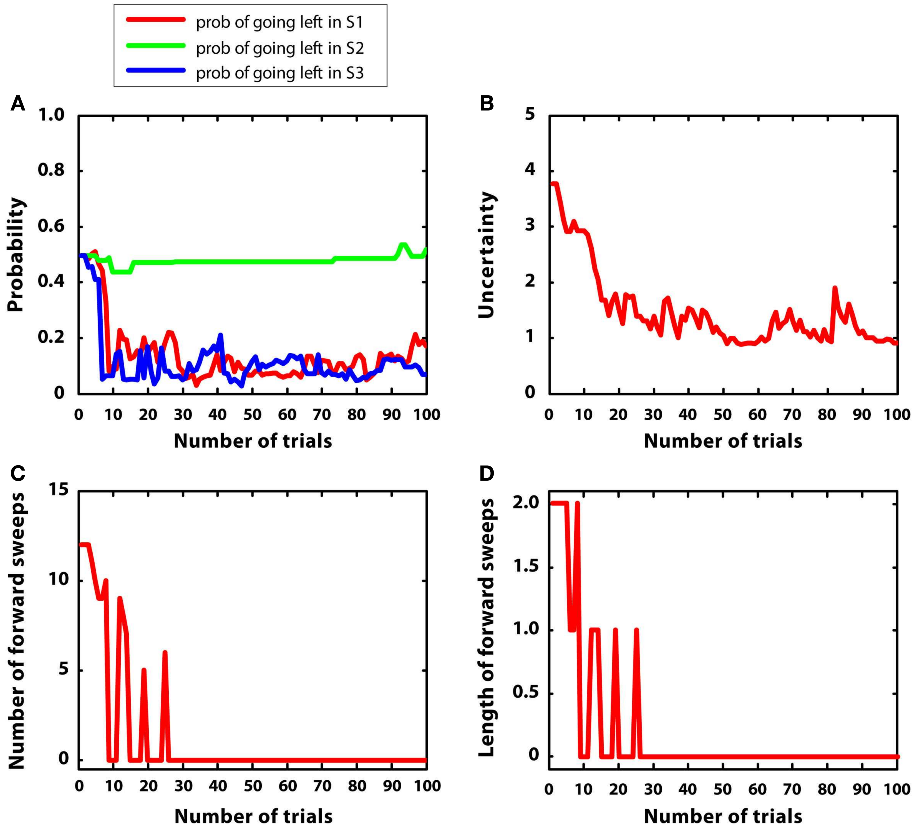

FIGURE 8 | Results of simulation 4, simple environment with high variance. (A–D): see **Figure 5**.

to model-based and model-free methods of reinforcement learning, respectively, and which assigns a key role to uncertainty. At the same time, the MIC departs from this model in that it assumes that model-based calculations are only used when the Value of Information they can furnish is higher than their costs. Another distinguishing point is the fact that while in the previous theory model-free and model-based processes produce two competing instrumental controllers, in the MIC they act in concert. First, although generally mental simulations are used to retrieve the rewards associated to future states, they can also retrieve Q values that permit to aggregate the value of several time steps, as it is typical of model-free algorithms. Combining these two methods (for instance, performing forward search until a reliable cached value is available) is typical in game playing set-ups (Baum and Smith, 1997) and understanding how the brain might do so is an important avenue for future research (Glascher et al., 2010; Simon and Daw, 2011a). Second, model-free and model-based processes provide complementary information to calculate action values. This is evident if one considers that, in equation (2), the cached Q value is used as a prior and updated using model-based calculations.

Another peculiarity of our model is the way mental simulation is realized. In the MIC, mental simulation is computationally implemented as a *sequential sampling* procedure using the graphical model described in **Figure 3**. The method we adopt consists in "clamping" one policy at a time (see Solway and Botvinick, 2012), which produces a serial process of (simulated) internal experience sampling. This method is different from the idea of a "tree search" as it is typically described in normative approaches (Niv et al., 2006), and from models of parallel "diffusion" processes for planning (Ivey et al., 2011). It produces a serial forward search that better captures the nature of forward sweeps in the hippocampus (see also Lengyel and Dayan, 2008; Bornstein and Daw, 2011). Furthermore, the specific algorithm used for the forward search, i.e., particle filtering, produces a (noisy) accumulation of evidence about rewards, which links well to sequential sampling dynamics used for perceptual decisions and memory search (Ratcliff, 1978) and the "ramping" activity of primate neurons during choice (typically, in the neuronal areas that control the effectors used for the choice; Shadlen and Newsome, 2001; Cisek and Kalaska, 2005; Ding and Gold, 2010). Overall, then, our mental simulation system describes the value-based computations of the hippocampus – ventral striatum circuit in terms that are analogous to those of perceptual-based decisions, and are coherent with the idea of "decision by sampling" (Stewart et al., 2006).

All these characteristics distinguish the MIC from the model of Daw et al. (2005) and from several others, which we shortly review below. Similar to the MIC, it has been recently proposed that model-based computations are activated only when the Value of Information they add is bigger than the cost of waiting they entail (Keramati et al., 2011). Similar to the MIC model, the Value

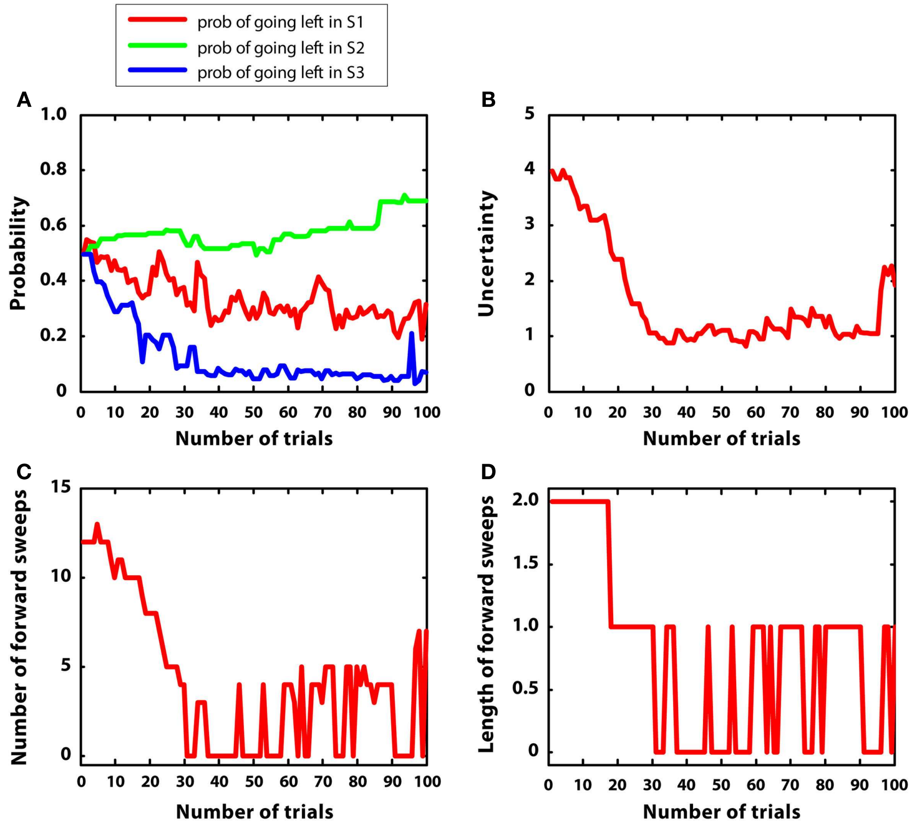

FIGURE 9 | Results of simulation 5, complex environment with high variance. (A–D): see **Figure 5**.

of Information is computed by considering the uncertainty and the distance between alternative action values; however, different from the MIC model the model-based component is expected to have perfect information. The major difference between the model of Keramati et al. (2011) and the MIC is that how model-based computations are performed and used. Indeed, the former model shifts completely from habitual to goal-directed control when the Value of Information is sufficiently high. Conversely, the latter model integrates "cached" values and model-based estimation, and thus results in a "mixed" control. In addition, in the MIC model model-based computations are performed using a serial sampling process; the samples vary in number and length and model-based computations can be performed only for a sub-set of available actions. These features have been adopted to fit better with the evidence available on rats' forward sweeps, which are thought to correspond to model-based computations.

The aforementioned models (Daw et al., 2005; Keramati et al., 2011) and others (Simon and Daw, 2011b) assume that model-based and model-free methods can only compete, not cooperate. The DYNA model is one of the few systems in which model-based and model-free methods cooperate (Sutton, 1990). In DYNA, only the habitual system is responsible for making decisions, but the goal-directed system can train it by providing off-line predictions. A recent study uses the DYNA system to explain the shift between habitual and goal-directed systems and retrospective revaluation (Gershman et al., 2012). In the MIC model mental simulations are used on-line, during the choice, and are responsible for the forward sweeps in the hippocampus at decision points. Below we discuss a straightforward extension of the MIC model that uses mental simulations both on-line and off-line.

An alternative view of the memory consolidation process is that it consists in a *chunking* of action sequences. In this view, model-free methods are not used: all actions are first executed in a model-based way and then gradually chunked and transformed into habits (Dezfouli and Balleine, 2012). Different from this theory, the MIC uses both model-free and model-based methods, and describes the transition from goal-directed to habitual behavior in terms of changed Value of Information rather than chunking.

5.1. FUTURE IMPROVEMENTS OF THE MIC MODEL

There are several aspects of the MIC model that can be further elaborated. First, the MIC currently uses simplified methods to calculate Value of Information and the costs of simulation. The method we devised has several limitations; for instance, it does not consider the absolute value of the actions but only their relative values, and only uses a fixed threshold. The current formal analyses of Value of Information take some of these aspects into consideration but are computationally impractical; furthermore, it is unclear how they link to neural computations (Howard, 1966). As our knowledge of how the brain addresses these problems increases, better methods can be devised that permit to quantify the costs and benefits of mental exploration, and to realize a better cost-benefits analysis.

The proposed model can be easily extended by permitting the model-based part to train the model-free part off-line and in absence of overt behavior, similar to other RL algorithms such as DYNA (Sutton, 1990) and prioritized sweeping (Moore and Atkeson, 1993). The values of C and Q can be updated even when the agent is not acting by endogenously steering mental simulations to produce "fake" reward observations O, and then using the same learning methods as described in sec. 2.4. With this straightforward extension the MIC can benefit from both on-line and off-line mental simulations using the same mechanisms. We chose not to use off-line mental simulations in our experiments because in the scenarios we simulated there could be too little time to complete the off-line training within experimental trials (otherwise we would never observe forward sweeps at decision points). Rather, we hypothesize that off-line training could have a more prominent role when there is enough time for memory consolidation (e.g., during pauses and sleep, but also when there is enough time between experimental trials). In the proposed "extended" version of the MIC model, mental simulations support both decision-making (when used on-line) and memory consolidation (when used off-line). Indeed, there are various demonstrations that the rat hippocampus replays (forward and backward) sequences of neural activity experienced during overt behavior both when the animal pauses (and is awake) and when it is asleep (Foster and Wilson, 2006; Diba and Buzski, 2007; Koene and Hasselmo, 2008; Peyrache et al., 2009; Gupta et al., 2010; Carr et al., 2011); still the behavioral significance of these findings is disputed. Some studies emphasize the importance of forward sweeps for decision-making (van der Meer and Redish, 2009), while other studies highlight the consolidation of recent memories into long-term memory and the formation of "cognitive maps" of the environment (Tolman, 1948; O'Keefe and Dostrovsky, 1971; Morris et al., 1982). We hypothesize that these apparently distinct views can be reconciled if one considers the aforementioned distinction between on-line and off-line uses of mental simulations in the MIC model. It is worth noting that the precise mechanisms regulating off-line mental simulations remain to be established. Off-line training could be regulated by similar principles of optimization as in the meta-choice we described. For example, the agent could simulate being at a decision point, decide whether or not to activate the model-based component using the Value of Information computations of equation (1), and use the particle filtering algorithm of sec. 2.4 for training the habitual system. Alternatively, it could eschew the Value of Information computations and only consider the accuracy of the habitual system (e.g., the variance of Q values) or more simply try to systematically update all the Q values. The plausibility of these and other hypotheses remains to be established.

The proposed model can also be improved by explicitly modeling environmental volatility. The MIC is implicitly sensible to volatility and changed reward contingencies. However, it is plausible that living organisms explicitly model volatility (Behrens et al., 2007; Kepecs et al., 2008). In turn, an estimate of volatility permits to better regulate the Value of Information (as in volatile environment uncertainty cannot be reduced using mental simulation), to adjust learning rates adaptively, and to modulate the rate of overt exploration (which is at the moment sidestepped using a parameterized *softmax* function). A related issue is considering the quality of the internal model and the controllability of the environment when choosing a controller; computational modeling studies suggest that it might be favorable to select closed-loop methods in well-modeled regions and open-loop methods in regions that are not (or cannot) be modeled with high accuracy (Kolter et al., 2010).

Another important direction for future studies is devising biologically plausible and scalable algorithms to implement the proposed model-based computations. At the moment, model-based methods are computationally prohibitive for large state spaces, but progresses on sampling methods (Doucet et al., 2000) and Monte Carlo search (Silver and Veness, 2010) are encouraging. Not only these methods are interesting from a computational viewpoint, but they could also shed light on how mental simulations and forward planning are mechanistically implemented in the brain, as suggested by recent studies that link brain activity with probabilistic computations (Ma et al., 2006; Doya et al., 2007) and sampling methods (Fiser et al., 2010; Berkes et al., 2011).

Furthermore, the MIC uses model-based computations and mental simulations for action selection and learning, but it leaves unspecified if they can be also used for other purposes. An intriguing proposal is that mental simulations can be used to monitor actions initiated by the habitual system until their successful completion (Alexander and Brown, 2011). This would permit a rapid initiation of action, and also its subsequent revision if mental simulation uncovers negative consequences that the habitual system did not take into consideration. It is worth noting that this mechanism could be another way how model-free and model-based methods cooperate.

We have linked the model-based computations of the MIC to a neural circuit formed by the hippocampus and the ventral striatum. The reason for our choice is that this circuit has been linked to goal-directed computations in spatial navigation (i.e., the scenario that we chose to exemplify the characteristics of the MIC). However, it is plausible that the brain uses additional (or different) neuronal circuits for model-based computations outside the spatial domain. We hypothesize that the MIC captures essential principles of instrumental control that are not restricted to goal-directed spatial navigation; however, understanding if the model-based computations of the MIC apply to instrumental choice at large remains an open objective for future research.

A further aspect to consider is how the MIC architecture could potentially include Pavlovian mechanisms. In relation to this, two possibilities should be considered. Pavlovian processes might substantially act in parallel with instrumental ones. Alternatively, Pavlovian and instrumental representations might largely overlap. Although contrasting findings have been reported, evidence suggests that Pavlovian and goal-directed values are segregated functionally and neurally. For instance, following devaluation, Pavlovian effects, contrary to goal-directed ones, are visible even without incentive learning. Moreover, lesions of different portions of amygdala, ventromedial prefrontal cortex, and striatum, have differential impact on Pavlovian and goal-directed mechanisms (Balleine and O'Doherty, 2009). Overall, this evidence suggests that Pavlovian and instrumental mechanisms work in parallel (see also Rigoli et al., 2012), and future implementations of the MIC should consider this fact.

Finally, the MIC model is currently limited in that it only considers one level of granularity of actions and states. In contrast, the control of behavior has been recently linked to hierarchical reinforcement learning models (Botvinick, 2008; Botvinick et al., 2009; Frank and Badre, 2012), in which actions can be specified at different levels of abstractness and temporal extension (see also Verschure et al., 2003). Extending the MIC with hierarchical action organization would provide extra flexibility, allowing it, for example, to select and plan actions at more abstract levels, and to connect with the growing literature on prefrontal control hierarchies (Fuster, 1997; Koechlin and Summerfield, 2007; Wise, 2008).

5.2. REAL-TIME DYNAMICS AND PUTATIVE NEURONAL ARCHITECTURE OF THE MIXED INSTRUMENTAL CONTROLLER

The MIC model offers a computational-level explanation of the interactions between habitual and goal-directed processes of choice in the context of spatial navigation. While the real-time dynamics of mental simulation are explicitly modeled using the particle filtering algorithm, the moment-by-moment dynamics of the action selection process are sidestepped using the process model described in **Figure 2**. Below we discuss how the MIC model could implement real-time dynamics of choice through a neural architecture.

We take as our starting point the *affordance competition hypothesis* (Cisek and Kalaska, 2010): a parallel model of decision-making that describes choice as a dynamic competition between two (or more) action alternatives (say, go left or right). In the affordance competition hypothesis, multiple plans for action are formed in parallel and compete over time until one has sufficient support to win the competition. In terms of the MIC, the default habitual processes (plausibly including mappings between stimuli and motor representations) mediate this selection by instructing previously reinforced stimulus-response associations. In this architecture, response dynamics correspond to the activity of neuronal populations in frontoparietal cortex, forming a sort of motor map for the potential responses (Cisek, 2006), whose selection is plausibly supported by the basal ganglia (Redgrave et al., 1999; Chersi et al., 2012; Lepora and Gurney, 2012). In the context of spatial navigation and the choice between spatially defined goals, the hippocampus is also involved to support (among the other things) spatial representation and processing. During the choice, the presence of an appropriate stimulus (say, the sight of a branch of the T-maze) could produce a strong peak of activation in the motor map in correspondence of the to-be-selected action. However, this is only effective when the stimulus-response associations are strong enough (e.g., after habitization). When the potential action plans have little support (e.g., before sufficient learning), or when the choice is highly uncertain, the motor map could encode several low-intensity and high variance peaks of activation. In these cases, cognitive control and monitoring mechanisms could inhibit action execution and allow for more information to be collected via model-based computations, until confidence is high or the costs of acquiring it surpasses the benefits.

In the MIC, there is not a univocal value representation, but different aspects of valuation correspond to different parts of the model; this is consistent with recent theories that recognize the contribution of different brain areas to utility representation and processing (Ito and Doya, 2011; Pennartz et al., 2011). State values (and reward expectancies) could be associated to ventral striatum (Lansink et al., 2009; van der Meer and Redish, 2009), ventrotegmental area, basolateral amygdala, and orbitofrontal cortex (Padoa-Schioppa and Assad, 2006; Yin et al., 2008; McDannald et al., 2012). In our model, state values correspond to $S \rightarrow R$ transitions; in a previous work we have also shown how these values can be modulated by the agent's internal motivational state (Pezzulo and Rigoli, 2011). Dorsolateral striatum could encode cached action values and could have a role in encoding uncertainty (Yin et al., 2004; Kepecs et al., 2008; represented in our model by Q and C, respectively). It is worth noting that although the dorsal/ventral division of the striatum (which we also re-propose here) has been associated to segregated habitual and goal-directed controllers, respectively, our model does not necessarily imply a complete segregation, but is compatible with the view that the controllers could partially overlap. The mapping of specific parts of the striatum with different computations (model-based and model-free) and modes of control (goal-directed and habitual) is still controversial (see Bornstein and Daw, 2011).

The MIC is consistent with the idea that the ventral striatum supports model-based reward representations (activated during forward sweeps), as suggested by van der Meer and Redish (2010). This idea is distinct from the standard view that the ventral striatum plays the role of "critic" in actor-critic RL theories, and is recruited exclusively during learning (Houk et al., 1995). However, the MIC is consistent also with an alternative possibility, coherent with the ventral striatal role as "critic." It is indeed possible that this structure encodes the "fictive" prediction error which, in the MIC model, is used to update prior Q "cached" values with pseudo-observations produced by mental simulation. This hypothesis generates the specific prediction that the signal in ventral striatum correlates with the "fictive" prediction error (i.e., with the discrepancy between "cached" and goal-directed values) rather than with goal-directed values. By using devaluation, for example, it could be possible to test these alternative hypotheses.

Another aspect of the MIC model is relative to the meta-choice, the calculation of the Value of Information, and the cognitive control of the computations. A relatively simple form of cognitive control has been linked to *optimal stopping* problems, in which it is necessary to consider the confidence of actions and the cost to be late before taking an action (Gold and Shadlen, 2001, 2007). It has been argued that optimal stopping and more sophisticated forms of meta-choice could be based on mechanisms for monitoring, uncertainty consideration and behavioral inhibition. With these mechanisms, the architectures for action specification and selection described before can become able of goal-directed choice and cognitive control, consistent with the view that these more advanced abilities could derive from elaborations of brain designs that solve simpler sensorimotor processes (Pezzulo, 2008, 2011; Pezzulo and Castelfranchi, 2009; Cisek and Kalaska, 2010; Cisek, 2012). In the MIC, these mechanisms could improve the choice by permitting model-based mechanisms to support or even substitute the default habitual control mode. In the current implementation, this is done by mentally simulating and collecting covert expectations of reward and goals, but anatomical considerations point also to more sophisticated mechanisms such as mental time travel and the construction of novel episodic memories (Schacter et al., 2007, 2012; Buckner, 2010). Although the neural underpinnings of the control architecture are incompletely known, we speculate that monitoring processes in the anterior cingulate cortex could signal the opportunity to overcome stimulus-bound responses (Botvinick et al., 2001; Alexander and Brown, 2011), the Value of Information computations could reuse cached action, and uncertainty values, and the passage from stimulus-bound to internally generated (simulated) contexts necessary for the model-based computations could be linked to rostral prefrontal cortex (Burgess et al., 2007).

These and other aspects of brain implementations of goal-directedness remain open objectives for future research. Indeed, our study is part of a large initiative investigating model-based decision-making in the brain (Balleine and Dickinson, 1998; Daw et al., 2005; Dayan, 2009; Green et al., 2010; Rao, 2010; Daw, 2012; Pezzulo and Rigoli, 2011; Simon and Daw, 2011b; Solway and Botvinick, 2012). Model-free RL methods have provided useful insights to study the neural neurobiology of action values and habitual behavior. Analogously, model-based RL mechanisms could help studying the neural underpinnings of mental simulations, outcome predictions, and goal-directed choice (O'Doherty, 2012). It is important to consider that there are many possible variants of model-based RL methods (as there are multiple forms of model-free RL computations), possibly linking to different neural substrates (Daw, 2012). So, it remains to be evaluated what computational proposals better capture the brain's ability to flexibly choose and act in a goal-directed manner.

ACKNOWLEDGMENTS

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement 270108 (Goal-Leaders). We thank Andrew Barto, Matthew Botvinick, and Peter Dayan for useful discussions.

REFERENCES

Alexander, W. H., and Brown, J. W. (2011). Medial prefrontal cortex as an action-outcome predictor. *Nat. Neurosci.* 14, 1338–1344.

Aston-Jones, G., and Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. *Annu. Rev. Neurosci.* 28, 403–450.

Balleine, B., and O'Doherty, J. (2009). Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology* 35, 48–69.

Balleine, B. W., and Dickinson, A. (1998). Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology* 37, 407–419.

Battaglia, F. P., Benchenane, K., Sirota, A., Pennartz, C. M. A., and Wiener, S. I. (2011). The hippocampus: hub of brain network communication for memory. *Trends Cogn. Sci. (Regul. Ed.)* 15, 310–318.

Baum, E. B., and Smith, W. D. (1997). A Bayesian approach to relevance in game playing. *Artif. Intell.* 97, 195–242.

Behrens, T. E. J., Woolrich, M. W., Walton, M. E., and Rushworth, M. F. S. (2007). Learning the value of information in an uncertain world. *Nat. Neurosci.* 10, 1214–1221.

Berkes, P., Orbn, G., Lengyel, M., and Fiser, J. (2011). Spontaneous cortical activity reveals hallmarks of an optimal internal model of the environment. *Science* 331, 83–87.

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning.* New York: Springer.

Bornstein, A. M., and Daw, N. D. (2011). Multiplicity of control in the basal ganglia: computational roles of striatal subregions. *Curr. Opin. Neurobiol.* 21, 374–380.

Botvinick, M., Niv, Y., and Barto, A. (2009). Hierarchically organized behavior and its neural foundations: a reinforcement learning perspective. *Cognition* 113, 262–280.

Botvinick, M. M. (2008). Hierarchical models of behavior and prefrontal function. *Trends Cogn. Sci. (Regul. Ed.)* 12, 201–208.

Botvinick, M. M., and An, J. (2008). "Goal-directed decision making in prefrontal cortex: a computational framework," in *Advances in Neural Information Processing Systems (NIPS)*, Lake Tahoe.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., and Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychol. Rev.* 108, 624–652.

Bromberg-Martin, E. S., and Hikosaka, O. (2011). Lateral habenula neurons signal errors in the prediction of reward information. *Nat. Neurosci.* 14, 1209–1216.

Buckner, R. L. (2010). The role of the hippocampus in prediction and imagination. *Annu. Rev. Psychol.* 61, 27–48, C1–C8.

Burgess, P. W., Dumontheil, I., and Gilbert, S. J. (2007). The gateway hypothesis of rostral prefrontal cortex (area 10) function. *Trends Cogn. Sci. (Regul. Ed.)* 11, 290–298.

Carr, M., Jadhav, S., and Frank, L. (2011). Hippocampal replay in the awake state: a potential substrate for memory consolidation and retrieval. *Nat. Neurosci.* 14, 147–153.

Chersi, F., Mirolli, M., Pezzulo, G., and Baldassarre, G. (2012). A spiking neuron model of the cortico-basal ganglia circuits for goal-directed and habitual action learning. *Neural Netw.* doi:10.1016/j.neunet.2012.11.009. [Epub ahead of print].

Chersi, F., and Pezzulo, G. (2012). Using hippocampal-striatal loops for spatial navigation and goal-directed decision-making. *Cogn. Process.* 13, 125–129.

Cisek, P. (2006). Integrated neural processes for defining potential actions and deciding between them: a computational model. *J. Neurosci.* 26, 9761–9770.

Cisek, P. (2012). Making decisions through a distributed consensus. *Curr. Opin. Neurobiol.* 22, 927–936.

Cisek, P., and Kalaska, J. F. (2005). Neural correlates of reaching decisions in dorsal premotor cortex: specification of multiple direction choices and final selection of action. *Neuron* 45, 801–814.

Cisek, P., and Kalaska, J. F. (2010). Neural mechanisms for interacting with a world full of action choices. *Annu. Rev. Neurosci.* 33, 269–298.

Cohen, J. D., McClure, S. M., and Yu, A. J. (2007). Should i stay or should i go? how the human brain manages the trade-off between exploitation and exploration. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 362, 933–942.

Daw, N. (2012). "Model-based reinforcement learning as cognitive search: neurocomputational theories," in *Cognitive Search: Evolution, Algorithms and the Brain*, eds P. Todd and T. Robbins (MIT Press).

Daw, N. D., Niv, Y., and Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat. Neurosci.* 8, 1704–1711.

Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B., and Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature* 441, 876–879.

Dayan, P. (2009). Goal-directed control and its antipodes. *Neural Netw.* 22, 213–219.

Dezfouli, A., and Balleine, B. W. (2012). Habits, action sequences and reinforcement learning. *Eur. J. Neurosci.* 35, 1036–1051.

Diba, K., and Buzski, G. (2007). Forward and reverse hippocampal place-cell sequences during ripples. *Nat. Neurosci.* 10, 1241–1242.

Dickinson, A. (1985). Actions and habits: the development of behavioural autonomy. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 308, 67–78.

Dindo, H., Zambuto, D., and Pezzulo, G. (2011). "Motor simulation via coupled internal models using sequential Monte Carlo," in *Proceedings of IJCAI 2011*, Barcelona, 2113–2119.

Ding, L., and Gold, J. I. (2010). Caudate encodes multiple computations for perceptual decisions. *J. Neurosci.* 30, 15747–15759.

Doucet, A., Godsill, S., and Andrieu, C. (2000). On sequential Monte Carlo sampling methods for Bayesian filtering. *Stat. Comput.* 10, 197–208.

Doya, K., Ishii, S., Pouget, A., and Rao, R. P. N. (eds). (2007). *Bayesian Brain: Probabilistic Approaches to Neural Coding*, 1st Edn. Cambridge: MIT Press.

Erdem, U. M., and Hasselmo, M. (2012). A goal-directed spatial navigation model using forward trajectory planning based on grid cells. *Eur. J. Neurosci.* 35, 916–931.

Fiser, J., Berkes, P., Orbn, G., and Lengyel, M. (2010). Statistically optimal perception and learning: from behavior to neural representations. *Trends Cogn. Sci. (Regul. Ed.)* 14, 119–130.

Foster, D., and Wilson, M. (2006). Reverse replay of behavioural sequences in hippocampal place cells during the awake state. *Nature* 440, 680–683.

Frank, M. J., and Badre, D. (2012). Mechanisms of hierarchical reinforcement learning in corticostriatal circuits 1: computational analysis. *Cereb. Cortex* 22, 509–526.

Fuster, J. M. (1997). *The Prefrontal Cortex: Anatomy, Physiology, and Neuropsychology of the Frontal Lobe.* Philadelphia, PA: Lippincott-Raven.

Gershman, S., and Daw, N. (2011). "Perception, action and utility: the tangled skein," in *Principles of Brain Dynamics: Global State Interactions*, eds M. Rabinovich, M. K. Friston, and P. Varona (MIT Press).

Gershman, S., Markman, A., and Otto, A. (2012). Retrospective revaluation in sequential decision making: a tale of two systems. *J. Exp. Psychol. Gen.* PMID:23230992. [Epub ahead of print].

Glascher, J., Daw, N., Dayan, P., and O'Doherty, J. P. (2010). States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron* 66, 585–595.

Glimcher, P., Fehr, E., Camerer, C., and Poldrack, R. (eds). (2009). *Neuroeconomics: Decision Making and the Brain.* New York: Academic Press.

Gold, J., and Shadlen, M. (2001). Neural computations that underlie decisions about sensory stimuli. *Trends Cogn. Sci. (Regul. Ed.)* 5, 10–16.

Gold, J. I., and Shadlen, M. N. (2007). The neural basis of decision making. *Annu. Rev. Neurosci.* 30, 535–574.

Green, C. S., Benson, C., Kersten, D., and Schrater, P. (2010). Alterations in choice behavior by manipulations of world model. *Proc. Natl. Acad. Sci. U.S.A.* 107, 16401–16406.

Gupta, A. S., van der Meer, M. A. A., Touretzky, D. S., and Redish, A. D. (2010). Hippocampal replay is not a simple function of experience. *Neuron* 65, 695–705.

Houk, J. C., Adams, J. L., and Barto, A. G. (1995). "A model of how the basal ganglia generates and uses neural signals that predict reinforcement," in *Models of Information Processing in the Basal Ganglia*, eds J. C. Houk, J. Davis, and D. Beiser (Cambridge: MIT Press), 249–270.

Howard, R. (1966). Information value theory. *IEEE Trans. Syst. Sci. Cybern.* 2, 22–26.

Ito, M., and Doya, K. (2011). Multiple representations and algorithms for reinforcement learning in the cortico-basal ganglia circuit. *Curr. Opin. Neurobiol.* 21, 368–373.

Ivey, R., Bullock, D., and Grossberg, S. (2011). A neuromorphic model of spatial lookahead planning. *Neural Netw.* 24, 257–266.

Jog, M. S., Kubota, Y., Connolly, C. I., Hillegaart, V., and Graybiel, A. M. (1999). Building neural representations of habits. *Science* 286, 1745–1749.

Johnson, A., and Redish, A. D. (2007). Neural ensembles in ca3 transiently encode paths forward of the animal at a decision point. *J. Neurosci.* 27, 12176–12189.

Kepecs, A., Uchida, N., Zariwala, H. A., and Mainen, Z. F. (2008). Neural correlates, computation and behavioural impact of decision confidence. *Nature* 455, 227–231.

Keramati, M., Dezfouli, A., and Piray, P. (2011). Speed/accuracy trade-off between the habitual and the goal-directed processes. *PLoS Comput. Biol.* 7:e1002055. doi:10.1371/journal.pcbi.1002055

Koechlin, E., and Summerfield, C. (2007). An information theoretical approach to prefrontal executive function. *Trends Cogn. Sci. (Regul. Ed.)* 11, 229–235.

Koene, R. A., and Hasselmo, M. E. (2008). Reversed and forward buffering of behavioral spike sequences enables retrospective and prospective retrieval in hippocampal regions ca3 and ca1. *Neural Netw.* 21, 276–288.

Kolter, J. Z., Plagemann, C., Jackson, D. T., Ng, A. Y., and Thrun, S. (2010). "A probabilistic approach to mixed open-loop and closed-loop control, with application to extreme autonomous driving," in *Proceedings of the International Conference on Robotics and Automation*, Anchorage.

Lansink, C. S., Goltstein, P. M., Lankelma, J. V., McNaughton, B. L., and Pennartz, C. M. A. (2009). Hippocampus leads ventral striatum in replay of place-reward information. *PLoS Biol.* 7:e1000173. doi:10.1371/journal.pbio.1000173

Lengyel, M., and Dayan, P. (2008). "Hippocampal contributions to control: the third way," in *Advances in Neural Information Processing Systems*, Vol. 20, eds J. Platt, D. Koller, Y. Singer, and S. Roweis (Cambridge, MA: MIT Press), 889–896.

Lepora, N. F., and Gurney, K. N. (2012). The basal ganglia optimize decision making over general perceptual hypotheses. *Neural Comput.* 24, 2924–2945.

Ma, W. J., Beck, J. M., Latham, P. E., and Pouget, A. (2006). Bayesian inference with probabilistic population codes. *Nat. Neurosci.* 9, 1432–1438.

McDannald, M. A., Takahashi, Y. K., Lopatina, N., Pietras, B. W., Jones, J. L., and Schoenbaum, G. (2012). Model-based learning and the contribution of the orbitofrontal cortex to the model-free world. *Eur. J. Neurosci.* 35, 991–996.

Moore, A. W., and Atkeson, C. (1993). Prioritized sweeping: reinforcement learning with less data and less real time. *Mach. Learn.* 13, 103–130.

Morris, R. G., Garrud, P., Rawlins, J. N., and O'Keefe, J. (1982). Place navigation impaired in rats with hippocampal lesions. *Nature* 297, 681–683.

Murphy, K. P. (2002). *Dynamic Bayesian Networks: Representation, Inference and Learning.* Ph.D. thesis, Computer Science Division, University of California Berkeley, Berkeley.

Niv, Y., and Chan, S. (2011). On the value of information and other rewards. *Nat. Neurosci.* 14, 1095Ð–1097D.

Niv, Y., Joel, D., and Dayan, P. (2006). A normative perspective on motivation. *Trends Cogn. Sci. (Regul. Ed.)* 8, 375–381.

O'Doherty, J. P. (2012). Beyond simple reinforcement learning: the computational neurobiology of reward-learning and valuation. *Eur. J. Neurosci.* 35, 987–990.

O'Keefe, J., and Dostrovsky, J. (1971). The hippocampus as a spatial map. preliminary evidence from unit activity in the freely-moving rat. *Brain Res.* 34, 171–175.

Padoa-Schioppa, C., and Assad, J. A. (2006). Neurons in the orbitofrontal cortex encode economic value. *Nature* 441, 223–226.

Pennartz, C. M. A., Ito, R., Verschure, P. F. M. J., Battaglia, F. P., and Robbins, T. W. (2011). The hippocampal-striatal axis in learning, prediction and goal-directed behavior. *Trends Neurosci.* 34, 548–559.

Penner, M. R., and Mizumori, S. J. Y. (2012). Neural systems analysis of decision making during goal-directed navigation. *Prog. Neurobiol.* 96, 96–135.

Peyrache, A., Khamassi, M., Benchenane, K. and Wiener, S., and Battaglia, F. (2009). Replay of rule-learning related neural patterns in the prefrontal cortex during sleep. *Nat. Neurosci.* 12, 12, 919–926.

Pezzulo, G. (2008). Coordinating with the future: the anticipatory nature of representation. *Minds Mach.* 18, 179–225.

Pezzulo, G. (2011). Grounding procedural and declarative knowledge in sensorimotor anticipation. *Mind Lang.* 26, 78–114.

Pezzulo, G., and Castelfranchi, C. (2009). Thinking as the control of imagination: a conceptual framework for goal-directed systems. *Psychol. Res.* 73, 559–577.

Pezzulo, G., and Couyoumdjian, A. (2006). "Ambiguity-reduction: a satisficing criterion for decision making," in *Proceedings of the 28th Annual Meeting of Cognitive Science Society*, Vancouver, 669–674.

Pezzulo, G., Hoffmann, J., and Falcone, R. (2007). Anticipation and anticipatory behavior. *Cogn. Process.* 8, 67–70.

Pezzulo, G., and Rigoli, F. (2011). The value of foresight: how prospection affects decision-making. *Front. Neurosci.* 5:79. doi:10.3389/fnins.2011.00079

Rao, R. P. N. (2010). Decision making under uncertainty: a neural model based on partially observable Markov decision processes. *Front. Comput. Neurosci.* 4:146. doi:10.3389/fncom.2010.00146

Ratcliff, R. (1978). A theory of memory retrieval. *Psychol. Rev.* 85, 59–108.

Redgrave, P., Prescott, T. J., and Gurney, K. (1999). The basal ganglia: a vertebrate solution to the selection problem? *Neuroscience* 89, 1009–1023.

Rigoli, F., Pavone, E. F., and Pezzulo, G. (2012). Aversive pavlovian responses affect human instrumental motor performance. *Front. Neurosci.* 6:134. doi:10.3389/fnins.2012.00134

Schacter, D. L., Addis, D. R., and Buckner, R. L. (2007). Remembering the past to imagine the future: the prospective brain. *Nat. Rev. Neurosci.* 8, 657–661.

Schacter, D. L., Addis, D. R., Hassabis, D., Martin, V. C., Spreng, R. N., and Szpunar, K. K. (2012). The future of memory: remembering, imagining, and the brain. *Neuron* 76, 677–694.

Schultz, W., Dayan, P., and Montague, P. (1997). A neural substrate of prediction and reward. *Science* 275, 1593–1599.

Shadlen, M. N., and Newsome, W. T. (2001). Neural basis of a perceptual decision in the parietal cortex (area lip) of the rhesus monkey. *J. Neurophysiol.* 86, 1916–1936.

Shadmehr, R. (2010). Control of movements and temporal discounting of reward. *Curr. Opin. Neurobiol.* 20, 726–730.

Silver, D., and Veness, J. (2010). "Monte-Carlo planning in large POMDPs," in *NIPS*, eds J. D. Lafferty, C. K. I. Williams, J. Shawe-Taylor, R. S. Zemel, and A. Culotta (Curran Associates, Inc), 2164–2172.

Simon, D. A., and Daw, N. D. (2011a). "Environmental statistics and the trade-off between model-based and td learning in humans," in *NIPS*, eds J. Shawe-Taylor, R. S. Zemel, P. L. Bartlett, F. C. N. Pereira, and K. Q. Weinberger (Granada), 127–135.

Simon, D. A., and Daw, N. D. (2011b). Neural correlates of forward planning in a spatial decision task in humans. *J. Neurosci.* 31, 5526–5539.

Solway, A., and Botvinick, M. M. (2012). Goal-directed decision making as probabilistic inference: a computational framework and potential neural correlates. *Psychol. Rev.* 119, 120–154.

Stewart, N., Chater, N., and Brown, G. D. A. (2006). Decision by sampling. *Cogn. Psychol.* 53, 1–26.

Sutton, R., and Barto, A. (1998). *Reinforcement Learning: An Introduction.* Cambridge, MA: MIT Press.

Sutton, R. S., and Barto, A. G. (1981). An adaptive network that constructs and uses an internal model of its environment. *Cogn. Brain Theory* 4, 217–246.

Sutton, R. S. (1990). "Integrated architectures for learning, planning, and reacting based on approximating dynamic programming," in *Proceedings of the Seventh International Conference on Machine Learning*, Bellevue, 216–224.

Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychol. Rev.* 55, 189–208.

van der Meer, M., Kurth-Nelson, Z., and Redish, A. D. (2012). Information processing in decision-making systems. *Neuroscientist* 18, 342–359.

van der Meer, M. A. A., and Redish, A. (2010). Expectancies in decision making, reinforcement learning, and ventral striatum. *Front. Neurosci.* 4:6. doi:10.3389/neuro.01.006.2010

van der Meer, M. A. A., and Redish, A. (2011). Ventral striatum: a critical look at models of learning and evaluation. *Curr. Opin. Neurobiol.* 21, 387–392.

van der Meer, M. A. A., and Redish, A. D. (2009). Covert expectation-of-reward in rat ventral striatum at decision points. *Front. Integr. Neurosci.* 3:1. doi:10.3389/neuro.07.001.2009

Verschure, P. F. M. J., Voegtlin, T., and Douglas, R. J. (2003). Environmentally mediated synergy between perception and behaviour in mobile robots. *Nature* 425, 620–624.

Watkins, C. J. C. H., and Dayan, P. (1992). Q-learning. *Mach. Learn.* 8, 279–292.

Wise, S. P. (2008). Forward frontal fields: phylogeny and fundamental function. *Trends Neurosci.* 31, 599–608.

Yin, H. H., Knowlton, B. J., and Balleine, B. W. (2004). Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. *Eur. J. Neurosci.* 19, 181–189.

Yin, H. H., Ostlund, S. B., and Balleine, B. W. (2008). Reward-guided learning beyond dopamine in the nucleus accumbens: the integrative functions of cortico-basal ganglia networks. *Eur. J. Neurosci.* 28, 1437–1448.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 10 July 2012; accepted: 08 February 2013; published online: 04 March 2013.

Citation: Pezzulo G, Rigoli F and Chersi F (2013) The Mixed Instrumental Controller: using Value of Information to combine habitual choice and mental simulation. Front. Psychol. ***4****:92. doi: 10.3389/fpsyg.2013.00092*

This article was submitted to Frontiers in Cognition, a specialty of Frontiers in Psychology.

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 20 November 2012
doi: 10.3389/fpsyg.2012.00514

Building a bridge into the future: dynamic connectionist modeling as an integrative tool for research on intertemporal choice

Stefan Scherbaum†, Maja Dshemuchadse*† and Thomas Goschke

Department of Psychology, Technische Universität Dresden, Dresden, Germany

Edited by:
Dorit Wenke, Humboldt University at Berlin, Germany

Reviewed by:
Petroc Sumner, Cardiff University, UK
Nick Chater, University College London, UK

***Correspondence:**
Maja Dshemuchadse, Department of Psychology, Technische Universität Dresden, Zellescher Weg 17, 01062 Dresden, Germany.
e-mail: maja@psychologie.tu-dresden.de

† Stefan Scherbaum and Maja Dshemuchadse have contributed equally to this work.

Temporal discounting denotes the fact that individuals prefer smaller rewards delivered sooner over larger rewards delivered later, often to a higher extent than suggested by normative economical theories. In this article, we identify three lines of research studying this phenomenon which aim (i) to describe temporal discounting mathematically, (ii) to explain observed choice behavior psychologically, and (iii) to predict the influence of specific factors on intertemporal decisions. We then opt for an approach integrating postulated mechanisms and empirical findings from these three lines of research. Our approach focuses on the dynamical properties of decision processes and is based on computational modeling. We present a dynamic connectionist model of intertemporal choice focusing on the role of self-control and time framing as two central factors determining choice behavior. Results of our simulations indicate that the two influences interact with each other, and we present experimental data supporting this prediction. We conclude that computational modeling of the decision process dynamics can advance the integration of different strands of research in intertemporal choice.

Keywords: decision making, temporal discounting, intertemporal choice, date-delay effect, impulsivity, time perception, dynamic systems, connectionist modeling

INTRODUCTION

Humans' self-image of being a "higher" species is justified in large part by reference to our extended mental abilities. In particular, our ability to anticipate the future enables us to defy momentary temptations or impulses and to make decisions based on foresight and long-term goals (Suddendorf et al., 2009; Goschke, 2012). Conversely, we are alerted when humans appear to ignore the future consequences of their behavior. Thus, researchers have been especially interested in understanding why sometimes human choices deviate from rationality standards as defined, for instance, by the economical rule of utility maximization (Fishburn, 1968). A prominent example of such a deviation can be found in intertemporal decision making, when humans have to choose between sooner and later delivered rewards. For such decisions, the original discounted utility model prescribes that the subjective value of a delayed option should decrease as an exponential function of the time until delivery (Samuelson, 1937). In contrast to this model, empirical studies found that individuals often discount rewards more steeply, especially for small time intervals (see Frederick et al., 2002 for a review). These and other observations suggesting that human choice behavior often deviates from normative rationality standards instigated an extensive research program on intertemporal choice behavior.

Within this broad field, different lines of research can be distinguished depending on whether their primary focus is on description, explanation, or prediction. In the following, we will shortly summarize core features of these three approaches and argue for an integrative approach that combines computational modeling with experimental studies of the process dynamics of choice behavior. As an initial step, we propose dynamic connectionist modeling as a tool supporting this integration and provide a first example of its potential benefits.

The descriptive approach originated from the original discounting model (Samuelson, 1937) and has led to the development of a range of formal models proposing various mathematical functions to fit the observed temporal discounting behavior (for an overview see Doyle, 2010). Comparisons of different discounting functions including exponential, hyperbolic, and hyperbola-like functions revealed that temporal discounting is often better described by hyperbola-like functions with more than one parameter (e.g., Green et al., 1994; McKerchar et al., 2009). However, although such models carry the promise of providing precise descriptions of the outcome of intertemporal decisions, they leave open the question which information-processing mechanisms underlie the observed deviations from normative rational choice standards.

The explanatory approach aims to fill this gap and has produced a wide range of different theories which attempt to explain the general pattern of hyperbolic temporal discounting in terms of underlying cognitive mechanisms that operate at different stages of the decision process (e.g., Stewart et al., 2006; Ebert and Prelec, 2007; Killeen, 2009; Zauberman et al., 2009; Scholten and Read, 2010; Trope and Liberman, 2010). Commonly the decision process is viewed as a transformation of a sensory input into a motor output through several consecutive stages, including the stage of option representation, the stage of value representation, and the stage of the final choice (cf. Sugrue et al., 2005; Rangel et al., 2008). At the stage of option representation, hyperbolic temporal discounting has been explained by an insensitive subjective perception of prospective durations leading to a logarithmic instead of a linear perception of temporal delays (Zauberman et al., 2009). At the stage of value representation, it has been proposed that the subjective value of an option is inferred by adding the utility of a good to the disutility of a delay thus leading to hyperbolic discounting (Killeen, 2009). At the stage of the final choice, Stewart et al. (2006) proposed a continuous accumulation of a frequency count of favorable binary comparisons between the offered options and value samples retrieved from memory, with hyperbolic discounting resulting from the real-world distribution of attribute values of gains, losses, and delays. Even this exemplary set of theories shows that a multitude of plausible explanations for the hyperbolic shape of the discounting function have been proposed. This raises the question, which of the proposed mechanisms (or which combination of mechanisms) is at work in a specific decision context and which variables determine to which degree a specific decision.

The predictive approach aims to provide answers to this question and is focused on the search for specific factors influencing the result of intertemporal decisions. Amongst the multitude of possible influences, two factors gained particular attention: self-control and contextual framing (cf. Berns et al., 2007). The ability to exert self-control is assumed to reduce the extent to which behavior is determined by automatic impulses triggered by an immediately available reward (Laibson, 1997; Hofmann et al., 2009; Heatherton and Wagner, 2011). This hypothesis is supported by clinical studies showing stronger discounting in patients with disorders presumably associated with higher impulsivity such as addiction and attention deficit hyperactivity disorder (e.g., Bickel and Marsch, 2001; Wittmann and Paulus, 2008). The role of contextual framing is emphasized by findings indicating that systematic biases strongly influence the degree of discounting (e.g., Loewenstein and Prelec, 1992). For example changing the framing of the time information from delays (e.g., "in 7 days") to calendar dates (e.g., "on the 13th of November") reduces temporal discounting, resulting in the so-called date-delay effect (Read et al., 2005; LeBoeuf, 2006). Altogether, the empirical studies have revealed numerous contextual factors modulating and moderating intertemporal choices.

While all three strands of research reviewed so far have yielded valuable insights into intertemporal choice behavior, they have to date often been pursued relatively segregated from each other with little cross-fertilization. To further advance the understanding of mechanisms and determinants of intertemporal choice, an integration of the different empirical findings and theoretical mechanisms is needed. We therefore propose an approach, based on computational modeling and a focus on the dynamical properties of decision processes, as an approach which could offer the required integrative and explanatory power. While a dynamic, process-oriented approach is common in research on perceptual decision making (Bogacz et al., 2007; Wang, 2008), it has only recently begun to find its way into research on economic decision making where a focus on stepwise mechanisms and decision results still dominates (e.g., Summerfield and Tsetsos, 2012). However, recent empirical work demonstrates the fruitfulness of a dynamic approach. For instance, in our own recent research we investigated specific influences on temporal discounting by tracking the decision process continuously over time using a mouse tracking procedure (cf. Spivey et al., 2005; Scherbaum et al., 2011). Results indicated an interaction of the influences of self-control and contextual framing (Dshemuchadse et al., 2012): less direct choice trajectories for later/larger options indicated more reflection (i.e., enhanced self-control) in contrast to choices of the sooner/smaller options. However, this difference was reduced when time was framed in calendar dates in contrast to delays.

In the following, we aim to combine this dynamic, process-oriented approach with connectionist models, that have already demonstrated their predictive power for multiattributive choice (Roe et al., 2001; Usher and McClelland, 2001; Glöckner and Betsch, 2008; Otter et al., 2008; for a comparison of the former two models see Tsetsos et al., 2010). We will explore the potential benefit of modeling intertemporal choice within a dynamic connectionist framework in two steps. First, we develop a neural network model that integrates several of the mechanisms and influencing factors described above. This model combines a logarithmic perception of time (cf. Zauberman et al., 2009), an additive valuation process (cf. Killeen, 2009), and an accumulation process based on the statistics of our environment (cf. Stewart et al., 2006). Additionally, the model accounts for the effects of the two central factors self-control and time framing and their interaction (e.g., Wittmann and Paulus, 2008; Dshemuchadse et al., 2012). Second, we validate the proposed computational model through an empirical study exploring the interaction of the two factors self-control and time framing. In this experiment, we varied time pressure to manipulate the amount of self-control and used different framings of the time information. This way, we aimed to dissociate the influence of the two factors and test the model predictions against empirical data.

A COMPUTATIONAL MODEL OF TEMPORAL DISCOUNTING

To model intertemporal choice behavior, we implemented the process of option evaluation (e.g., Busemeyer and Townsend, 1993; Johnson and Busemeyer, 2010) in a connectionist model (see also Roe et al., 2001). In a parallel distributed network model (Rumelhart and McClelland, 1986) options are represented as different activation patterns competing with each other (e.g., Usher and McClelland, 2001; Busemeyer and Johnson, 2004). The option represented by the pattern reaching the response threshold first wins the competition and determines the final choice behavior[1] (cf. Wang, 2008). The model incorporates the following five assumptions.

First, the activation of the option patterns accumulates gradually over time, following a non-linear activation function (cf. Usher and McClelland, 2001; Bogacz et al., 2007; Wang, 2008). The accumulation is terminated when one of the pattern reaches a threshold (cf. Busemeyer and Townsend, 1993; Wang, 2008).

---

[1] Although the representations of the magnitudes of the value and the time interval were not specified in the network architecture, one could implement a distributed representation within a specific network layer and the value of this representation could even be learned. However, the simplistic assumption of activation strength representing the magnitude of input variables is common in modeling decision processes (e.g., Roe et al., 2001; Usher and McClelland, 2001) and should not affect critical aspects of the dynamics of the accumulation process.

Second, an option receives activation by network units representing the option attributes time interval and value (cf. Roe et al., 2001) reflecting an additive valuation process (cf. Killeen, 2009). These option attribute units represent the properties of the two options through rate coding (cf. Shadlen and Newsome, 1998; van Rullen and Thorpe, 2001). Longer time intervals are represented by less activation (and hence less support for the option), following a non-linear function as has been proposed by previous empirical work (cf. Zauberman et al., 2009). Higher values are represented by increased activation (and hence more support). Taken together, this varying activation mirrors the preference for sooner and larger options.

Third, the speed of accumulation depends on the kind of information. Specifically, we assume that time information accumulates faster than value information, leading to a general dominance of time information and hence increased temporal discounting (cf. Dshemuchadse et al., 2012). Such an increased accumulation could be the result of differences in the connection weights resulting from the statistics of our environment (cf. Stewart et al., 2006).

Fourth, the degree of self-control influences the response threshold: less self-control will lower the response threshold, thereby leading to faster responses (cf. Busemeyer et al., 2006; Wittmann and Paulus, 2008; Kim and Lee, 2011).

Five, the contextual framing of information influences the accumulation rate of information: time information presented in terms of delays accumulates faster than time information presented in terms of dates (cf. Read et al., 2005). We assume that the more complex format of calendar dates requires increased processing and therefore longer decision times in comparison to delays. This assumption is in line with similar assumptions in models of perceptual decision making, which also postulate increased processing times for more difficult stimulus material (cf. Ratcliff and Smith, 2004; Palmer et al., 2005).

To validate our model, we will analyze its discounting behavior for effects described in the literature, namely the effects of self-control and the date-delay effect. We will then report the results of a behavioral experiment based on the simulated setup and compare the empirical data to the model predictions.

SIMULATION

MODEL AND HYPOTHESES

In the following, we will outline the model used in our simulation (for details on the architecture and parameters see the Appendix).

Layers and connectivity

The model architecture represents a feed forward leaky competing accumulator model containing two input layers, one for time information and one for value information, and a response layer, which integrates accumulating information and indicates the tendency to choose one of the two presented options (see **Figure 1**, for details, please see the Appendix).

The two units within each layer inhibit each other laterally (Usher and McClelland, 2001) while activating themselves, thus supporting non-linear dynamics (Wang, 2008). Each unit within an input layer is connected to and activates the respective response unit. Hence, unit 1 in the time layer and unit 1 in the value layer both activate unit 1 in the response layer, representing option 1 (the sooner/smaller option); unit 2 in the time layer and unit 2 in the value layer both activate unit 2 in the response layer, representing option 2 (the later/larger option).

Activation dynamics

While layers and connections define the static architecture of the model, the unit's activation dynamics define its reaction to an input, which is determined by the activation function. In line with previous connectionist/dynamic models, a non-linear sigmoid activation function was chosen (Cohen et al., 1992; Erlhagen and Schöner, 2002; Scherbaum et al., 2012). This ensures that each unit participates in the interaction between units only to the extent that its activation exceeds a soft threshold modeled by the sigmoid function (Erlhagen and Schöner, 2002). Hence, activation of attributes and responses and their interaction show non-linear properties. The non-linear dynamics is further enhanced by the recurrent excitatory connections, which lead to a competitive attractor dynamics (cf. Usher and McClelland, 2001; Bogacz et al., 2007; Wang, 2008).

SIMULATED PARADIGM

We implemented an intertemporal choice task in which simulated participants had to decide which of two options they preferred: the sooner but smaller or the later but larger option.

For each participant, we orthogonally varied the interval between the options (1, 3, 5, 8, 11, and 14 days) and the value of the sooner option in percentages of the value of the later option (20, 50, 70, 85, 95, and 99%). Additionally, we orthogonally varied the time of the sooner option (0 and 7 days).

Two variables were manipulated orthogonally between simulated participants (also see the Appendix): the response threshold, simulating an impulsive (low threshold) or a self-controlled (high threshold) choice, and the timescale of accumulation for the time information, simulating the framing of the time information as dates (slower accumulation) or delays (faster accumulation). Overall, we simulated 52 participants, leading to 13 participants per condition.

At the start of each trial, two options were presented to the simulated participants. A choice was made when one of the two response units reached the response threshold.

DATA PROCESSING

To examine the amount of discounting, we determined individual discounting functions for every simulated participant in two steps. First, we identified for each block separately the indifference point, i.e., the value difference for a particular time interval where a given simulated participant chose indifferently between the two options. As an estimate of the indifference point, we determined the point of inflection of a logistic function fitted to the individual choices (sooner/smaller vs. later/larger) as a function of increasing value differences (expressed in the ration sooner/later, cf. Ballard and Knutson, 2009). In the second step, we fitted for each participant a hyperbolic function[2] to the estimated indifference points over the

---

[2] The fitting of the hyperbolic function was performed by applying Matlab's multidimensional unconstrained nonlinear minimization function to the hyperbolic function $1/(1+k*x) = y$, with x denoting time interval, y denoting subjective value, and k denoting the discounting parameter.

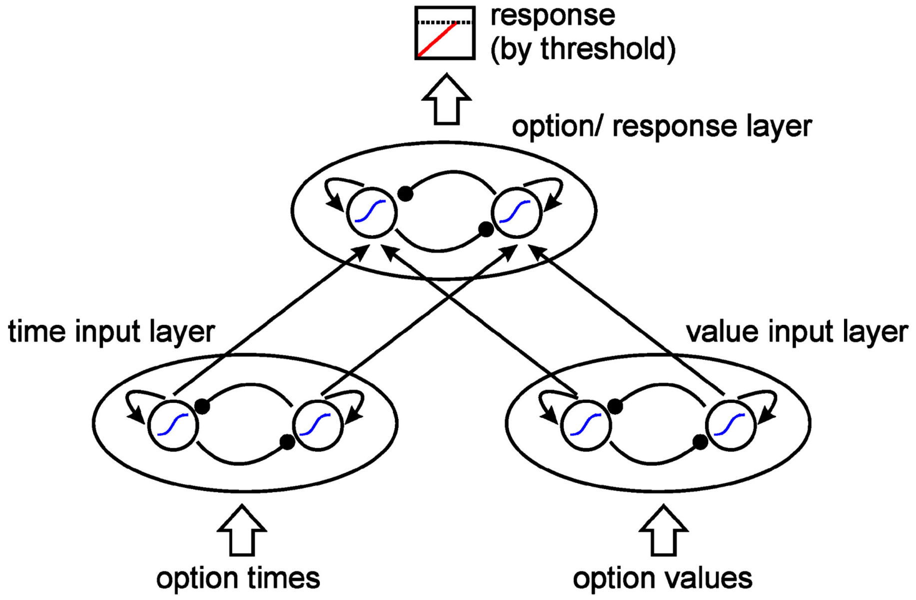

FIGURE 1 | Model architecture. Two input layers represent times and values of the two options. A response layer represents the choice preference of the network. Each unit in the input layers excites its respective response unit. Response is elicited by reaching a response threshold. All units follow a sigmoid activation function, show recurrent excitation, and inhibit each other within each layer.

different intervals and extracted the *k*-parameter of this function (Green et al., 1994).

RESULTS

As expected, simulated participants showed temporal discounting varying in steepness between the four different conditions (**Figure 2**) and varying in strength as measured by the *k*-parameter of hyperbolic functions fitted to the subjective values. Participants in the fast accumulation – low threshold condition exhibited the strongest discounting [$M(k) = 0.077, \mathrm{SD}(k) = 0.006$], followed by participants in the fast accumulation – high threshold condition [$M(k) = 0.047, \mathrm{SD}(k) = 0.004$]. The slow accumulation condition showed the weakest effects of discounting. Importantly, in this condition, there was no difference between the low [$M(k) = 0.026$, $\mathrm{SD}(k) = 0.003$] and the high threshold condition [$M(k) = 0.027$, $\mathrm{SD}(k) = 0.004$].

This indicates that the response threshold manipulation (simulating the degree of self-control) influenced the amount of discounting only when time information accumulated quickly (which by assumption is the case when time information is framed in terms of delays). The influence of the threshold vanishes when time information is accumulated slowly. However, the accumulation speed itself also influences discounting. This model behavior fits well previous empirical findings showing an interaction of the two factors self-control and contextual framing (Dshemuchadse et al., 2012).

Looking at the activation dynamics in the response layer suggests an explanation for these results (**Figure 3**). If time information accumulates faster, the activation of the sooner/smaller option dominates in the first part of a trial. If the threshold is sufficiently low, this option is actually chosen. However, with a higher threshold the activation of the later/larger option catches up due to the stronger, but delayed, activation elicited by its larger value, thereby leading to a reversal of the preferred option within the trial. This difference between thresholds vanishes when time accumulates slowly, since the activation of the later/larger option dominates during the entire trial.

To corroborate this analysis, we determined the number of activation reversals within the response layer for the different conditions. For each simulated participant, we counted the number of trials in which both options were dominant at least for some time within the trial. The number of such reversal trials mirrored the observed discounting pattern, with participants in the fast accumulation – high threshold condition showing the highest number of reversals of response activation within a trial ($M = 26.31$, $\mathrm{SD} = 1.49$), followed by the fast accumulation – low threshold condition ($M = 22$, $\mathrm{SD} = 1$), and the two very similar slow accumulation conditions (low threshold: $M = 4.62$, $\mathrm{SD} = 1.55$; high threshold: $M = 2.92$, $\mathrm{SD} = 1.5$).

Hence, the higher likelihood of within-trial preference reversals in the fast accumulation – high threshold condition was associated with less discounting, as this condition offered enough time for the later but larger choice tendency to overcome a premature choice of the sooner but smaller option. In contrast to this, the low likelihood of preference reversals in the two slow accumulation conditions and the low degree of discounting were due to a dominance of the later but larger option during the entire trial.

DISCUSSION

As expected, we found stronger discounting in the condition simulating time framing in terms of delays (which was assumed to lead to faster accumulation of time information) compared to the

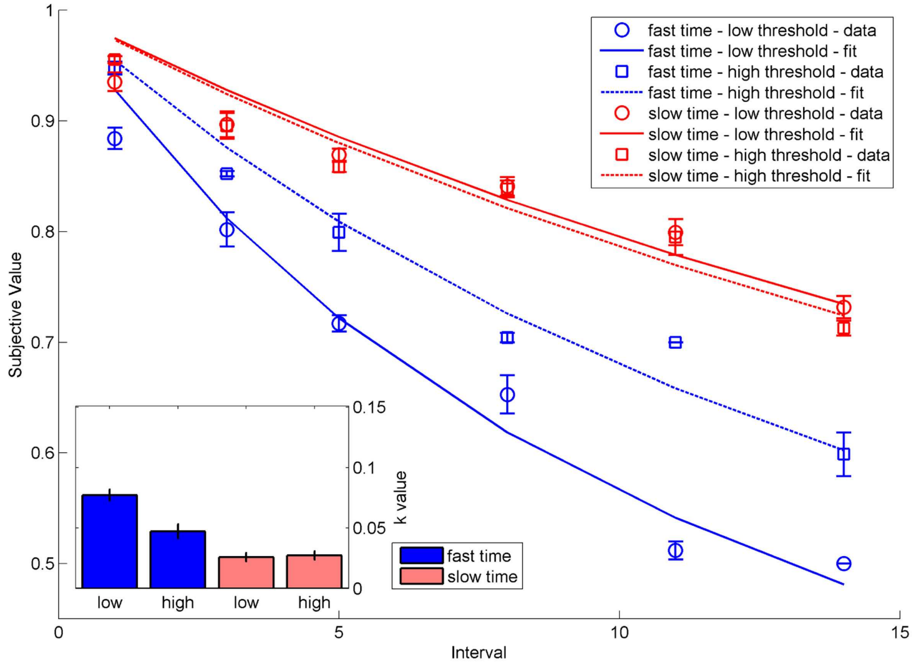

FIGURE 2 | Indifference points and hyperbolic functions depicting the decrease in subjective value as a function of intervals for the four conditions fast time accumulation – low threshold, fast time accumulation – high threshold, slow time accumulation – low threshold, slow time accumulation – high threshold. Error bars indicate standard errors. The inset shows the *k*-values of hyperbolic functions fit to the respective data. Error bars show standard deviations.

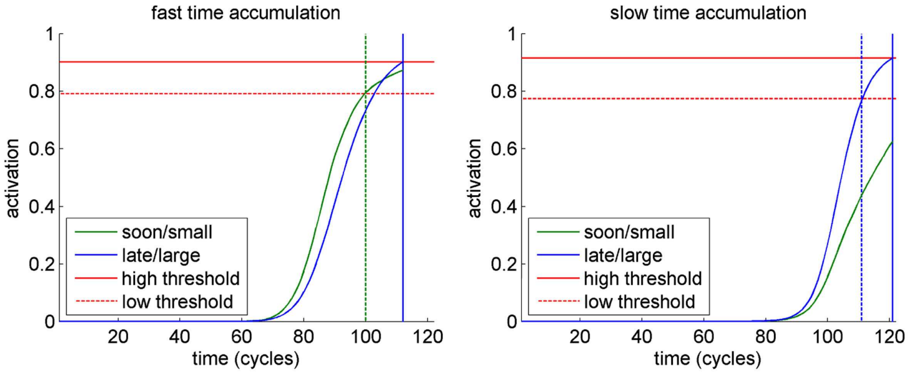

FIGURE 3 | Activation dynamics of the units in the response layer in a representative trial of one simulated participant in the different simulated conditions. In the fast time accumulation condition (left), the activation of the sooner/smaller response dominates in beginning of a trial. Hence, lowering the threshold changes the final choices (responses indicated by vertical lines). In the slow time accumulation condition (right), there is no difference in choice between the high and the low threshold.

condition simulating time framing in terms of dates (assumed to lead to slower accumulation of time information). Additionally, the model reproduced previous data (Dshemuchadse et al., 2012) in that it showed an effect of the simulated degree of self-control (which was implemented as a high vs. low response threshold) only in the delay condition, but not in the date condition.

An explanation for this behavior was suggested by the analysis of the activation dynamics within the response layer, which revealed reversals of the dominance of the choice options over the course of a trial (cf. Busemeyer and Townsend, 1993). In the delay condition, the faster accumulation of the time information had the effect that time information initially dominated the option preference, as it exerted a stronger influence on the activation of the option patterns than the value information. In a decision situation with low self-control – assumed to be associated with a lower response threshold – the final decision is predominantly driven by the more rapidly accumulating time information, leading to an overvaluation of time information and thus stronger temporal discounting. In contrast, in a decision situation with a higher degree of self-control – assumed to be associated with a higher response threshold – the accumulation process is prolonged, which leaves more time for the slowly accumulating value information to exert its influence on the option preferences and thus leads to less temporal discounting. In the date condition, however, this pattern changes, since by assumption the accumulation rate of the time information is reduced, leading to a more balanced influence of time and value. Therefore, in this condition the amount of temporal discounting is by and large independent from the degree of self-control.

In summary, our computational model of temporal discounting integrates theoretical assumptions derived from the explanatory approach and assumptions concerning the influence of specific factors (self-control and contextual framing) derived from a predictive approach, by specifying the non-linear dynamics of information accumulation during the option evaluation process. To validate the predictions of our model, we conducted an experiment in which we operationalized the simulated factors self-control and contextual framing and examined whether they would exert effects on human choice behavior mimicking the model predictions.

EXPERIMENT

The aim of the experiment was to investigate in an intertemporal choice task whether the two factors self-control and contextual framing would interact in the same way as predicted by our computational model. Firstly, to manipulate the amount of self-control, we imposed a response deadline forcing subjects to respond quickly and thus severely restricting the opportunity for deliberate reflection about the choice outcomes (cf. Kim and Lee, 2011). This way we aimed to reduce the length of the accumulation process in a way comparable to a lowered response threshold or an increased initial activation of response units in connectionist models (e.g., Botvinick et al., 2001). Under these circumstances, we predicted stronger temporal discounting compared to a control condition without a response deadline. Secondly, to vary the contextual framing of time information, we capitalized on the so-called date-delay effect (Read et al., 2005; LeBoeuf, 2006), which denotes the observation that time discounting is reduced when times are presented as calendar dates instead of delays. We assumed that framing time in calendar dates would lead to slower accumulation of the time information due to the more complex format. From our simulation results we derived the prediction that framing time in calendar dates should lead to less discounting and a reduced effect of the response deadline manipulation. In summary, by independently manipulating (i) the amount of reflection and/or self-control during intertemporal choices (via imposing a response deadline) and (ii) the accumulation rate of time information (via the contextual framing), we aimed to provide empirical evidence that these two factors exert an interactive influence on temporal discounting as predicted by our model simulation.

MATERIALS AND METHODS

Participants

Fifty students (32 female, mean age = 23.75) of the Technische Universität Dresden took part in the experiment and were assigned at random to the two framing (date vs. delay) conditions. All participants had normal or corrected to normal vision. They gave informed consent to the study and received class credit or 5 € payment.

Apparatus and stimuli

Stimuli were presented in white or gray on a black background on a 17 inch screen running at a resolution of 1280 × 1024 pixels (75 Hz refresh frequency). The experiment was controlled by the Eprime 1.2 software (Psychology Software Tools) running on a Windows XP SP2 personal computer. Subjects had to press the key X on a standard German computer keyboard to choose the sooner/smaller option and the key M to choose the later/larger option.

Two types of screens were presented to the subjects: preparation screens and choice screens (see **Figure 4**). On both types of screens, the two choice options were presented on the midline of the screen, with one option on the left side (sooner and smaller option) and one option on the right side (later and larger option). The font used for the presentation was Courier New with a size of 18 points. On the preparation screen only the values of the options (e.g., "20, 23 Euro") were presented in white color. On the choice screen, the values were presented again (albeit in gray color) and directly beneath them the corresponding delays, e.g., "in 3 Tagen" ("in 3 days") or the corresponding dates, e.g., "19 Juli" (July, 19) were shown in white color.

Procedure

On each trial participants had to decide which of two options they preferred: the left (sooner but smaller) or the right (later but larger) option. They were instructed to respond to the hypothetical choices as if they were real choices. Trials were grouped into mini-blocks of 14 trials (**Figure 4**). For each mini-block, the two monetary values remained constant and only the times of the two options were varied. Each mini-block consisted of a preparation screen followed by 14 choice screens. The preparation screen only showed the option values and was presented for 5 s. This procedure was chosen to allow participants to encode the value information in advance, because we suspected that otherwise the amount of information especially in the response deadline condition might lead to a neglect of some of the information. However, the main goal of the response deadline was not to restrict encoding of the option information but rather to restrict the time available for subsequent reflection about the options and their anticipated future outcomes. After the preparation phase, each of the subsequent 14 choice screens additionally showed the varying option times. Each choice screen was preceded by a fixation cross presented for 500 ms. Upon the presentation of each choice screen, participants had to indicate their choice, starting 300 ms after the screen onset. In the response deadline condition, they had to respond within a time window of 1800 ms after onset of the presentation screen. If they responded too late, a feedback screen was presented indicating an error. In the control condition, no deadline was imposed and participants were free respond at any time they chose.

Design

The experimental design was similar to the simulated paradigm, with a slightly increased number of time intervals and value

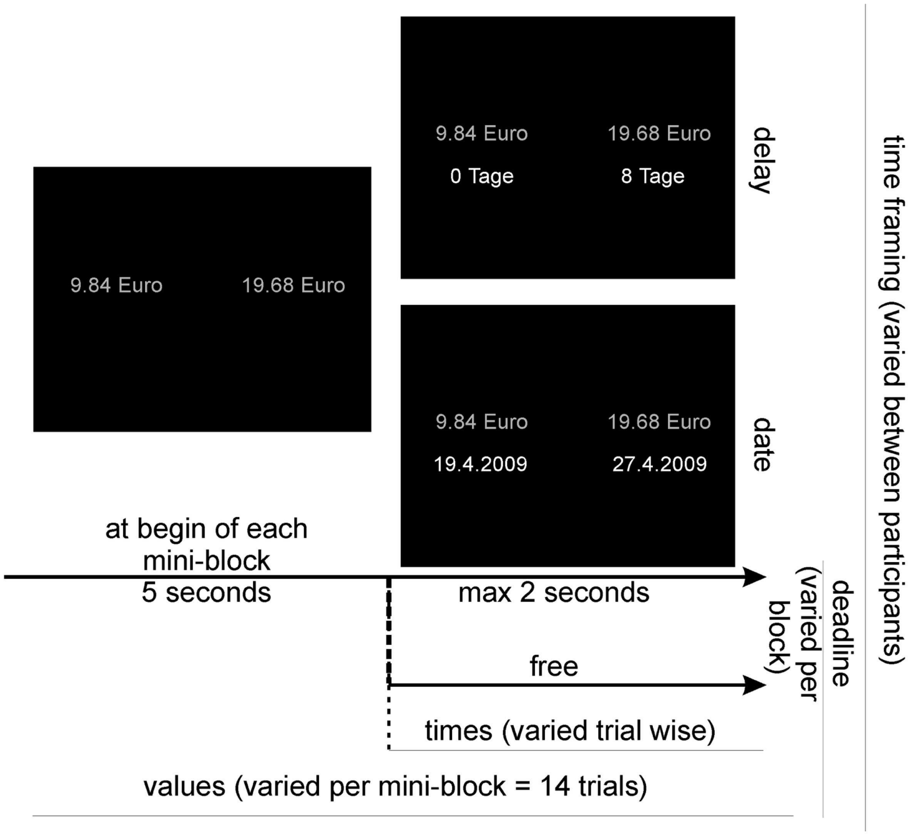

FIGURE 4 | Procedure and setup of the experiment. The subjects were divided into two groups (varying the time framing) and the experiment was split up into blocks (varying the response deadline), consisting each of 16 mini-blocks (varying option values), consisting in turn of 14 trials (varying option delays). Before each mini-block, the option values were presented for 5 s.

differences. Hence, for each participant and block, we orthogonally varied the time interval between the options (1, 2, 3, 5, 7, 10, and 14 days) and the value of the sooner option as percentages of the value of the later option (20, 50, 70, 80, 88, 93, 97, and 99%). The percentage of the value of the sooner option was varied between mini-blocks, while the time interval between the options was varied randomly between trials within each mini-block.

Additionally, we orthogonally varied the time of the sooner option (0 and 7 days) and the value of the later option (19.68 and 20.32 Euro). The time of the sooner option was varied to control for effects that may be specific for decisions where one of the options is immediately available (i.e., today) in contrast to decisions where both the sooner and later options are delayed. The value of the later option was varied to collect a sufficiently large number of data points without repeating identical trials, which could have induced memory effects. As neither of these two factors had any reliable effects, data was collapsed across them in the analyses reported below.

The response deadline was varied between blocks: one block with a response deadline of 1800 ms and another block without deadline were presented in random order. The framing of time (delay vs. date) was varied between subjects, who were randomly assigned to one of the two framing groups.

RESULTS

Experimental data

On 1.62% of the trials (SD = 1.83) in the deadline condition, responses were too slow and hence not included in the analyses. As expected, participants showed varying degrees of temporal discounting in the four different conditions (**Figure 5**). A mixed analysis of variance (ANOVA) with the within-subjects variable response deadline (with vs. without) and the between-subjects variable time framing (delay vs. date) and the k-parameter of the discounting function as the dependent variable (extracted from the data analogous to the simulation data processing) revealed a significant main effect of time framing, $F(1,48) = 4.78$, $p < 0.05$, and a significant interaction between response deadline and time framing, $F(1,48) = 7.34$, $p < 0.01$. The main effect reflected steeper discounting when time was framed in terms of delays compared to dates. The interaction reflected the fact that subjects showed steeper discounting in the response deadline condition compared to the condition without deadline, but this was only the case when the time was framed in terms of delays [delays: $t(24) = 2.12$, $p < 0.05$; dates: $t(24) = -1.16$, $p = 0.26$].

To examine the effectiveness of the deadline manipulation, an analogous ANOVA was computed with response time as the dependent variable. This analysis revealed a main effect of response

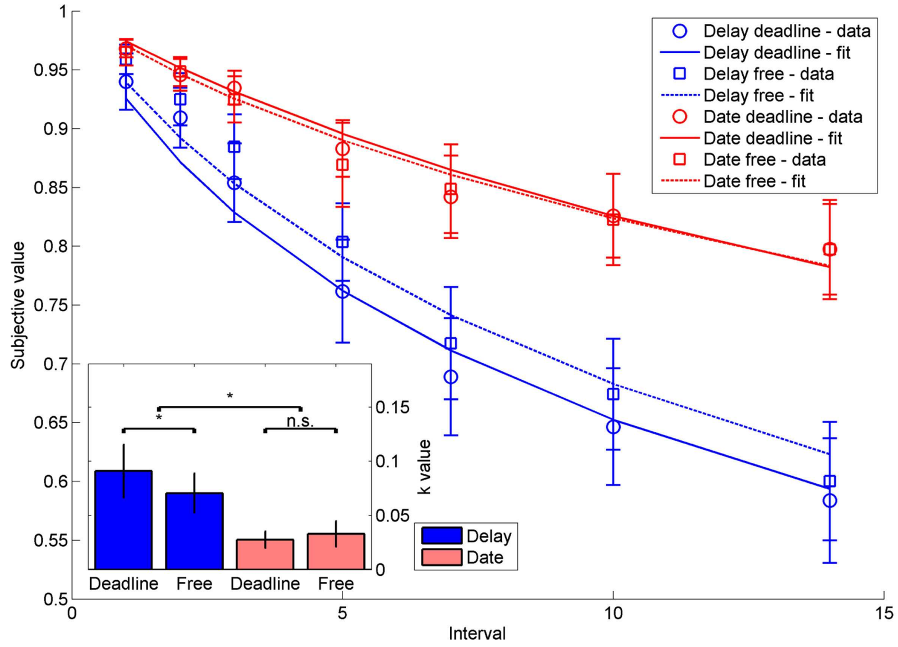

FIGURE 5 | Indifference points depicting the decrease in subjective value as a function of intervals for the four conditions delay-deadline, delay-free, date-deadline, and date-free. Error bars indicate standard errors. The inset shows the *k*-values of hyperbolic functions fit to the respective data. Stars mark statistical significance at $p < 0.05$, error bars indicate standard errors.

deadline as the only reliable result, $F(1,48) = 25.518$, $p < 0.001$ (all other ps > 0.3), indicating faster responses when subjects had to respond within the deadline ($M = 760$ ms, SD = 165 ms) compared to the condition without response deadline ($M = 1013$ ms, SD = 376 ms).

Comparison with the model predictions

To compare the simulated and experimental data, we performed two correlation analyses on the mean indifference curves in the different conditions and on the mean *k*-values in the different conditions. As expected, the indifference curves of simulated and experimental data were highly correlated ($r^2 = 0.96$, $p < 0.001$), as were the simulated and empirically obtained *k*-values ($r^2 = 0.97$, $p < 0.001$). There was thus a very good fit between model and experimental data.

DISCUSSION

As we had predicted from our computational model, the two variables response deadline and time framing not only exerted reliable effects on temporal discounting, but the experiment also yielded the expected interaction of the two variables. First, we replicated the standard date-delay effect, which was reflected in steeper discounting when time was framed in terms of delays compared to when it was framed in terms of dates. Secondly, we found that imposing a response deadline of 1800 ms induced steeper discounting compared to when participants responded at their leisure. However, most importantly, we obtained a reliable interaction between the two variables response deadline and time framing, which reflected the fact that the effect of the response deadline was only present in the delay but not in the date condition. The experiment thus replicated the critical results of the model simulation.

Interestingly, the influence of time framing was numerically much stronger than the influence of the response deadline. At present we do not know whether this reflects a genuine difference in the relative strength of the two factors or just reflects the fact that the deadline imposed in our experiment was to lenient to produce stronger effects on choice behavior. Although the effectiveness of the deadline manipulation was demonstrated by the fact that decisions times were reliably shorter when the response deadline was imposed, it must be noted that decision times in the condition without deadline were also relatively fast and on average well below the response deadline of 1.8 s. It is thus well conceivable that a stricter deadline, which would impose more severe restrictions on subjects' opportunity to recruit self-control would exert stronger effects on choice behavior and lead to a higher proportion of choices of sooner/smaller rewards.

Furthermore, the experimental setup differs slightly from the model concerning the presentation of the options. In the experimental setup, the values of the options are presented in advance. In contrast to this, the accumulation process for time interval and value starts simultaneously in the model. We assume that, although the values have been processed prior, the option evaluation process only starts when all information is presented. In line with a previous study (Dshemuchadse et al., 2012), our results support this assumption, since time information still dominates the final decision reflected in temporal discounting.

One general concern with computational models is the number of degrees of freedom when fitting model and empirical data due to the number of parameters that could be manipulated. It is therefore important to note that the model showed temporal discounting across a wide range of parameter configurations. Likewise, the critical effects of the response threshold and the accumulation rates were obtained across a wide range of parameter settings. Furthermore, we constrained the number of free parameters by setting several parameters such as the amount of lateral inhibition to a fixed value in all layers (for more details, see the Appendix). Last but not least it should be noted that, even though care has to be taken in choosing the parameters for a model, not every simple model will succeed in producing specific results and interaction patterns simply be fine-tuning of parameters (for further discussion see McClelland, 2009). In conclusion, the present empirical results validate core predictions derived from our computational model and indicate that different framings of time information are associated with more or less complex processing operations, which influence the accumulation rate of time information and thus the impact of this information on the option preferences particularly in the early phase of the decision process. As a result, presenting time in terms of delays increases the likelihood of choosing the sooner/smaller option, due to the stronger impact of the rapidly accumulating time information compared to the more slowly accumulating time information in the date condition.

GENERAL DISCUSSION

In this article, we presented a dynamic connectionist model of intertemporal choice behavior by which we attempted to integrate theoretical mechanisms derived from an explanatory approach and influencing factors (i.e., self-control and contextual framing) derived from a predictive approach. Our modeling approach builds on previous connectionist models of the process of option evaluation in multiattributive choice (Roe et al., 2001; Usher and McClelland, 2001). In our simulation of an intertemporal choice task, we modeled differences in the amount of self-control by varying response thresholds (assuming that a low response threshold – by promoting rapid decisions – reduces the likelihood that time consuming self-control processes are recruited prior to the final choice). Secondly, we modeled differences in the framing of time information (dates vs. delays) by varying the activation accumulation rates in the time input layer. The simulation yielded the typical date-delay effect: in the delay condition the model exhibited increased temporal discounting compared to the date condition. Furthermore, the simulation yielded evidence for an interaction between time framing and response threshold: a reduced response threshold (assumed to reflect less self-control) increased discounting, but this was the case only in the delay condition. This pattern was related to the frequency of re-decisions or "changes of mind" (Resulaj et al., 2009) within a trial and fits with results of a previous study, in which we used movement trajectories to investigate the time course of intertemporal decision making (Dshemuchadse et al., 2012). These model predictions were further successfully validated in a new behavioral experiment, in which we manipulated the hypothesized degree of self-control by imposing a response deadline and induced different time framings via the standard date-delay manipulation.

The present model and empirical data can be viewed as an initial proof of principle demonstrating the possible gain and feasibility of an approach to intertemporal choice, that focuses on the dynamical properties of the decision process and tests specific predictions derived from computational (e.g., connectionist) modeling. In the following, we will evaluate our dynamic, process-oriented approach, and discuss the integrative benefits in the context of the three research approaches to intertemporal choice distinguished in the introduction: the descriptive, the explanatory, and the predictive approach.

The descriptive approach provides mathematical functions to formalize central aspects of temporal discounting. This approach is integrated into our data analysis, where we fitted a hyperbolic function to the discounting curves. However, in contrast to findings indicating an optimal fit for models using functions with two or more parameters (e.g., Green et al., 1994; McKerchar et al., 2009) we choose a single-parameter hyperbola for two reasons. First, since the *k*-parameter and the hyperbolic model has been widely used in other studies of discounting (e.g., Kable and Glimcher, 2007; Ballard and Knutson, 2009), we attempted to make our results directly comparable to these studies. Second, since our primary goal was to compare model predictions with the empirical data, the single-parameter hyperbolic function offers a parsimonious characterization of discounting curves in terms of a single-parameter compared to models with several interdependent parameters. In conclusion, we capitalized on insights from the descriptive approach to derive a compact quantitative description of core aspects of decision behavior (Doyle, 2010).

The explanatory approach proposes theoretical mechanisms that apply at different stages of the decision process. Three theoretical assumptions concerning mechanisms were integrated into our computational model. First, the assumption of a logarithmic perception of time (cf. Zauberman et al., 2009) was embedded into the non-linear activation function of the network units representing the option attribute "time of delivery of a reward." Second, an additive valuation process (cf. Killeen, 2009) was implemented by having separate networks units represent the option attributes value and time, which then activated simultaneously the respective option. Third, we assumed that the accumulation of evidence (cf. Stewart et al., 2006) resulting in the final choice occurs with varying speed depending on the type of information.

Although we incorporated several mechanisms as postulated in other theories of choice behavior, we obviously also had to ignore other assumptions of these theories as well as a wide range of alternative theories not directly relevant for our dynamic modeling approach. On the one hand, we followed a process-oriented approach stemming from perceptual decision making (Bogacz et al., 2007; Wang, 2008; Summerfield and Tsetsos, 2012). Such an approach stands in contrast to theories of intertemporal choice building on stepwise mechanisms and focusing on the result of the decision (Trope and Liberman, 2003; Killeen, 2009; Loomes, 2010). On the other hand, our computational model was based on models of multiattributive choice (Roe et al., 2001; Usher and McClelland, 2001; Otter et al., 2008) with a competition process between options at its core: options are represented by different network units that inhibit each other and the choice is determined by the unit that is more strongly activated. This assumption stands in contrast to the assumptions and mechanisms of other models. For example, Stewart et al. (2006) proposed a competition between statistical frequencies: each option is compared with samples from memory, the frequency of favorable comparisons is counted, and the option with the higher frequency count is chosen. A further comparison mechanism was proposed by Scholten and Read (2010) between attributes: the attributes of the options are compared, the difference between the attributes is weighed against each other, and the more valued option is chosen. Finally, typical brain systems approaches are based on the competition between different subsystems of the brain. Metcalfe and Mischel (1999), for example, proposed that a hot brain system usually favors the sooner/smaller option and a cool brain system favors the later/larger option. Since the two systems do interact, the dominating system determines which option is chosen.

In summary, we made an attempt to integrate several mechanisms postulated within the explanatory approach into our computational model to demonstrate the potential gains of a dynamic process-oriented modeling approach to intertemporal choice. It has to be admitted, however, that in its current form our computational model is primarily intended as a proof of principle and will have to be elaborated further to explain a wider range of findings and to examine whether and in what respects its explanatory power may supersede that of alternative models of intertemporal choice (e.g., Stewart et al., 2006; Loomes, 2010; Scholten and Read, 2010). As integrative enhancements, the interaction between the different option attributes time and value (Scholten and Read, 2010) could be implemented via inhibition between the two layers; the finding of greater discounting rates for gains than for losses (Thaler and Shefrin, 1981) could be implemented via different accumulation rates as it has been done for the different time framings; the effect of stronger discounting under memory-load (Hinson et al., 2003, but see Franco-Watkins et al., 2006) can be explained with memory-load restraining resources and hence restricting deliberate reflection comparable to the influence of time restriction.

The third general approach discussed in the introduction, that we termed the predictive approach, aims to identify factors influencing intertemporal choices. Two such factors were included into our computational model and the reported experiment: the amount self-control and the contextual framing of time information. The amount of self-control was manipulated by varying the response threshold in the model and by imposing a response deadline in the experiment. Lowering the response threshold in the model led to faster responses due to a shorter process of evidence accumulation (cf. Busemeyer et al., 2006). Alternatively, one could have varied the baseline activation level to prolong or speed up responses, which, however, leads in most cases to similar results (see, e.g., Botvinick et al., 2001 in the context of a model accounting for post-error slowing). By imposing a response deadline in the experiment, we forced subjects to respond quickly, which should likewise reduce the duration of the evidence accumulation process and is known to induce more impulsive choices (Kim and Lee, 2011). Our assumption that a lowered response threshold (as induced by a response deadline) leads to reduced self-control is consistent with the fact that these processes are time consuming and fits with evidence indicating that a lack of self-control is associated with impaired behavioral inhibition and more impulsive choices (Soubrie, 1986; Stein et al., 1993). While this relatively general use of the term self-control suffices for the purposes of the present investigation, it should be noted that self-control is a multifaceted construct (e.g., Evenden, 1999; Santisteban and Arce, 2006) allowing for alternative implementations as, for instance, in theories postulating multiple decision systems (e.g., Thaler and Shefrin, 1981; Fudenberg and Levine, 2006).

To examine the influence of contextual framing – and specifically the framing of time information – on intertemporal choice, we manipulated the accumulation rate of time information in the model and the presentation format (delay vs. calendar dates) in the experiment. The manipulation of the accumulation rate rests on the assumption that the processing of dates is more complex than the processing of delays. This should lead to different rates at which time information accumulates in the respective processing layer, in a manner analogously to what has been assumed in models of perceptual decision making (cf. Ratcliff and Smith, 2004; Palmer et al., 2005). Our manipulation of the format of the time information in the experiment relied on findings from previous studies of the date-delay effect (Read et al., 2005; LeBoeuf, 2006) and yielded findings consistent with this earlier work. Nevertheless, it should be mentioned that alternative interpretations of the date-delay effect have been proposed (Read et al., 2005; LeBoeuf, 2006). While in the present study we examined two critical factors influencing intertemporal choice – self-control and contextual framing – it is an aim for future investigations to extend the present model to account for other relevant factors (see, e.g., Frederick et al., 2002) and different forms of contextual framing (see, e.g., Kahneman and Tversky, 1984).

In summary, the present model and empirical results provide an initial demonstration of the gain and feasibility of a dynamical, process-oriented approach to intertemporal choice based on computational modeling. By combining connectionist modeling and experimental data, we obtained evidence that self-control and time framing exert interactive effects on temporal discounting, which can be accounted for by dynamic properties of the decision process, in particular, the interaction of different accumulation rates and different response thresholds.

ACKNOWLEDGMENTS

This research was partly supported by the German Research Council (DFG; Grant SFB 940 2012).

REFERENCES

Amari, S. (1977). Dynamics of pattern formation in lateral-inhibition type neural fields. *Biol. Cybern.* 27, 77–87.

Ballard, K., and Knutson, B. (2009). Dissociable neural representations of future reward magnitude and delay during temporal discounting. *Neuroimage* 45, 143–150.

Berns, G. S., Laibson, D., and Loewenstein, G. (2007). Intertemporal choice – toward an integrative framework. *Trends Cogn. Sci. (Regul. Ed.)* 11, 482–488.

Bickel, W. K., and Marsch, L. A. (2001). Toward a behavioral economic understanding of drug dependence: delay discounting processes. *Addiction* 96, 73–86.

Bogacz, R., Usher, M., Zhang, J., and McClelland, J. L. (2007). Extending a biologically inspired model of choice: multi-alternatives, nonlinearity and value-based multidimensional choice. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 362, 1655–1670.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., and Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychol. Rev.* 108, 624–652.

Busemeyer, J. R., and Johnson, J. G. (2004). Computational models of decision making. *Judgm. Decis. Mak.* 1, 133–154.

Busemeyer, J. R., Johnson, J. G., and Jessup, R. K. (2006). "Preferences constructed from dynamic microprocessing mechanisms," in *The Construction of Preference*, eds S. Lichtenstein, and P. Slovic (New York: Cambridge University Press), 220–234.

Busemeyer, J. R., and Townsend, J. T. (1993). Decision field theory: a dynamic-cognitive approach to decision making in an uncertain environment. *Psychol. Rev.* 100, 432–459.

Cohen, J. D., Servan-Schreiber, D., and McClelland, J. L. (1992). A parallel distributed processing approach to automaticity. *Am. J. Psychol.* 239–269.

Doyle, J. (2010). *Survey of Time Preference, Delay Discounting Models.* Social Science Research Network. Available at: http://ssrn.com/abstract=1685861

Dshemuchadse, M., Scherbaum, S., and Goschke, T. (2012). How decisions emerge: action dynamics in intertemporal decision making. *J. Exp. Psychol. Gen.* [Epub ahead of print].

Ebert, J. E. J., and Prelec, D. (2007). The fragility of time: time-insensitivity and valuation of the near and far future. *Manage. Sci.* 53, 1423–1438.

Erlhagen, W., and Schöner, G. (2002). Dynamic field theory of movement preparation. *Psychol. Rev.* 109, 545–572.

Evenden, J. L. (1999). Varieties of impulsivity. *Psychopharmacology (Berl.)* 146, 348–361.

Fishburn, P. C. (1968). Utility theory. *Manage. Sci.* 14, 335–378.

Franco-Watkins, A. M., Pashler, H., and Rickard, T. C. (2006). *Does Working Memory Load Lead to Greater Impulsivity*? Commentary on Hinson, Jameson, and Whitney (2003). Available at: http://psycnet.apa.org/journals/xlm/32/2/443/ [accessed November 1, 2012].

Frederick, S., Loewenstein, G., and O'Donoghue, T. (2002). Time discounting and time preference: a critical review. *J. Econ. Lit.* 40, 351–401.

Fudenberg, D., and Levine, D. K. (2006). A dual-self model of impulse control. *Am. Econ. Rev.* 1449–1476.

Glöckner, A., and Betsch, T. (2008). *Modeling Option and Strategy Choices with Connectionist Networks: Towards an Integrative Model of Automatic and Deliberate Decision Making.* SSRN eLibrary. Available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1090866 [accessed September 10, 2012].

Goschke, T. (2012). "Volition in action: intentions, control dilemmas and the dynamic regulation of intentional control," in *Action Science: Foundations of an Emerging Discipline*, eds W. Prinz, A. Beisert, and A. Herwig (Cambridge, MA: MIT Press) (in press).

Green, L., Fristoe, N., and Myerson, J. (1994). Temporal discounting and preference reversals in choice between delayed outcomes. *Psychon. Bull. Rev.* 1, 383–389.

Heatherton, T. F., and Wagner, D. D. (2011). Cognitive neuroscience of self-regulation failure. *Trends Cogn. Sci. (Regul. Ed.)* 15, 132–139.

Hinson, J. M., Jameson, T. L., and Whitney, P. (2003). Impulsive decision making and working memory. *J. Exp. Psychol. Learn. Mem. Cogn.* 29, 298–306.

Hofmann, W., Friese, M., and Strack, F. (2009). Impulse and self-control from a dual-systems perspective. *Perspect. Psychol. Sci.* 4, 162–176.

Johnson, J. G., and Busemeyer, J. R. (2010). Decision making under risk and uncertainty. *Cogn. Sci.* 1, 736–749.

Kable, J. W., and Glimcher, P. W. (2007). The neural correlates of subjective value during intertemporal choice. *Nat. Neurosci.* 10, 1625–1633.

Kahneman, D., and Tversky, A. (1984). Choices, values, and frames. *Am. Psychol.* 39, 341–350.

Killeen, P. R. (2009). An additive-utility model of delay discounting. *Psychol. Rev.* 116, 602–619.

Kim, S., and Lee, D. (2011). Prefrontal cortex and impulsive decision making. *Biol. Psychiatry* 69, 1140–1146.

Laibson, D. (1997). Golden eggs and hyperbolic discounting. *Q. J. Econ.* 112, 443–477.

LeBoeuf, R. A. (2006). Discount rates for time versus dates: the sensitivity of discounting to time-interval description. *J. Mark. Res.* 43, 59–72.

Loewenstein, G., and Prelec, D. (1992). Anomalies in intertemporal choice: evidence and an interpretation. *Q. J. Econ.* 107, 573–597.

Loomes, G. (2010). Modeling choice and valuation in decision experiments. *Psychol. Rev.* 117, 902–924.

McClelland, J. L. (2009). The place of modeling in cognitive science. *Top. Cogn. Sci.* 1, 11–38.

McKerchar, T. L., Green, L., Myerson, J., Pickford, T. S., Hill, J. C., and Stout, S. C. (2009). A comparison of four models of delay discounting in humans. *Behav. Processes* 81, 256–259.

Metcalfe, J., and Mischel, W. (1999). A hot/cool-system analysis of delay of gratification: dynamics of willpower. *Psychol. Rev.* 106, 3–19.

Otter, T., Johnson, J., Rieskamp, J., Allenby, G. M., Brazell, J. D., Diederich, A., et al. (2008). Sequential sampling models of choice: some recent advances. *Mark. Lett.* 19, 255–267.

Palmer, J., Huk, A. C., and Shadlen, M. N. (2005). The effect of stimulus strength on the speed and accuracy of a perceptual decision. *J. Vis.* 5, 376–404.

Rangel, A., Camerer, C., and Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nat. Rev. Neurosci.* 9, 545–556.

Ratcliff, R., and Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychol. Rev.* 111, 333–367.

Read, D., Frederick, S., Orsel, B., and Rahman, J. (2005). Four score and seven years from now: the date/delay effect in temporal discounting. *Manage. Sci.* 51, 1326–1335.

Resulaj, A., Kiani, R., Wolpert, D. M., and Shadlen, M. N. (2009). Changes of mind in decision-making. *Nature* 461, 263–266.

Roe, R. M., Busemeyer, J. R., and Townsend, J. T. (2001). Multialternative decision field theory: a dynamic connectionist model of decision making. *Psychol. Rev.* 108, 370–392.

Rumelhart, D. E., and McClelland, J. L. (1986). *Parallel Distributed Processing. Explorations in the Microstructure of Cognition: Foundations*, Vol. 1. Cambridge: MIT Press.

Samuelson, P. A. (1937). A note on measurement of utility. *Rev. Econ. Stud.* 4, 155–161.

Santisteban, C., and Arce, E. (2006). Impulsivity: a review. *Psicothema* 18, 213–220.

Scherbaum, S., Dshemuchadse, M., Ruge, H., and Goschke, T. (2012). Dynamic goal states: adjusting cognitive control without conflict monitoring. *Neuroimage* 63, 126–136.

Scherbaum, S., Fischer, R., Dshemuchadse, M., and Goschke, T. (2011). The dynamics of cognitive control: evidence for within-trial conflict adaptation from frequency-tagged EEG. *Psychophysiology* 48, 591–600.

Scholten, M., and Read, D. (2010). The psychology of intertemporal tradeoffs. *Psychol. Rev.* 117, 925–944.

Shadlen, M. N., and Newsome, W. T. (1998). The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *J. Neurosci.* 18, 3870–3896.

Soubrie, P. (1986). Reconciling the role of central serotonin neurons in human and animal behavior. *Behav. Brain Sci.* 9, 319–335.

Spivey, M. J., Grosjean, M., and Knoblich, G. (2005). Continuous attraction toward phonological competitors. *Proc. Natl. Acad. Sci. U.S.A.* 102, 10393–10398.

Stein, D. J., Hollander, E., and Liebowitz, M. R. (1993). Neurobiology of impulsivity and the impulse control disorders. *J. Neuropsychiatry Clin. Neurosci.* 5, 9–17.

Stewart, N., Chater, N., and Brown, G. D. A. (2006). Decision by sampling. *Cogn. Psychol.* 53, 1–26.

Suddendorf, T., Addis, D. R., and Corballis, M. C. (2009). Mental time travel and the shaping of the human mind. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 364, 1317–1324.

Sugrue, L. P., Corrado, G. S., and Newsome, W. T. (2005). Choosing the greater of two goods: neural currencies for valuation and decision making. *Nat. Rev. Neurosci.* 6, 363–375.

Summerfield, C., and Tsetsos, K. (2012). Building bridges between perceptual and economic decision-making: neural and computational mechanisms. *Front. Neurosci.* 6:70. doi:10.3389/fnins.2012.00070

Thaler, R. H., and Shefrin, H. M. (1981). An economic theory of self-control. *J. Polit. Econ.* 392–406.

Trope, Y., and Liberman, N. (2003). Temporal construal. *Psychol. Rev.* 110, 403–421.

Trope, Y., and Liberman, N. (2010). Construal-level theory of psychological distance. *Psychol. Rev.* 117, 440–463.

Tsetsos, K., Usher, M., and Chater, N. (2010). Preference reversal in multiattribute choice. *Psychol. Rev.* 117, 1275–1291.

Usher, M., and McClelland, J. L. (2001). The time course of perceptual choice: the leaky, competing accumulator model. *Psychol. Rev.* 108, 550–592.

van Rullen, R., and Thorpe, S. J. (2001). Rate coding versus temporal order coding: what the retinal ganglion cells tell the visual cortex. *Neural Comput.* 13, 1255–1283.

Wang, X. J. (2008). Decision making in recurrent neuronal circuits. *Neuron* 60, 215–234.

Wittmann, M., and Paulus, M. P. (2008). Decision making, impulsivity and time perception. *Trends Cogn. Sci. (Regul. Ed.)* 12, 7–12.

Zauberman, G., Kim, B. K., Malkoc, S. A., and Bettman, J. R. (2009). Discounting time and time discounting: subjective time perception and intertemporal preferences. *J. Mark. Res.* 46, 543–556.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 04 June 2012; accepted: 01 November 2012; published online: 20 November 2012.

Citation: Scherbaum S, Dshemuchadse M and Goschke T (2012) Building a bridge into the future: dynamic connectionist modeling as an integrative tool for research on intertemporal choice. Front. Psychology ***3****:514. doi: 10.3389/fpsyg.2012.00514*

This article was submitted to Frontiers in Cognition, a specialty of Frontiers in Psychology.

APPENDIX

MODEL ARCHITECTURE

The model consists of two input layers and a response layer, with two units per layer. Activation of each unit is calculated by non-linear first order differential equations as has been done previously for patterns of neural activation (Amari, 1977; Erlhagen and Schöner, 2002). Simulated by numerical integration, results were obtained using Matlab 2006a running under Windows XP SP3. The difference equation over time t for the activation u of units in a layer had the following form:

$$\tau\dot{u}(t) = -u(t) - h + wi \cdot \sigma(u(t)) + w \cdot \text{Input}(t) + N$$

Here, τ denotes the timescale, h the resting level, wi the interaction weight within the layer (self excitation and lateral inhibition), and w the weight of inputs into the layer; Input defines the input into the layer, N denotes random noise (distributed normally with $M = 0$ and $\text{SD} = 0.0025$), and σ denotes a sigmoid non-linearity, mirroring neural population dynamics:

$$\sigma(x) = 1\big/\left(1 + e\left(-\beta \cdot (x - \alpha)\right)\right).$$

Hence, each unit contributes to interactions in the network only to the extent that its activation exceeds a soft threshold (Cohen et al., 1992; Erlhagen and Schöner, 2002).

Following this scheme, the equation for the input layers was:

$$\tau\dot{u}(t) = -u(t) - h + wi_i \cdot \sigma(u(t)) + w_{si} \cdot S(t) + N$$

Here, wi_i denotes the interaction weight within the input layers, w_{si} the weight of external inputs into the layers, $S(t)$ represents the external stimulus input into the layer. To ensure baseline levels of activations for external stimulations for all possible inputs, inputs were defined by

$$S(t) = S_{\max} - w_{si} + S_{\text{raw}}(t) \cdot w_{si}.$$

Here, $S_{\max}$ denotes the maximum strength of the input signal, set to 7, S_{raw} denotes the signals defined by the values and times of the respective options (ranging from 0 to 1, see model input, below), and w_{si} denotes the weight of the input with respect to $S_{\max}$, set to 0.3 for times and 0.7 for values.

For the input layer representing value information, we set $\tau_V = 30$. For the input layer representing time information, we simulated different speeds of information accumulation, by setting $\tau_T = 10$ for the fast accumulation condition and setting $\tau = 30$ for the slow accumulation condition.

Analogously to the input layers, the equation for the response layer was:

$$\tau\dot{u}(t) = -u(t) - h + wi_r \cdot \sigma(u(t)) + w_{i_1 r} \cdot I_1(t) + w_{i_2 r} \cdot I_2(t) + N$$

Here, wi_r denotes the interaction weight within the response layer,$w_{i_1 r}$ $w_{i_2 r}$ denote the strength of input from the input layers, and I_1 and I_2 represent the signal from the input layers. Responses were considered as made when $\sigma(u(t))$ reached a response threshold. This threshold was sampled at random from a normal distribution with an SD of 0.0075 and a mean of 0.9 for the high threshold condition and a mean of 0.77 in the low threshold condition. The timescale of information accumulation was set to the same value as for the value input layer, τ_V, hence $\tau = 30$.

The weight matrices are shown in the following. The interactions within the input layers, and the response layer were defined by

$$wi_i = \begin{pmatrix} 1 & -2 \\ -2 & 1 \end{pmatrix},.\ wi_r = \begin{pmatrix} 1 & -2 \\ -2 & 1 \end{pmatrix}$$

Hence, within all layers, there was the same strong lateral inhibition compared to a weaker self excitation of each node.

Signal transmission from each input layer to the response layer was defined by

$$w_{i_1 r} = w_{i_2 r} = \begin{pmatrix} 5 & 0 \\ 0 & 5 \end{pmatrix}.$$

Hence, the input layers were associated equally with the response layer and each unit within an input layer representing the time or value of an option activated the response unit representing the preference for this option.

The other parameters where chosen as follows: $h = 5$, $\alpha = 0$, $\beta = 1.5$.

The parameters $h, \alpha, \beta, S_{\max}$ and the connection weights w_{si}, wi_i, wi_r,$w_{i_1 r}$ $w_{i_2 r}$ were chosen to produce classical discounting behavior. By choosing equal values for wi_i and wi_r as well as for $w_{i_1 r}$ $w_{i_2 r}$ we aimed to minimize the number of free parameters and keep the model as simple as possible. Within these constraints, the model's discounting behavior was qualitatively similar across a wide range of parameter choices.

The two critical parameter variations concerned the response threshold (0.9 vs. 0.77) and the accumulation rate of the time information τ_T. Again, we set τ_T, τ_R, and τ_T to equal values (in the slow accumulation condition) to minimize the number of free parameters. Hence, the only free parameter was τ_T in the fast accumulation condition, with the constraint $\tau_{T(fast)} < \tau_{T(fast)}$. Within these constraints, the presented effects were qualitatively stable across a wide range of parameter combinations.

CALCULATION OF MODEL INPUT AND PROCEDURE

The input to the input layers representing time and value information for options 1 and 2 was defined by the input vectors $S_T = (T_1, T_2)$ and $S_V = (V_1, V_2)$. S_T and S_V were varied orthogonally (see description of paradigm). For S_T, T_1 was chosen from $\{0,7\}$. T_2 was defined by $T_1 + T_I$, the interval between the options, with T_I chosen from $\{1, 3, 5, 8, 11, 14\}$. S_T was then transformed to normalized input values by

$$S_T = 1 - \sqrt{S_T\big/T_{\max}}.$$

Hence, time was normalized to the maximum possible time value, transformed non-linearly to mirror non-linear time perception (see main text) and inverted, so that lower times lead to higher input activation, mirroring the preference for smaller delays.

For S_V, V_2 was set to 1, and V_1 was chosen so that the ration V_1/V_2 was {0.2, 0.5, 0.7, 0.85, 0.95, 1}. Since S_V was already normalized to a maximum value of 1, no further transformation was necessary to receive normalized input values. Hence, higher values lead to higher input activation, mirroring the preference for high values.

Each trial began with an inter-trial interval of 50 cycles without input, followed by the activation of the input vectors. The trial ended when the output activation of one of the two response units reached the response threshold and, hence, a choice was performed.

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 08 October 2012
doi: 10.3389/fpsyg.2012.00383

What's next: recruitment of a grounded predictive body model for planning a robot's actions

***Malte Schilling*[1,2]* *and Holk Cruse*[2]**

[1] *International Computer Science Institute, Berkeley, CA, USA*
[2] *Center of Excellence 'Cognitive Interaction Technology', University of Bielefeld, Bielefeld, Germany*

Edited by:
Dorit Wenke, Humboldt University, Germany

Reviewed by:
Lei Wang, RWTH Aachen University, Germany
Martin V. Butz, Eberhard Karls University of Tübingen, Germany

****Correspondence:***
Malte Schilling, International Computer Science Institute Berkeley, 1947 Center Street, Suite 600, Berkeley, CA 94704, USA.
e-mail: malteschilling@googlemail.com

Even comparatively simple, reactive systems are able to control complex motor tasks, such as hexapod walking on unpredictable substrate. The capability of such a controller can be improved by introducing internal models of the body and of parts of the environment. Such internal models can be applied as inverse models, as forward models or to solve the problem of sensor fusion. Usually, separate models are used for these functions. Furthermore, separate models are used to solve different tasks. Here we concentrate on internal models of the body as the brain considers its own body the most important part of the world. The model proposed is formed by a recurrent neural network with the property of pattern completion. The model shows a hierarchical structure but nonetheless comprises a holistic system. One and the same model can be used as a forward model, as an inverse model, for sensor fusion, and, with a simple expansion, as a model to internally simulate (new) behaviors to be used for prediction. The model embraces the geometrical constraints of a complex body with many redundant degrees of freedom, and allows finding geometrically possible solutions. To control behavior such as walking, climbing, or reaching, this body model is complemented by a number of simple reactive procedures together forming a procedural memory. In this article, we illustrate the functioning of this network. To this end we present examples for solutions of the forward function and the inverse function, and explain how the complete network might be used for predictive purposes. The model is assumed to be "innate," so learning the parameters of the model is not (yet) considered.

Keywords: prediction, anticipation, recurrent neural network, internal body model, internal simulation, minimal cognitive system, robotic architecture

INTRODUCTION

The capability of not only reacting to actual stimuli, but also predicting future stimuli, was for a long time attributed to "higher animals" and therefore tightly connected to properties of (some) vertebrate brains. Now, however, not even "simple" animals like insects are considered merely reactive; it is now known that they are able to anticipate future situations. Anticipation, i.e., the use of information about what will be next, is used to guide actions. Examples include the prediction of the future position of a moving object, which can be used to visually pursue or reach for it, and the estimation of the mass of an object to be lifted. To allow for such prediction, internal models of the environment are required. Seen from the brain's point of view (Cruse, 1999), an essential part and a starting point is a model of the body. Therefore, internal models not only refer to objects in the external environment, but also have to include a simulation of – at least parts of – the body.

Usually, two types of models are distinguished (Kawato, 1999, p. 718):

> "Internal models are neural mechanisms that can mimic the input/output characteristics, or their inverses, of the motor apparatus. Forward internal models can predict sensory consequences from efference copies of issued motor commands. Inverse internal models, on the other hand, can calculate necessary feedforward motor commands from desired trajectory information." Here, we add a third function, namely sensor fusion and want to explain these function in more detail.

INVERSE MODELS

Classical paradigms for inverse models are targeted and goal-directed movements that fundamentally rely on an internal model. The simple ability to grasp an object seems to be carried out without any explicit planning of the movement but by application of controllers using sensory feedback. But the action is not merely controlled through visual feedback. Targeted movements can be accomplished without sight and so fast that a feedback control loop, which inevitably would include certain delays, would be too slow to account for the behavior (Miall et al., 1993; Desmurget and Grafton, 2000). As a possible solution to this problem, it has been assumed that the controller implements a transformation of the target description onto the actuator dynamics. The target position – given through visual input – may be defined in an ego-centric Cartesian space. To reach the target, the position, and the reaching movement must, however, be described in terms of joint or muscle activations in some form. A transformation between these two reference systems represents a mapping from Cartesian space to joint space. This is called an inverse model (Wolpert and Kawato, 1998).

Solving this inverse kinematic problem is difficult because, as is the case in most if not all biological control problems, the controlled system, in our case the limb, contains extra degrees of freedom (DoF), i.e., more joints than necessary for the solution to the task (Bernstein, 1967). This "ill-posedness" means that there is not only one but many solutions. Therefore, the controller has to select one out of these many possible solutions.

Visually guided reaching movements have been the subject of many studies in humans (Castiello, 2005; Shadmehr and Wise, 2005), as well as in other animals. But targeted limb movements can be found also in insects. An example is the optomotor response in crickets. The antennae of crickets can follow moving targets that are visually recognized (Honegger, 1981). Another example is the targeted leg movement in locusts that can be elicited by a tactile stimulus. When stimulating a locust by touching its forewing with a paintbrush, the animal will react with aimed scratching movements, usually of the ipsilateral leg (Matheson and Dürr, 2003; Page et al., 2008). In walking stick insects the swing movement of a leg aims at the current foothold position of the anterior leg (Cruse, 1979). All these aimed movements rely on a connection between sensory information and muscle activation. This mapping solves the inverse kinematic problem and therefore establishes an inverse model.

FORWARD MODELS

As mentioned, motor control in general requires feedback information to guide a movement. The whole cycle of motor control, for example the movement to a target, is affected by disturbances, such as misperception of the target position or the target distance and noise in the signal conductance from sensors or toward the actuators. To counteract all these disturbances, sensory feedback is required to supervise the movement, detect deviations from the intended movements, and adjust the control signal. However, in fast movements the controller cannot rely solely on sensory feedback to guide the movement because of delay inherent to the sensory and motor pathways. The question arises: how it is possible that humans as well as other animals actually are capable of such fast movements? A possible solution is that humans predict sensory consequences instead of waiting for their real values. Therefore, control of movements, in particular fast movements, relies crucially on the ability to predict sensory and motor consequences.

A solution for a fast prediction of the real feedback could be provided by a forward model (Miall et al., 1993; Desmurget and Grafton, 2000) as forward models can be used to determine spatial location when joint angles are given. Combined with an inverse model of the body, a forward model can detect a possible error more quickly than one that relies only on proprioceptive feedback. When participating in dynamical tasks, such as catching a ball, an actor must be able to predict the movement of target objects, and therefore must have a forward model of parts of the world that forecasts future states from the current state.

Today, there are many lines of evidence supporting the existence of such models in the brain. Especially for manual or bimanual tasks in humans, much work has been devoted to the influence of prediction on control tasks (Wolpert and Ghahramani, 2000; Wolpert and Flanagan, 2001). An experiment by Strauss and Pichler (1998) suggests that the fruit fly *Drosophila* is able to construct a dynamic representation of a steadily moving optical pattern that disappears behind an occluder. As a consequence, the pattern is expected to appear again on the other side of that occluder. Li and Strausfeld (1999) have found evidence suggesting that the mushroom bodies in crickets differentiate between stimulation as a consequence of intended motor actions and stimulation as an external imposed stimulation. Webb (2004) reviews further examples that involve predictive models and could be termed forward models, such as those that stabilize the visual field in flying insects.

SENSOR FUSION

A distinctive feature of animals and humans is the large number of sensors for each modality. This multitude of sensory channels is in sharp contrast to technical systems, which usually use only a handful of different sensors measuring disjunct qualities. In animals, many sensors measure the same or closely related features of the environment, but in different ways.

Each sensory channel may employ its own way of "representing" information. For example, a position of an arm may be described by the visual system in a Cartesian and body-centered coordinate system, while proprioceptive sensors use some kind of muscle length or joint angle-like representation.

A recent review Makin et al. (2008) concluded that a representation of the hand's position relies on sensory information coming from skin, joints, muscles, eyes, and even ears (Ernst and Banks, 2002). An advantage of redundant systems is that errors due to inconsistencies or to loss of sensors can be canceled out and variances can be compensated for. This presupposes an integration of the sensory information. The integration seems to be realized as a weighted summation of the different information (Makin et al., 2008).

Quite similar results can be found for targeted limb movements in insects. Niven et al. (2010) have shown that desert locusts use vision as well as tactile information from the antennae to guide where they put their limbs when walking on a horizontal ladder. In this situation, the animals are required to make accurate targeted leg placements on rungs to find a foothold, especially when the distance between rungs is variable. On the one hand, the animals directly find footholds for the front legs even when they have not touched the rung with their antennae. The visual information is in this case sufficient. On the other hand, leg placement in insects is strongly influenced by tactile information from the antennae, which is used in searching movements to find footholds for the legs (Dürr and Schütz, 2011). Locusts with occluded eyes are still able to walk over the ladder. Importantly, a deterioration in either modality has a corresponding deterioration in ladder-walking performance.

As mentioned, multiple redundant modalities in a system compensates for errors and disturbances. This, however, presupposes some kind of integration mechanism of the sensory information (see, e.g., Wolpert et al., 1995; van Beers et al., 2002). Such an integration of visual and proprioceptive/tactile information (Botvinick and Cohen, 1998; Muller et al., 2009) requires an internal model of parts of the body, which may be termed a sensor-fusion model and can apparently be found even in animals like insects (Wessnitzer and Webb, 2006).

POSSIBLE NEURONAL ARCHITECTURES

How might such models be coded neuronally? Recent studies have shown that neuronal systems controlling behavior are constructed in a modular fashion. Flash and Hochner (2005) have reviewed results that lead to the interpretation that "many different movements can be derived from a limited number of stored primitives." Davidson and Wolpert (2004) demonstrate that internal models underlying grasp can be additively combined. Results of Cothros et al. (2006) suggest that there are distinct neural representations of objects and limb dynamics. Briggman and Kristan (2008) review the arguments for modular architectures, concentrating on the question concerning functional vs. morphological modules. Anderson (2010) reviews a huge body of results supporting the idea of "neural reuse," i.e., the hypothesis that new modules have been evolved by "massive redeployment" of earlier existing modules.

Specifically, Wolpert and Kawato (1998) proposed a modular architecture, where an individual model is required for each task and each behavioral element. In this approach, not only are predictive and control functions separated, but dedicated modules are used in the context of single behaviors (Wolpert and Kawato, 1998). Such an approach requires a large number of specialized and redundant modules, and excludes the possibility of transferring knowledge between different contexts, e.g., adapting only once to changes of the body geometry or the inclusion of tools into a bodily representation (Maravita and Iriki, 2004).

In contrast, we argue that this type of specialization is not necessary and propose another approach. As each behavior has to be performed with the body, why should separate body models be applied for each of these many procedures? We propose one holistic model that, on the one hand, addresses both control and predictive function, and, on the other hand, which is one core representation that can be recruited by different behaviors and has not to be remodeled in each and every behavior anew. First, we will explain the structure of our model, which is realized as a recurrent neural network (RNN) allowing for pattern completion (Schilling, 2011a). Therefore one and the same model can be applied as an inverse model, for sensor fusion as well as a forward model, i.e., for prediction. An important characteristic of this model is that it can deal with redundant structures, in our case a complex body with 22 DoF arranged in series or in parallel. Complex redundant manipulators are a challenge for many modeling approaches as redundancy allows for multiple solutions and requires some form of decision which solution to choose. For example, the human arm consists at least of seven DoF. Many points close to a person can be reached by many different arm configurations. Instead of introducing an explicit criterion for selecting one solution, in our approach the redundancy is exploited. The complexity of the body is divided into trivial relationships and the Mean of Multiple Computation principle is a mechanism to integrate these multiple relationships. We will not refer to biological structures that possibly reflect this network. Rather, we will use it as a simple example providing a proof of concept for an integrative model that does not need a huge number of dedicated modules. We will also not discuss how the internal body model as such could be learned.

Second, we will explain and discuss how this internal model can be combined with a decentralized architecture consisting of sensorimotor procedures, i.e., be embodied in a biologically inspired control framework for the control of a walking robot (Schilling and Cruse, 2008, submitted). On the one hand, the body model serves reactive control, i.e., the network is applied as an inverse model for the control of the leg movements and as a filter to improve erroneous sensory data. On the other hand, we want to explain how the predictive capabilities of the network can be exploited to anticipate consequences of the application of novel or existing behaviors in – possibly harmful or dangerous – situations. This faculty allows the system to mentally simulate an action before carrying out a possibly unsuitable action in reality. In this way, predictive capabilities of a model can make cognition as planning ahead possible (following the definition of McFarland and Bösser (1993). How the complete model might be used for planning will be discussed in Section "Conclusion and Future Work." In the Section "Discussion," we will contrast this approach with approaches in robotics and movement science that rely on a multitude of very specific internal models.

MATERIAL AND METHODS: THE MEAN OF MULTIPLE COMPUTATION MODEL

In the following, we present a holistic model that can be used in different contexts. This model solves all three problems discussed above. The model is based on an integration principle – the mean of multiple computation (MMC) principle (Cruse and Steinkühler, 1993). The general idea is that the model describes relationships between body parts and that these kinematic descriptions are encoded into a RNN. Although the underlying principle of calculating a mean value between different influences is supported by biological findings on sensory integration (Makin et al., 2008), this network is not meant as a model of one specific part of the brain, nor do we propose that there is one single dedicated body modeling area. Rather, we only want to show the feasibility of such a model as a proof of concept. It is important that the principle proposed for the integration offers to merge multiple sites of information in a coherent way while addressing the three tasks mentioned.

The core of the network describes the structure of the body to be represented – the network can be directly set up from the kinematic equations. Even a simple manipulator structure (like a human arm) can be quite complex, making a direct mathematical solution impossible. This complexity is a problem for control approaches and is usually circumvented by introducing restrictions. In our model, by contrast, the redundancy of the manipulator is not seen as a problem, but is exploited. When setting up the kinematic descriptions we do not encode a complete solution for the whole structure of the body, but we divide the complexity into smaller structures, which can easily be handled mathematically. This leads to more equations than the minimum number required, but they can be solved and solutions can be found easily. Specifically, the structure is split into relationships between three variables each. A variable is either one that describes a moveable joint and the connected segment, or a newly introduced variable capturing relationships between two other variables. The variables describe local relationships (e.g., the upper arm and the lower arm are two variables that construct a local relationship and form a diagonal vector/variable which connects these two, see **Figure 1**, D_2). Finding a solution for any of these three variables is straightforward and always leads to a solution. Each variable takes part in several such local relationships (see **Figure 1**) and in the end we can derive a whole set of such local and simple equations (for the example of the arm, the derived equations are presented in the Appendix). Solving each of the equations for each variable, we get multiple ways of describing each variable through its local relationships: there are Multiple Computations for each variable. Following the MMC principle, the multiple solutions for one variable can be integrated by calculating a (weighted) mean. This leads to an iterative way of calculating new values for each variable. At the same time the set of equations can be understood as constituting a neural network. The introduction of recurrent connections dampens and stabilizes the system as it introduces low-pass properties (the equations describing the resulting network are given in the Appendix, for more details see Schilling, 2011a).

While the multiple computations appear to introduce additional but unnecessary computations, this is true only while the network is in a harmonic state, meaning all the multiple computations for one variable lead to the same result. But when, due to a disturbance, the different computations lead to different values the network basically performs a form of pattern completion. It acts as an attractor network forming an autoassociator and integrates the different solutions in a coherent way constrained by the encoded relationships. This means that the network overall settles into a state consistent with the encoded relationships that basically span the activation space of the network. In this way the network can fill in missing information or correct wrong information. By that means, it can produce solutions for the inverse, forward, or any mixed task.

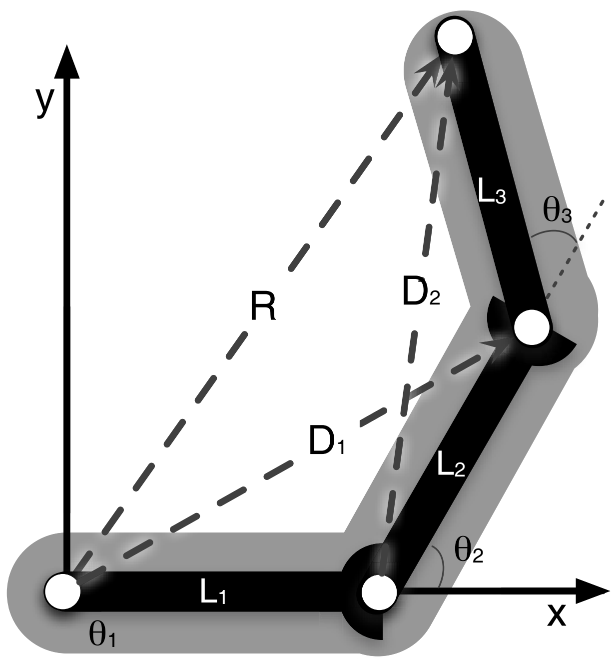

FIGURE 1 | Arm consisting of three segments (L_1, L_2, and L_3) that are connected by three hinge joints. The end-effector position is described by the vector R. D_1 and D_2 describe the diagonals. The arm can move in a two-dimensional plane, but has three DoF (joints), one more than necessary.

THE MMC BODY MODEL

In the following we want to explain how such a network can be setup as a body model for a simple animal such as a six-legged stick insect. We will start with the description of the kinematics of a single leg, which is comparable to the example of an arm. In the next step we will extend this network toward a model of the whole body, showing how different levels of representations can be integrated and how the model mediates between the different partial models. To this end, we show how this complete model can be applied in motor control and how a leg model can be utilized for the inverse model function in this task. Later, we will discuss how this model can be used for planning ahead.

The complete model has a two-layered structure (see **Figure 2**). The lower level contains six models, one for each leg (**Figure 2B**, right). The upper layer represents the thorax and the six legs, the latter, however, in an abstracted form (**Figure 2B**, left). We will begin with describing the model of the individual leg.

The leg model

Figure 3 shows the structure of a stick insect's leg that has been modeled. It only contains three DoF. We can set up a simple MMC network using redundant trigonometric relationships. Because of the kinematic structure, we can derive a specific solution for this type of manipulator. As the second and third joint act on a plane (**Figure 3C**) and their rotation axes are parallel, we can use basic trigonometric function to come up with a solution for these joint angles that hold true in this plane. The first joint angle can be derived from the projections of all leg segments on to the ground plane. Even though for this kind of structure a closed mathematical solution is possible, we restrict our solution to simple trigonometric relationships. This leads to multiple computations of the variables that can then be integrated into the model (more details on the derived equations are given in the Appendix).

As the model directly encodes the kinematic equations describing the structure of the leg, the local relationships basically represent the forward kinematics and in this way provide a means to translate movements of joints into displacements in three-dimensional space. The partial solutions are then combined through the shared connecting variables. When a set of joint values is given, the model adapts its internal values in a complementary way. The result is a leg configuration that is geometrically valid as the network activations are restricted by the encoded geometric constraints (e.g., fixed segment length, joint angle limits). This property is independent of the input given to the net being underdetermined or overdetermined.

While for a single leg the number of DoF is quite limited, the model as such is not limited in this respect and the MMC principle can and has been applied to model manipulators with many more DoF. It had been applied to three-segmented manipulators in general and it has been shown that it can be used in such scenarios

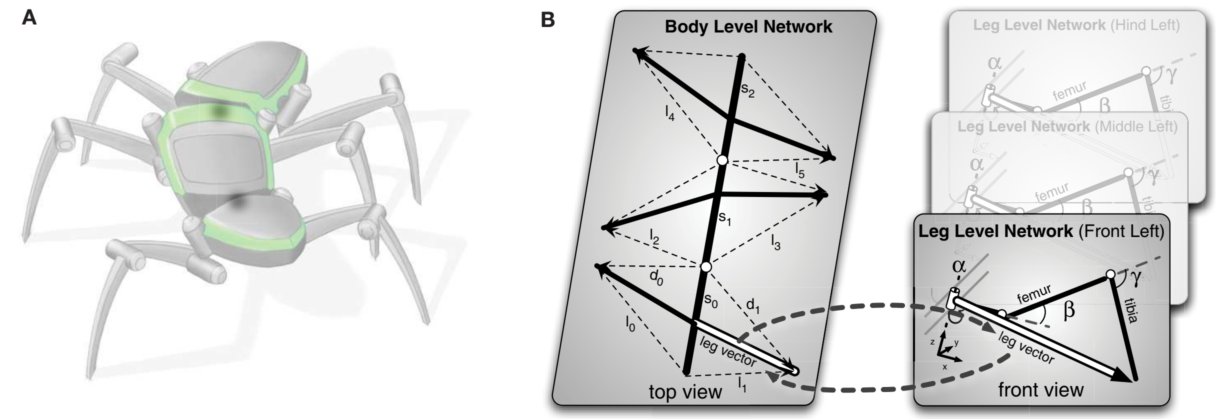

FIGURE 2 | The hierarchical body model. In **(A)** the vectors constituting the higher (body) level of the model are shown. Each leg is only represented by vectors to the end point of the leg. The detailed geometry is not reflected on this level. In **(B)** it is shown how the two levels are connected. Each leg is represented through a single leg network as shown in **Figure 3**. Each leg network shares the end point vectors with the higher body level network. During processing both levels mutually inform each other. In the inverse kinematic case the body level produces new leg vectors as target vectors that are forced onto the leg network, which comes up with corresponding joint angles for the target vectors.

with universal joints and nine DoF in total (Schilling, 2011a). The model can as well cope with additional constraints applied to it, for example, when modeling a human arm with seven DoF for the whole arm and an elbow joint that is restricted to movements in one dimension (Schilling et al., 2012).

The thorax model

When we want to look at the more complex case of a whole body, which for the insect corresponds to three body segments and six legs, we can divide the complexity of the problem into meaningful levels (**Figure 2**). To this end, the model is constituted of detailed models of the individual legs, as described above, while for the complete model of the whole body in the upper layer, the thorax model (**Figure 2B**, left), the legs are only abstracted to the vectors representing the end points of each individual leg (for more details on the representation on the body level see Schilling et al., in press). Such an approach has two immediate advantages. First, it divides the complexity into different levels and therefore reduces on each level the number of involved variables and as a consequence the number of redundant derived equations to a manageable set. Second, it introduces a form of explicit abstraction that is reflected in the structure of the model.

The different levels of the body model are connected as they share variables, in the case of the insect the vectors pointing to the tip of the leg. The computation of the different levels is tightly interwoven through these shared variables. This allows the model to be flexibly used in different scenarios. For example, we can use the body model to control the coordinated movement of the legs during the stance movement in forward and curved walking (Schilling et al., in press). In the upper level (**Figure 4A**), we initiate the movement of the body by pulling at the front segment (see **Figure 4B**, vector delta$_0$), while the other segments as well as the legs pick up the movement. Through the shared variables, the movement of the leg in the thorax model is given as an input to the leg networks and the leg networks provide the complementing joint movements for motor control.

PROCEDURAL MEMORY ELEMENTS AND MOTIVATION UNITS

The body model as such is not able to create specific behaviors. Its main function is to filter input data in such a way that the resulting output corresponds to the geometrical (and, in the extended version Schilling, 2009, dynamic) side conditions given by the body. To drive specific behaviors, a bank of procedural memories is required (see **Figure 10** for an overview of the decentralized control system for a single leg). Examples are given by a network called Walknet (Dürr et al., 2004) which, being based on behavioral studies on stick insects, produces descriptions of many complex behaviors (such as climbing over a gap that is wider than twice the step length of the animal Bläsing, 2006). The most important procedures with respect to walking concern the Swing-net and the Stance-net, controlling swing movement and stance movement, respectively. Both procedures exploit sensory feedback, joint angle position, or velocity to provide angular changes to be performed in the next moment of time. In the case of the Stance-net, the contribution of the individual joints is determined by the body model.

To control the temporal sequence underlying any behavior, for example the more or less regular sequence of swing and stance movements involved in walking, an additional neuronal structure is required. Inspired by Maes (1991), who was herself inspired by Konrad Lorenz, we equip each procedural element with a motivation unit that gates the output of the corresponding procedural element. These motivation units form a separate network as they may be coupled with mutual excitatory or inhibitory connections. This network can adopt a number of stable (attractor) states that provide the context for a specific procedure to be selected. In the examples, to allow for a simple explanation of the principle, we use only the Swing-net and a second procedure, Reach-net, explained below, together with their motivation units.

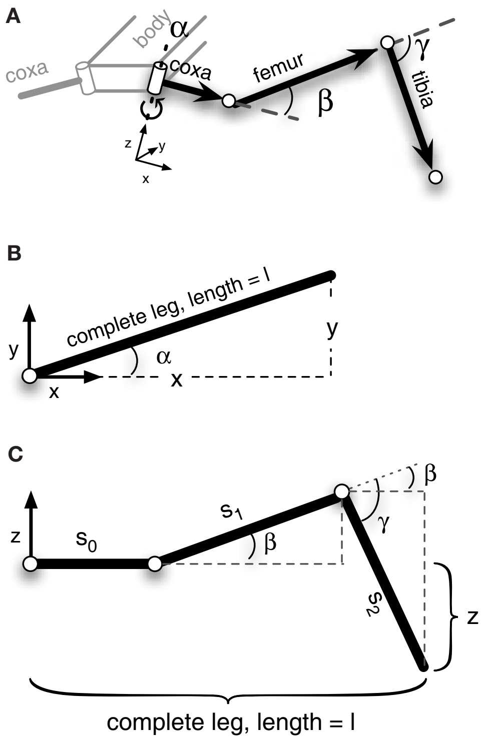

FIGURE 3 | Schematic figure of the left front leg of a stick insect. The leg consists of three hinge joints. **(A)** Shows a view of the complete leg attached to the body. **(B)** Top view of the leg. The α-joint moves the leg forward and backward. In **(C)** the leg is seen from the front. β- and γ-joint operate on a plane, meaning their axes are parallel to each other and are perpendicular to the leg plane. Lifting the leg equals a positive movement in the β-joint, and an outward going movement produced by the γ-joint is defined as a positive movement. The origins of the joint coordinate systems are set corresponding to leg positions in a standing walker (α is in a middle position, while β- and γ-joints are in a position in which the femur is approximately parallel to the ground and the tibia is nearly orthogonal to the femur).

RESULTS: SIMULATION OF THE MMC NETWORK

We will show two sets of simulations. The first one (application as a forward model) demonstrates the predictive capabilities of the MMC network. The second simulation demonstrates how the same network can be used in motor control to make targeted movements (application as an inverse model).

APPLICATION AS A FORWARD MODEL

To illustrate the basic function of the model, we will consider the scenario of a walker climbing in an environment on footholds that are sparsely distributed. Specifically, we assume that the walker is standing in front of a gap where a vertically oriented beam is positioned in the sagittal plane of the body and near enough so that the beam could be reached by a front leg (see **Figure 5**). We assume that the animal (or robot) does not exploit visual input nor does it use tactile input from the antennae. When the walker continues walking, it uses a procedural memory element called Swing-net. This network provides signals for how to move the joints during a swing movement. The latter is characterized by a trajectory that describes a movement forward that involves a lifting movement in the first part followed by a downward movement in the second part of the trajectory. In normal walking over flat terrain the swing movement ends as soon as the leg touches ground. In some versions of the Swing-net (Dürr, 2001; Bläsing, 2006), a somewhat regular searching movement is performed if no ground contact is given. During a swing of an insect standing in front of a gap, where only the vertical beam can provide ground contact, the leg may be moved until it finds a possible support at the vertical beam. Note that the body model does not contribute to control this swing movement. Nonetheless, during the swing movement, the actual values of the joint angles are given to the RNN forming the leg model, thus disturbing its actual state. As described above, this leg network starts to distribute the externally introduced disturbance onto all variables that are part of the network. As a consequence, all variables adopt values that complement the ones forced onto the leg network. As the network acts as an autoassociator, and as all the values are restricted by the encoded geometrical and kinematic structure of the modeled body, the network also contains the vector describing the end position of the leg. This information will be exploited in the second example explained below. **Figure 6** shows a simulation run in which the front left leg is making a swing movement driven by Swing-net. Shown is the real configuration of the leg as given through the joint angles (solid lines in the figure) as well as the vector pointing to the tip of the leg (dashed lines). As the figures show, the leg position estimated by the body model is quite close to the real position. Thereby the network solves the direct kinematic task.

In the example given in **Figure 6**, we showed how the body model is able to determine the end position of the leg during a swing movement. To give an impression concerning the behavior of our model, we test how well the vector pointing to the tip of the leg corresponds to the actual position determined by the joint angles. Therefore, we tested our model on a number of movements between 36 pre-defined postures (see **Figure 7**). These result from four different joint angles used for the alpha joint (87, 37, −13, −63°), three variations for the beta (15, 40, and 65°), and the gamma joint (36, 86, and 136°).

In 1260 simulation runs in all, we now produced movements from each posture to every other posture. Initially, in each run the network is provided with the joint angles of the start posture as an input and iterated for 100 iteration steps, so that the network is in a settled state and represents the start posture adequately. Then the actual test begins. For 25 time steps, each joint is moved from its start angle to the target angle. The joint angles change linearly over time and these joint values are used as input to the leg level of the body model, which is iterated as input is provided. The body model predicts the end position of the leg. **Figure 8** shows the Euclidean distance between the predicted end point and the target point over time. This distance is normalized with respect to the overall distance between the starting point and the target point. As can be seen from the figure, the body model follows nicely the imposed movement. There is an expected time lag as the used model does not anticipate the continuation of the movement, but merely integrates the current sensory data into the old estimated position and therefore underestimates the overall movement. (In an extension of the MMC network, we introduced dynamic influences and integrated equations representing velocities and accelerations in the

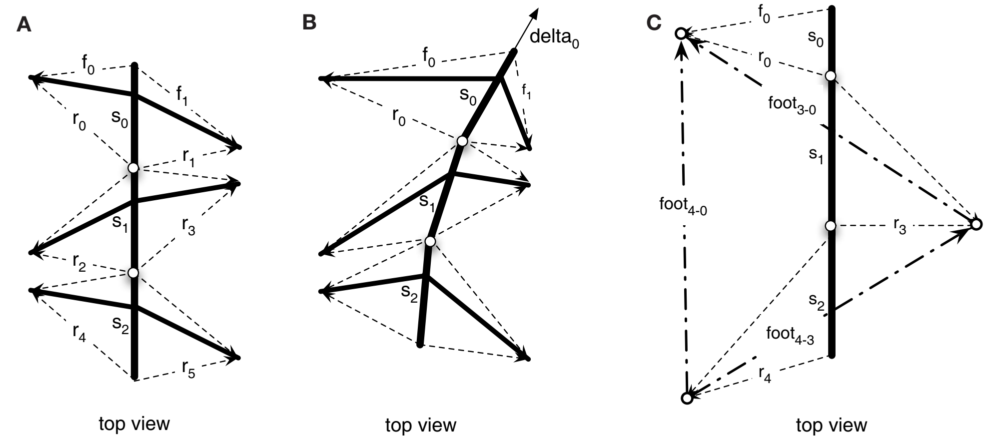

FIGURE 4 | Vectors constituting the thorax model (view from above). In **(A)** the vectors describing the foot point with respect to the segments are shown. **(B)** Shows how these vectors are changed when the model is pulled at the front (delta vector) and the foot points are kept in place. **(C)** Shows an example configuration during walking, with only three legs on the ground (front left, middle right, hind left).

network. As an effect, such a network can also successfully predict the ongoing movement and the lag is reduced correspondingly; Schilling, 2009. Including dynamic influences also counteracts the exponential slowing down at the end of the movement.) After 25 additional iteration steps the body model has settled close to the target position. The mean distance between target position as given through the joint angles and the estimated end position of the leg provided by the body model is 0.1598 (SD ± 0.112) at iteration step 25 (when the movement of the input is finished) and 0.0084 (SD ± 0.026) at iteration step 50. This is a normalized distance with respect to the overall distance between start and target position. A side effect of this normalization is that some movements that actually are quite close in three-dimensional space nonetheless require substantial movements in the joints. In such cases the normalized distance over time gets inflated by the normalization process. Looking at individual results we found that small positional differences between starting and target posture had substantially higher normalized distances, which increased the error measurement and the SD.

APPLICATION AS AN INVERSE MODEL

In the next simulation, our goal is to demonstrate how the internal body model can be used as an inverse model. We show that after the left front leg has found foot contact on the beam (see Application as a Forward Model), the contralateral, right front leg can make a targeted movement to the same spot at which the left front leg found a foothold. The left leg was driven by a simple behavioral module, Swing-net (see first simulation in Application as a Forward Model), and we used the body model to estimate its position. In the next step, the contralateral, right, leg should aim for this position. The information transfer between these two legs is mediated via the upper level of the complete body model (see The Thorax Model). Parts of the body model are vectors describing the relative position of the tips of all legs (see **Figure 4C**). For example, in **Figure 4C**, vector foot_{3-0} connects the foot of the left front leg (#0) with that of the right middle leg (#3). Correspondingly, vector foot_{1-0} (not shown in **Figure 4C**) connects the left front leg with the right front leg (#1). Therefore, to control a direct, targeted movement of the right front leg toward the current position of the left front leg, we need another procedure, termed Reach-net, that simply sets vector foot_{1-0} to zero and thereby enforces the body network to adopt a foot_{1-0} vector of length zero. In this way, the body network will generate a new target vector for the right front leg which is then given to the lower-level leg network. As the network has to satisfy this constraint, the right front leg of the model will approach the position of the left front leg, thus solving the inverse kinematic task.

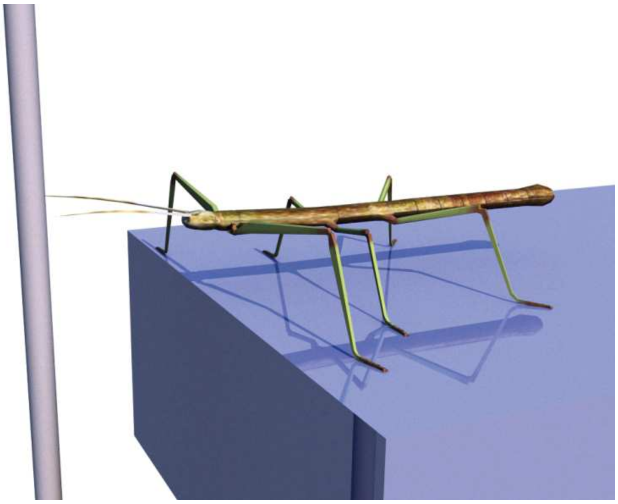

FIGURE 5 | Insect in front of a gap. Left front leg will perform a searching movement that is controlled by the Swing-net. The leg network can be used in this case to estimate the end position of the leg (forward function).

Figure 9 shows a simulation run. The position of the left leg touching the beam is given by solid gray lines. At $t > 0$ Reach-net

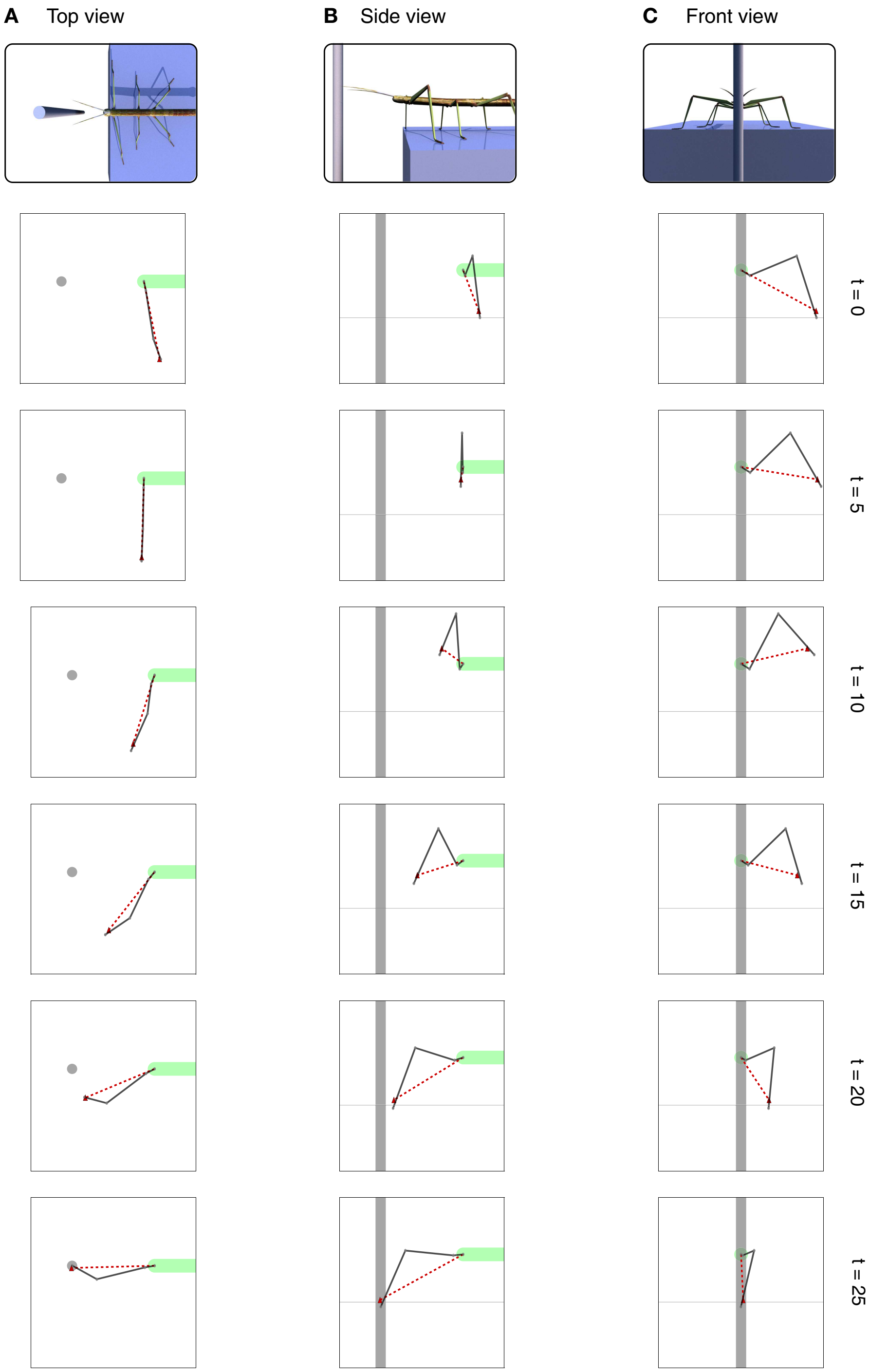

FIGURE 6 | Different views of the movement of the leg during the search movement. The dark solid line always shows the current leg configuration as described through the joint angles. The red dashed line shows the position of the tarsus as estimated by the MMC leg network. The horizontal dashed lines in **(B,C)** indicate the ground level. View from above is shown in **(A)**, side view is shown in **(B)** and view from the front is shown in **(C)**. Right: number of iterations.

is activated, which changes the target position for the right leg to the current position of the left leg. This change in target position is mediated by the upper level of the body model and depicted by the dashed line. As a consequence, the right leg (dark solid lines) is reaching for the target position and is therefore moved into the direction of the target position. The leg is moved to the

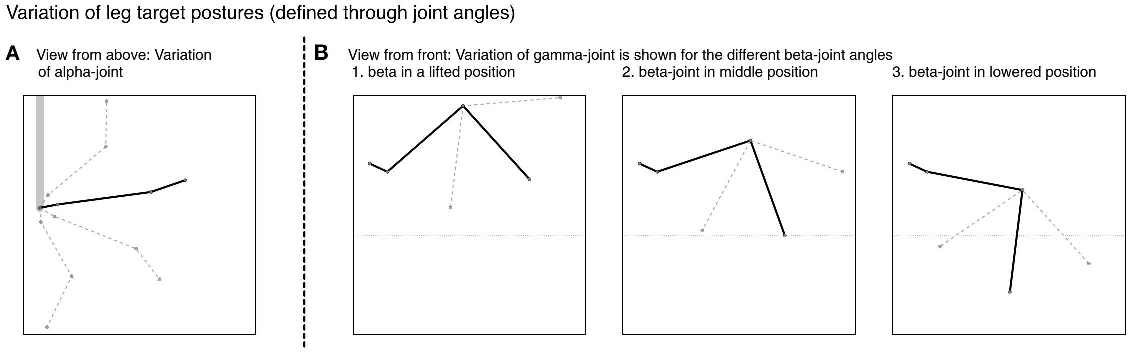

FIGURE 7 | Different leg postures – produced through variation of joint angles. In **(A)** the four different alpha joint angles used for the definition of the postures are shown. In **(B)** the three different figures show the different postures stemming from the variation of the beta joint, each showing the three gamma joint values applied.

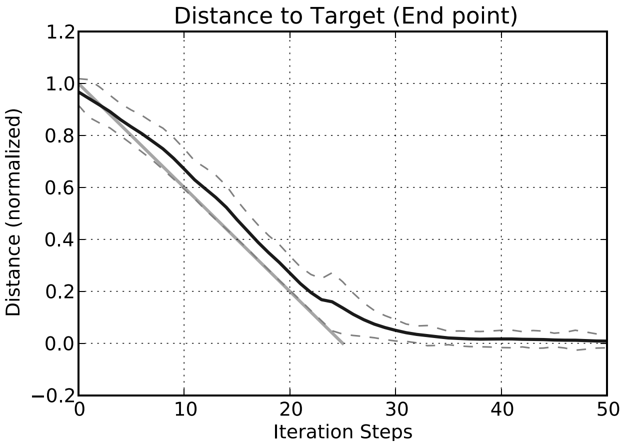

FIGURE 8 | Distance of the estimated leg end position to the position of the target posture in three-dimensional space over time. The mean normalized distance is calculated for all 1260 movements for each iteration step. The distances are normalized with respect to the distance between start and end posture in three-dimensional space (dashed lines show the SD around the mean value). The light gray line indicates a linear interpolation between the start and the target position (Importantly, the interpolation is done in joint space with a constant velocity. As a result, the interpolation of a single movement is better described by a curve (a geodesic), but over all movements we use a straight line as a simplification to indicate the general expected movement characteristics).

front through a movement of the first (the alpha) joint and then reaches out to the target position by moving both the second and third joint. In a couple of iteration steps the leg closes in on the target position and touches the beam meeting the left leg. The leg network is able to provide matching joint angles for a given target position and in this way solves the inverse kinematic problem.

DISCUSSION

The MMC network can be recruited as a body model in diverse tasks as it serves different function. The body model can address the three functions of forward modeling, inverse modeling, and sensor fusion. We have used similar models in the past to solve the inverse kinematic (Schilling, 2011a; Schilling et al., 2012) and inverse dynamic (Schilling, 2009) problems. In this article we showed how the model can serve as a forward model and predict from motor commands given as joint angles (or movements) goal positions of legs in Euclidean space. In the following, we first discuss how our approach compares to other approaches employing internal models and approaches to solving the inverse and forward kinematic problem. Second, we address how the model will be embedded in our control framework reaCog (Schilling and Cruse, submitted). There, due to its flexibility, the model can serve all functions of an internal model. In particular, the predictive capabilities allow recruitment of the model in planning ahead and use of the model as a grounded internal representation to anticipate action consequences. We will discuss connections to other motor control approaches utilizing internal models for prediction in the sense of planning ahead.

INTERNAL BODY MODELS

An important notion in the context of motor control is the internal body model, a representation of an organism's own body and its environment. Even though the work on embodiment has pointed out that complex behavior is possible without an explicit representation and can rely on the "body itself as its own best model" (Brooks, 1991), the intention was not to abandon internal representations, but to focus on grounded internal representation (Steels, 2003). Following this line of research, internal models have to be in service for some lower-level function or behavior before they can be used in a different context. One important part of such a model is a model of the body (Cruse, 1999) as it provides a starting point for models of the environment, i.e., the way the environment relates to an organism's body. The MMC model is an example of such a model that, at first, can serve behavior (targeted movements), but then is flexible enough to allow for prediction and sensor fusion and in this way may be employed for planning ahead (see Internal Models Used for Planning Ahead Through Internal Trial-and-Error).

Until now, we have focused in this paper on the forward function of the model and how this allows predictions of consequences

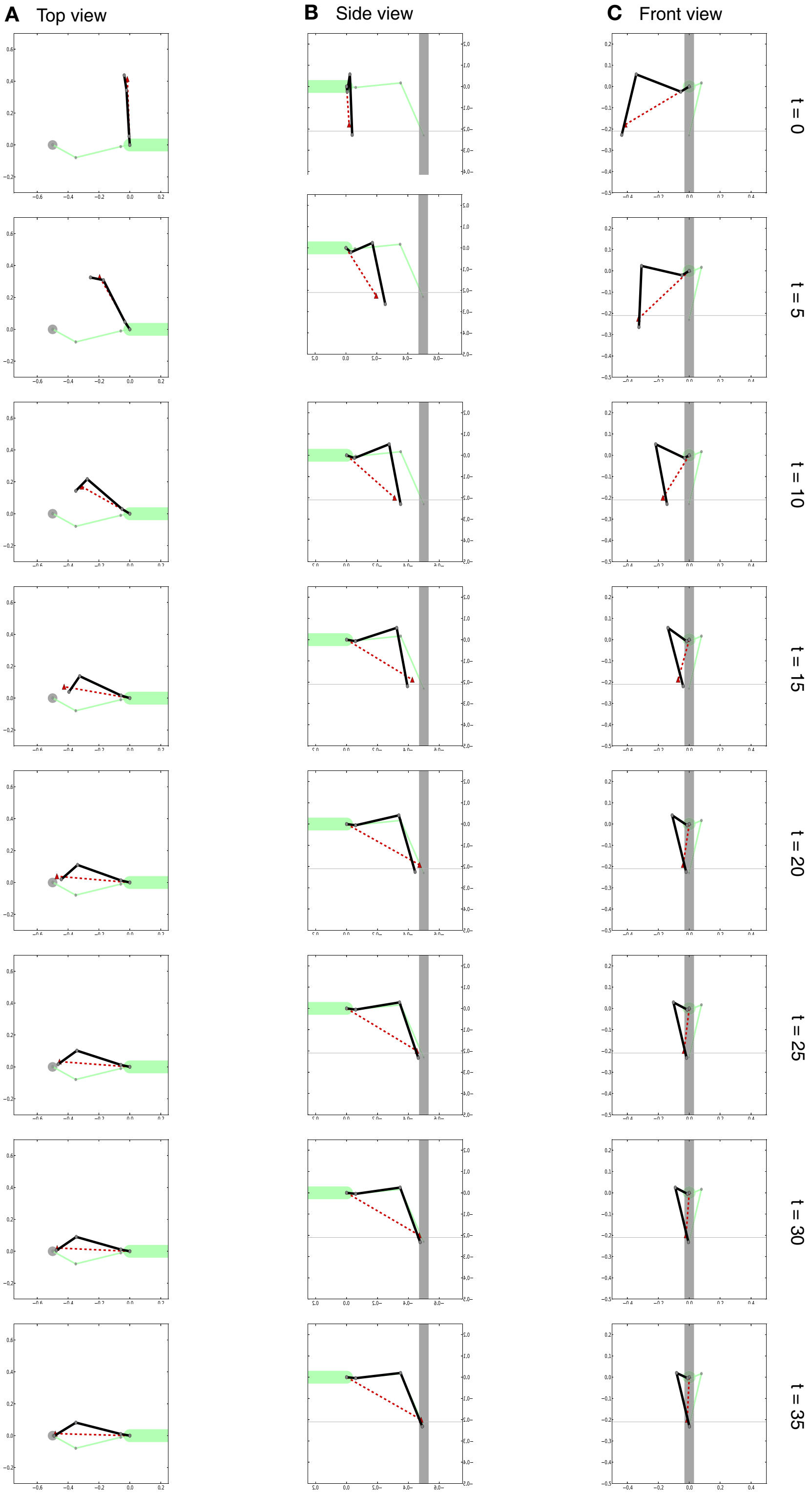

FIGURE 9 | Different views of the movement of the right front leg (dark solid lines) during the targeted movement toward the front. [Views as shown in **Figure 6**. Note that as we are looking at the right leg in **(B)** we are looking from the other (the right) side.] Shown is the movement over time. The configuration of the left leg is shown as a solid light green line. The red dashed line shows the target position provided by the body model as a target vector for the leg network of the right front leg. View from above is shown in **(A)**, side view is shown in **(B)** and view from the front is shown in **(C)**. Right: number of iterations.

of actions. In the following, we briefly discuss the properties of the proposed model with respect to aspects of internal models as raised by Haggard and Wolpert (2005). According to them, the term "body schema" stands for the unaware spatial coding of body parts (Paillard, 1999) and is comparable to our notion of an internal body model. (In contrast, the term "body image" is a visual and conscious representation of the body seen from the outside.) In the following, the different aspects (we leave out those related to phenomenological experience) are listed together with an explanation of how they refer to our MMC model:

- *Spatially coded*: The internal model represents the body and the configuration of the body. In our MMC network, the configuration of the whole body can be described by the joint configuration. The positions and relations in space result from the forward kinematic function.
- *Modular*: The brain is assumed to represent the body in a modular manner and in different neural modules (Imamizu and Kawato, 2008). The different modules must be able to interact. Hierarchical MMC networks allow a representation to be modularized easily. The complexity can be distributed on different levels of the hierarchy. The different layers of the network can cooperate by using shared variables describing their geometric relations.
- *Updated with movement*: Haggard and Wolpert (2005) demand that a body model used for the production of action has to continuously track positions and states of the body segments. It is essential for our approach to use the body model as a central part of the whole architecture. The MMC principle is basically an integration principle that allows a value for a variable to be derived from multiple values and influences. In the same way the system can be extended and used to integrate more influences and directly integrate sensory data. For a detailed discussion about how the body model can be used for sensor fusion see Schilling (2011a) or Schilling and Cruse (2008).
- *Adaptable*: Until now, the presented body model does not account for changes of the body geometry or learn even the loss of a leg. The body model is assumed to be innate and may later be modified by experiences and adapted to bodily changes (Funk et al., 2005).
- *Supramodal and interpersonal*: There are distinct areas in the brain that are responsible for processing sensory data from a single modality. The information from the different modality-specific regions is integrated by association areas (Gallese and Lakoff, 2005). The body schema is referred to as such an association area, where the integration of sensor data from different sources is an essential aspect. The MMC principle provides a basic mechanism through which multiple inputs and influences can be integrated and which could be applied there.
- Haggard and Wolpert (2005) further propose that the body scheme is not only used to represent one's own body, but also to represent the bodies of others. In a scenario with two agents we applied the body model for perception and control of action (Schilling, 2011b). One agent was making targeted arm movements using the body model to provide motor commands. The second agent observed the movements from a fixed point of view. The movements resulted in postures lying in the viewing plane of the observing agent. Lower-order visual moments were used to represent the visual input. It was the task of the observing agent to predict these visual descriptors from the current stream of sensory data. We used a RNN for this prediction consisting of one hidden layer. The structure of the hidden layer was fixed and the hidden layer was identical to the body model used for production of the movements. The observing agent was able to learn the input and output mappings in an unsupervised fashion. The dynamics of the hidden layer were exploited to reproduce the dynamics of the observed movement and to predict the movement correctly. This is a first step toward a multimodal representation. A mapping of the visual impression of another body onto one's own body model is established (Schilling, 2011b). As the body model is utilized in action and perception it provides a connection between action representation and perceptual effects as proposed by the common coding theory (Prinz, 1997).

In contrast to our approach, various authors have tried to address kinematic problems through individual models. In an early and interesting approach, Morasso and Sanguineti (1994) connected the individual models for the inverse and forward kinematic function. The output of the inverse model was routed to the forward model and vice versa. In this way, a RNN is constituted which is able to perform pattern completion similar to our approach. But it presupposes forward and inverse kinematic models, which may be hard to learn for complex structures. The advantage of the MMC approach is that it is based only on simple local relationships.

Other approaches to implementing forward and inverse functions usually separate both functions and employ independent models for each function. A classic example of such models is the MOSAIC model, which proposes pairs of inverse and forward models to represent individual motor programs. A single motor primitive (a procedural motor program representing the controller of a behavior; overall the motor primitives constitute the motor memory) is defined through the inverse model, which captures the dynamic relation between a goal state and the corresponding motor commands (Thoroughman and Shadmehr, 2000). In the case of targeted movements a goal position is described in Euclidean space and the inverse model would provide movements of the individual joints as motor commands. A motor primitive following the MOSAIC approach consists of a collection of such inverse models, each one paired with a forward model. While both models can be learned at the same time, the main function of the forward model is to offer a prediction of the currently issued motor commands. This prediction can be, first, used as a prediction of the slower sensory feedback. Second, the prediction can be later compared to the actual feedback the system receives. When the predicted value and the actual feedback are in good agreement, the respective model is modeling the current behavior well. Because in the MOSAIC framework these pairs of models are used in parallel and predictions are derived for all forward models, the comparison can be used to choose the current behavior. Therefore, the advantage of such pairs of forward and inverse models as well as learning them in combination is that the switching of motor primitives can be directly linked to the motor primitives themselves. Each motor primitive provides a measurement of how good that behavior fits the current context.

This is in contrast to the architecture we use, in which all motor primitives compete in a winner-take-all fashion (on the level of motivation units) about which primitives should be active, although merging of procedures is not precluded. The activation of a motor primitive is given through the situational context that depends on current sensory states and the current internal state of the system. One important problem for control frameworks in general is adding new behaviors. In the case of the MOSAIC controller, it is hard to decide when a behavior should be regarded as a new behavior or when it should just be understood as a variation of an existing behavior (e.g., reaching in a different direction). While this problem holds true in the same way for our approach, in the abstraction we introduced through the higher level of motivation units, quite complex and adaptive motor primitives may be built on top of the lower level, which can simply be separated by sensory signals.

In the DAC series of robots, Verschure et al. (2003) introduced a hierarchy of abstraction levels similar to ours. In his approach, the lower-level motor primitives were learned together with a more high-level and abstract representation that basically defines in which context a behavior should become active. Learning a motor primitive would be possible in the same way for our system, but currently our system consists of a pre-defined set of motor primitives that are biologically inspired from experiments on the walking of stick insects.

A serious disadvantage of the MOSAIC control framework compared to our approach is the enormous redundancy of the information. For each behavior a new pair of forward and inverse models has to be learned. Each of these models has to incorporate all the aspects required by Haggard and Wolpert (2005) as listed above, i.e., each model has to capture the basic geometric constraints and relationships and basic assumptions concerning the dynamics of movement. Not only would such a redundant system be unnecessary as it represents all these relationships multiple (and presumably a large number of) times, but it also would be difficult to adapt changes of body geometry as these would have to be changed in all the dependent models. In our model, changes in body geometry have to be applied only once to the system and not to each and every individual motor primitive. In addition, as argued above, it has been found that internal models are also recruited in perception (Loula et al., 2005) and therefore must be quite flexible and may not be restricted to specific body sizes.

The essential aspect of the MMC model is not constituted by the body dimensions as such, but is formed by the generic geometric relationships between body parts that hold true for other people's bodies as well. In this sense, the MMC model may only provide a core representation of the kinematic constraints that can be used by different motor primitives.

Such a core representation of the body is supported by experimental findings. A distinction between an internal model of the body's kinematics or dynamics and task- or behavior-specific models has been found by Cothros et al. (2006). In their experiments, subjects learned targeted goal-directed reaching movements while at the same time holding a robotic-device that applied novel force fields to the arm during the movement. The representation of the dynamics of the behavior appear to be separated from the representation of the body dynamics and kinematics. After adaptation to the force field subjects performed the same movements either in free space or in a null field holding the robot. Aftereffects during movements in free space were significantly smaller compared to those in a null field. Furthermore, no reduction in retention was observed when subjects returned to the force field after moving in free space. The representation of the object-related dynamics appear to be separated from the representation of the body dynamics and kinematics.

Another approach related to ours is the work of Bongard et al. (2006). These authors have used an internal model of the body in a starfish-like four-legged robot. In their system, the internal model was used in internal simulation loops to evolve locomotion controllers. The internal model was used to predict sensory consequences of the generated motor primitives and to access the quality of the resulting behaviors. After learning a suitable new locomotion motor primitive this controller was then applied to the robot itself. From the difference between the predicted outcome of the motor primitive and the result when carried out on the real robot, the system was able to bootstrap over time changes of its own structure and to adapt its internal model of the body. It was, for example, able to recognize the shortening of a leg and to change its internal body model, as well as to adapt the locomotion motor primitive. Such an updating routine of the internal model could be similarly introduced into the way we are applying our model as our model is also predictive. In Bongard's approach the internal model is predictive and the forward function of the internal model is exploited in internal simulation. In addition, the model is refined over time, but lacks the flexibility of the MMC model as it is only a predictive model that cannot be used for other tasks. In addition, it is not biologically inspired or related to cognitive function as such, but only computes the forward function. Furthermore, the robot structure used consists only of eight DoF and it is difficult to imagine how this approach could easily be applied to a system able to control complex behaviors, as is the case for the insect-inspired hexapod robot.

A different approach has been proposed by Butz et al. (2007); Herbort et al. (2010) based on the SURE_REACH model. SURE_REACH is a posture-based theory (Rosenbaum et al., 1993, 2001) in which a set of postures is stored in neural population codes. Crucial for motor control are two mappings. First, for a given goal state (a hand position) an appropriate posture or combination of postures has to be selected. This requires an inverse model of the goal space to the posture space. The activation stemming from the goal state drives the activity in the posture space. Second, the changes in activity of the posture space can be projected to motor commands. The motor commands invoke the movement and therefore a change in posture which is fed back into the system into the posture space. The SURE_REACH model has been tested for an arm with three DoF acting in a two-dimensional plane. This manipulator is redundant and one of the strengths of this approach is that it can deal with the redundancy. The SURE_REACH model is able to learn the bidirectional mapping between joint and Euclidean space in an unsupervised fashion. It provides a population coding of the sensorimotor mappings that is in good agreement with neuroscientific findings (Doya et al., 2007) and allows for goal-directed movements while avoiding obstacles. Unfortunately, the redundant coding of the complete arm comes with high computational costs as the number of DoF increases. Therefore, in a recent paper Butz and colleagues conclude that model does not scale up to the complexity of nine DoF like in a human arm (Ehrenfeld and Butz, submitted). In consequence, they developed the Modular Modality Frame (MMF) approach in which the overall complexity of the manipulator is distributed onto local relationships between neighboring segments. This is quite similar to the MMC approach as it is based on redundant local relationships. This model is used for the representation and integration of sensory data of an arm. A central idea is that there are redundant representations and that position and orientation of a limb are represented at the same time with respect to multiple frames of reference. Similar to the MMC approach, the model is modular and relies on local relationships between adjacent limbs of the arm. Relative forward and inverse kinematic transformations are computed between adjacent limbs in the model. In addition, representations with respect to a global frame of reference are continuously updated. Each frame of reference can be connected to multiple sensory inputs. The sensory inputs are integrated and as a consequence the network is able to compensate for noise. In addition, the computation of a plausibility value allows the network to account for (systematical) errors of sensors. The MMF model has been introduced to account for sensor fusion and it has been shown how the model can integrate different sensory channels as well as how it can deal with systematic failure. At the same time the MMF model is based on local computation of forward and inverse kinematic computation in a similar way as the MMC network. In the future, we want to extend our MMC model toward multiple sensory inputs and might use similar ideas to realize the sensor fusion in our model (weighting of inputs, plausibility measurement). While Kalman filters (Wolpert et al., 1995) have been widely used for sensor fusion and integrating these values into a current state, a crucial problem of the Kalman filter approach is that it relies on a minimization procedure required by the inverse modeling step for complex manipulator structures. As a consequence, not all states possible for redundant systems can be realized by the system. Only specific solutions are found (Grush, 2004). Again, approaches based on local relationships circumvent this problem.

While both the MMC model and the MMF model are based on kinematic descriptions that are used to set up the model, there are some approaches in which body models are learned as mappings from visually observed movements to motor commands. Most of these models deal with quite simple robotic structures and are applied to robot arms with a small number of DoF (for a thorough review see Hoffmann et al., 2010). One nice example is the work by Sturm et al. (2009), in which a Bayesian network is used. The network identifies the kinematics of the robot arm just through self-observation over time. The model successfully learns kinematic relationships between neighboring segments of the arm depending on the relating joint variables. Therefore, the model is – similar to the MMC model – based on local relationships that can then be combined to construct the kinematics of the whole robot arm. The local models are learned through a nonparametric regression. It is searched for a best arrangement of these models in order to represent the full system. The forward model has then been applied to predict movement consequences and derivations between prediction and observation have been used to adapt the model. In this way the model was able to adapt to changes of the robot dimensions. This shows the feasibility of learning such mappings and has been used even for a manipulator with six DoF. Nonetheless, it appears difficult to scale such an approach to more complex structures like robots that not only consist of a series of limb, but also have parallel limbs, such as a hexapod walker, as the basic considerations provided by Sturm et al. (2009) on the complexity of learning point out. For such a case at least some basic information on the structure of the robot seems necessary.

INTERNAL MODELS USED FOR PLANNING AHEAD THROUGH INTERNAL TRIAL-AND-ERROR

We have shown how a specific type of body model, forming a holistic system, can be used as a forward model and as an inverse model. Because it represents a pattern completion system that is restricted to geometrically consistent output vectors (i.e., body configurations), the MMC model can likewise be used for sensor fusion. Forming a redundant representation, the model is able to distribute large errors over the whole system, thus decreasing the effect of the errors. This faculty will not be discussed here further, however.

Instead, we will point to the fact that the property of this body model to act as a forward model can also be exploited for prediction. Whereas the term "prediction" usually describes the ability to provide expected sensory signals that can then be compared with actual sensor values (allowing, e.g., for correction of errors in the model), here we address another property. Internal simulation can also be used for prediction of "higher-level" expectations, for example, whether in a specific situation walking can be successfully continued. Together with the ability to exploit various elements of the motor memory, new kinds of procedures could be tested through internal simulation on being successful or not, thus allowing for the faculty of "internal trial-and-error." In this sense, Schilling and Cruse (2008) have proposed a way of using the body model in cooperation with a procedural memory. Performing internal simulation is possible within this architecture when the output of the complete motor controller is not given to the body, i.e., the muscles (or in the case of a robot, the motors), but is instead directly projected back to the input of the body model. How this could be done is schematically illustrated in **Figure 10**. Only two procedural elements of one leg, the Swing-net and the Stance-net (for details see Dürr et al., 2004) of the right front leg, are depicted. The function of Swing-net has been explained earlier. The Stance-net is very simple as it contains only three Integral controllers, one for each joint. The reference values for these controllers are provided by the joint angles determined by the leg network (see The Leg Model). During normal walking, the output of these networks drives the leg muscles, as shown by switch 1 being in position 1. Proprioceptive feedback from the legs is given to the body model (switch 2 in position 1) which in turn provides information on joint angles to the procedures (Swing-net, Stance-net) thus closing the loop through the world. To allow the system to internally simulate a behavior, in our example simulate various

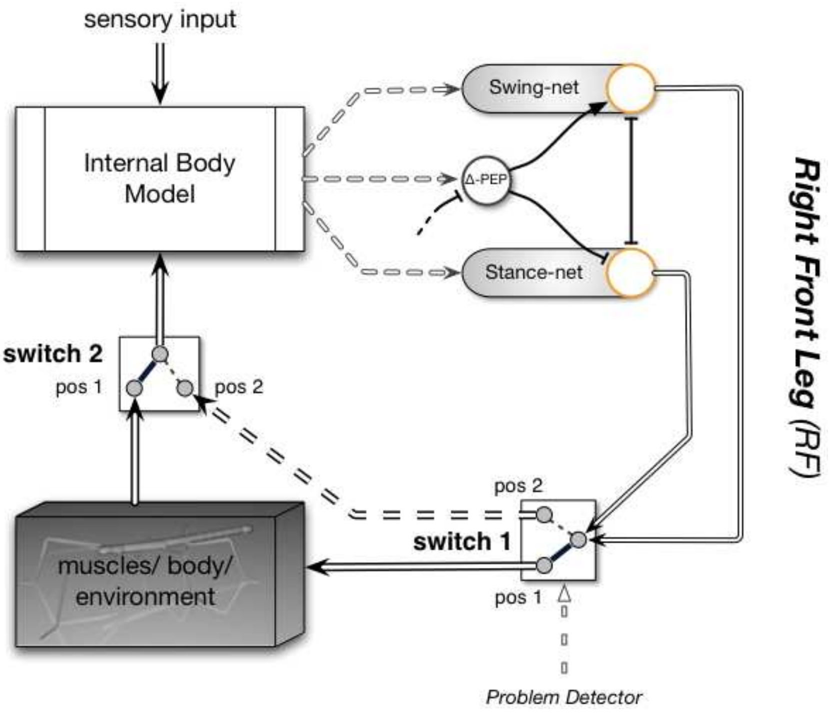

FIGURE 10 | The first step from the reactive architecture for the six-legged walker to a cognitive architecture: the controller includes an internal body model which is used for sensor fusion (and can be used to produce the trajectories for movements such as those during the stance phase – this is not shown in the figure). Only a part of the controller is shown (only some of the existing behaviors and only for the right front leg). During normal behavior, the internal body model (upper left) serves perception. Switch 2 being in pos 1 provides proprioceptive input (e.g., joint angles from the legs). The body model may also receive external sensory input (e.g., from the antennae or visual input, arrow from above). If the system runs into a problem, the body model is, together with the procedural networks (Swing-net, Stance-net), used for trying out variations of behaviors. In this case both switches are flipped from position 1 to position 2 and the motor control (double-lined arrows entering switch 1 on the right) is routed, not to the body anymore, but to the body model (dashed double line). This circuit is used for internal simulation and predicts the sensory consequences of the action. The whole process is repeated until a suitable behavior has been found. For further explanations see text.

ways of walking, both switches have to be moved to position 2. This causes the movement of the real body to stop. Instead the loop contains and drives only the body model and not the body. The more accurately the body model represents the properties of the real body (as well as selected properties of the environment, e.g., an obstacle), the better the simulated behavior corresponds to the behavior that would have been performed by the real body.

In our simulation approach, such an imagined behavior is elicited if during normal behavior a problem has been detected. A problem is characterized by a situation that cannot be handled by the currently performed behavior.

The network will, however, be able to find a solution to the problem only if the system can show some creativity. This means that new behaviors can be performed that are normally not elicited in the actual context. We assume that creativity is given by the faculty of the complete system to select new motivation units, i.e., procedural memory elements that are not activated in the actual context. We are currently working with a simple expansion of the motivation unit network to allow for creativity as characterized here (for more information see Schilling and Cruse, submitted).

If this approach turns out to be successful, we can distinguish between three levels of decision making (Cruse, 2009). The lowest level is characterized by a sensory-driven winner-take-all network, as for example is given in simple Braitenberg (1986) vehicles. The strongest sensory input determines which behavior will be performed (e.g., moving to stimulus A or to stimulus B). Noise plays merely a marginal effect as it will influence the decision only if both sensory inputs are very similar. A more complex “decision” structure would base its decision also on the current state of the system. In the case of our system the current state is represented by the motivation units as for example applied in the winner-take-all network controlling the swing-stance transitions. Both levels of decision making can be attributed to so-called reactive systems. The third level is characterized by the above mentioned system endowed with the property of internal simulation, i.e., with the ability to predict, in combination with the ability to test new behavioral solutions. As the search for new behavioral solutions, i.e., new procedures not used in the current context, is equipped with some stochasticity, an external observer cannot predict the new behavior invented by this system. This property, following the definition of Cleeremans (2005), may be characterized as comprising a volitional decision. In this way, the system is able to act in unknown and problematic situations. It is able to vary existing behaviors and, importantly, anticipates consequences of new behavioral plans before actually executing them. This allows the system to try different possible adaptations and select the one that predicts a desired outcome. The system uses the predictive capabilities of the internal model and becomes an anticipatory system following the definition of Pezzulo et al., (2008, p. 23) who provide a broad overview on anticipation and anticipatory artificial systems.

Möller and Schenck (2008) proposed another example of internal simulation to test if a specific behavior can be performed successfully. In their system, an inverse model is learned and required to suggest actions for a robot exploring a corridor with open and blocked doorways. Here, the robot acquired a forward model through active exploration. A sensorimotor representation has been constructed that is able to predict sensory consequences of a movement depending on the current state. Recruiting this forward model, the robot was able to internally simulate possible actions without actually performing them. An inverse model for selection of a suitable action was learned through ongoing internal simulations. Essentially, this model takes into account projected sensory consequences and only suggests actions that appear suitable in the long-run. Möller and Schenck (2008) relate this sensory representation to Gibson's theory of affordances, which states that an object is not represented simply by what can be sensed, but in the way the object relates to the robot (Gibson, 1979). The approach of Möller and Schenck (2008) learns an inverse model for the selection of appropriate action commands. Importantly, the possible commands are quite simple and elegant, but, due to its simple body structure, do not need to involve sophisticated control of a complex robot consisting of multiple parallel and serial joints. For the case of a hexapod walker with many (22) DoF this will become much more complicated as the computation of inverse models itself has shown to be problematic in such cases. This is especially true because this computation is closely intertwined with the sensory representation and the prediction of the sensory values. We assume that only a larger structure like the MMC model proposed here, which tightly integrates inverse and forward models, allows exploiting the flexibility of the internal model to play around with variations of existing behaviors, and to come up with new behaviors that can be tested in internal simulation.

A number of articles address the question of planning ahead on an even more abstract level. For a typical and interesting example we will briefly refer to the work of Toussaint (2006). Starting from Hesslow's (2002) notion of internal simulation as an activation of motor structures while suppressing execution, the core idea is that perceptions can be predicted as a simulation that directly leads to perceptual consequences. Central to their system is a sensorimotor map that couples sensor and motor signals in a joint representational layer. This layer is modeled as laterally connected neural layers (there are specific layers for the sensory representation and the motor commands as well as an intermediate layer coupling the two). In the same way as in the MMC network, a current state is represented through the activation of the network. The network can be driven by activations. In this way, anticipation is realized as the shifting of activity in the network triggered by external modulations provided by the motor commands. Toussaint (2006) used this network to demonstrate planning capabilities. The task was to navigate a maze. Initially, a sensorimotor map is learned through random explorations that represent the maze environment. Afterward, in the planning phase a goal stimulus is applied to the network that represents the goal position. This activation spreads through the network constrained by the topography of the maze as represented in the different networks. Here, the back and forth between sensory and motor network basically correspond to predicting sensory consequences of motor actions. For possible movements (the way is not blocked) the motor activation is maintained and can further spread. When in contrast a movement is predicted as not possible the motor activation is inhibited and here the spreading stops. In this way the networks explore the different possibilities. Although the work of Toussaint (2006) deals with even simpler motor commands than the approach of Möller and Schenck (2008), their work shows nicely how the idea of internal simulation can be understood in terms of neural computation and can be based on the spreading of activation. At the same time it demonstrates how this relies on the close coupling of sensory and motor representation and especially that this approach requires transformation mediating in both directions.

CONCLUSION AND FUTURE WORK

Anticipation of effects of action is crucial to motor control, but it is also a prerequisite for planning (Clark, in press). We have described an approach using an artificial RNN that constitutes an internal model of the body. The model is flexible and can address diverse tasks: We have shown how it can be used in motor control for targeted movements. But the model is also predictive in its nature. It is able to anticipate the effects of action and we have demonstrated how the model can estimate the resulting posture when a movement is executed. While we focused on joint position information, an extension of the model can be used to integrate dynamic influences and control signals like velocities. Following such an approach leads to natural and biological movement characteristics (Schilling, 2011a).

The model is a holistic model and as such it can be flexibly applied in other contexts serving other functions as well. We have used the model in perception in past work and used it during the observation of movements to reconstruct the observed movement (Schilling, 2011b).

Finally, we have explained how the internal model can be introduced in a robotic control structure for a hexapod robot and have briefly illustrated how the predictive capabilities can be exploited by the system in order to anticipate the effects of action before actually carrying out an action. This allows the controller to evaluate the consequences of an action and decide against performing it when it turns out to be dangerous.

Currently, we are realizing this control structure for a hexapod walker. As of now, the body model is applied in the stance controller. It is used in a similar way as described in Section "Application as an Inverse Model" (see also Schilling et al., in press). As the body model is already part of the control loop, we are going to extend the model and introduce additional redundant sensory information that is available on the real robot. As the model realizes an integration principle it can be used to fuse the sensory information of different modalities.

The control structure will be extended as explained above to account for new problematic situations to which none of the present motor primitive can react. Due to the predictive capabilities of the body model, the body model, it can be used for internal simulation. The controller can differentiate between different alternatives and variations of behaviors using their outcomes. In this way, the system can plan its action and becomes cognitive. The system takes the outcome of action into account to decide about future action. Even though the internal model is not what has changed in the system when becoming cognitive, the internal model of the body is the central part of the cognitive system. The predictive capabilities are crucial and it is the flexibility of the proposed internal body model that allows the model to be recruited in planning ahead (Anderson, 2010).

In the future, the control structure shall learn these new successful behaviors and integrate them into the overall controller structure which means that the new behaviors will also take part in the process of action selection. The model of the body is a central representation, but it is only a starting point. Even the simple body model relates to some parts of the environment where the tarsi are touching the ground. For example, the spatial arrangement of the foot points of the body model provide a simplistic representation of the environment in a way that is relevant to the animal and it's action. Our bottom-up approach allows introduction of such higher-level representations as grounded internal models as they are not detached from the lower levels of motor control. Instead, the higher levels of representation can be tightly interconnected and directly anchored in the lower levels of body representation.

ACKNOWLEDGMENTS

This work has been supported by the Center of Excellence 'Cognitive Interaction Technology' (EXC 277), by the EC project EMICAB, the ICSI Berkeley, and the DAAD (PostDoc grant to Malte Schilling). We want to thank Martin Butz and Lei Wang, whose comments helped to improve the manuscript.

REFERENCES

Anderson, M. L. (2010). Neural reuse: a fundamental organizational principle of the brain. *Behav. Brain Sci.* 33, 254–313.

Bernstein, N. A. (1967). *The Co-ordination and Regulation of Movements*. Oxford: Pergamon Press Ltd.

Bläsing, B. (2006). Crossing large gaps: a simulation study of stick insect behavior. *Adapt. Behav.* 14, 265–285.

Bongard, J., Zykov, V., and Lipson, H. (2006). Resilient machines through continuous self-modeling. *Science* 314, 1118–1121.

Botvinick, M., and Cohen, J. (1998). Rubber hands 'feel' touch that eyes see. *Nature* 391, 756–756.

Braitenberg, V. (ed.). (1986). *Vehicles – Experiments in Synthetic Psychology*. Cambridge: MIT Press.

Briggman, K. L., and Kristan, W. B. (2008). Multifunctional pattern-generating circuits. *Annu. Rev. Neurosci.* 31, 271–294.

Brooks, R. A. (1991). "Intelligence without reason," in *Proceedings of the 12th Inter-National Joint Conference on Artificial Intelligence (IJCAI-91)*, eds J. Myopoulos and R. Reiter (San Mateo, CA: Morgan Kaufmann Publishers Inc.), 569–595.

Butz, M. V., Herbort, O., and Hoffmann, J. (2007). Exploiting redundancy for flexible behavior: unsupervised learning in a modular sensorimotor control architecture. *Psychol. Rev.* 4, 1015–1046.

Castiello, U. (2005). The neuroscience of grasping. *Nat. Rev. Neurosci.* 6, 726–736.

Clark, A. (in press). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav. Brain Sci.*

Cleeremans, A. (2005). "Computational correlates of consciousness," in *Progress in Brain Research*, Vol. 150, ed. S. Laureys (Oxford: Elsevier), 81–98.

Cothros, N., Wong, J. D., and Gribble, P. L. (2006). Are there distinct neural representations of object and limb dynamics? *Exp. Brain Res.* 173, 689–697.

Cruse, H. (1979). The control of the anterior extreme position of the hindleg of a walking insect. *Physiol. Entomol.* 4, 121–124.

Cruse, H. (1999). Feeling our body – the basis of cognition? *Evol. Cogn.* 5, 162–173.

Cruse, H. (2009). "Substrate der Entscheidungsfindung," in *Das Ich und sein Gehirn, Die Herausforderungen der Neurobiologischen Forschung für das (Straf-) Recht. Göttinger Studien zu den Kriminalwissenschaften*, ed. G. Duttge (Germany: Universitätsverlag Göttingen), 63–74.

Cruse, H., and Steinkühler, U. (1993). Solution of the direct and inverse kinematic problems by a common algorithm based on the mean of multiple computations. *Biol. Cybern.* 69, 345–351.

Davidson, P. R., and Wolpert, D. M. (2004). Internal models underlying grasp can be additively combined. *Exp. Brain Res.* 155, 334–340.

Desmurget, M., and Grafton, S. (2000). Forward modeling allows feedback control for fast reaching movements. *Trends Cogn. Sci. (Regul. Ed.)* 4, 423–431.

Doya, K., Ishii, S., Pouget, A., and Rao, R. P. N. (2007). *Bayesian Brain: Probabilistic Approaches to Neural Coding*. Cambridge: The MIT Press.

Dürr, V. (2001). Stereotypic leg searching-movements in the stick insect: kinematic analysis, behavioural context and simulation. *J. Exp. Biol.* 204, 1589–1604.

Dürr, V., Schmitz, J., and Cruse, H. (2004). Behaviour-based modelling of hexapod locomotion: linking biology and technical application. *Arthropod Struct. Dev.* 33, 237–250.

Dürr, V., and Schütz, C. H. (2011). Active tactile exploration for adaptive locomotion in the stick insect. *Philos. Trans. R Soc. Lond. B Biol. Sci.* 366, 2996–3005.

Ernst, M. O., and Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* 415, 429–433.

Flash, T., and Hochner, B. (2005). Motor primitives in vertebrates and invertebrates. *Curr. Opin. Neurobiol.* 15, 1–7.

Funk, M., Shiffrar, M., and Brugger, P. (2005). Hand movement observation by individuals born without hands: phantom limb experience constrains visual limb perception. *Exp. Brain Res.* 164, 341–346.

Gallese, V., and Lakoff, G. (2005). The brain's concepts: the role of the sensory-motor system in conceptual knowledge. *Cogn. Neuropsychol.* 22, 455–479.

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. New Jersey: Lawrence Erlbaum Associates.

Grush, R. (2004). The emulation theory of representation: motor control, imagery, and perception. *Behav. Brain Sci.* 27, 377–442.

Haggard, P., and Wolpert, D. M. (2005). "Disorders of body scheme," in *Higher-Order Motor Disorders*, eds H.-J. Freund, M. Jeannerod, M. Hallett, and R. Leiguarda (Oxford: University Press), 261–271.

Herbort, O., Butz, M. V., and Pedersen, G. (2010). "The SURE_REACH model for motor learning and control of a redundant arm: from modeling human behavior to applications in robotics," in *From Motor Learning to Interaction Learning in Robots*, eds O. Sigaud and J. Peters (Berlin: Springer), 85–106.

Hesslow, G. (2002). Conscious thought as simulation of behaviour and perception. *Trends Cogn. Sci. (Regul. Ed.)* 6, 242–247.

Hoffmann, M., Marques, H., Arieta, A. H., Sumioka, H., Lungarella, M., and Pfeifer, R. (2010). Body schema in robotics: a review. *IEEE Trans. Auton. Ment. Dev.* 2, 304–324.

Honegger, H. W. (1981). A preliminary note on a new optomotor response in crickets: antennal tracking of moving targets. *J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol.* 142, 419–421.

Imamizu, H., and Kawato, M. (2008). Neural correlates of predictive and postdictive switching mechanisms for internal models. *J. Neurosci.* 28, 10751–10765.

Kawato, M. M. (1999). Internal models for motor control and trajectory planning. *Curr. Opin. Neurobiol.* 9, 718–727.

Li, Y., and Strausfeld, N. J. (1999). Multimodal efferent and recurrent neurons in the medial lobes of cockroach mushroom bodies. *J. Comp. Neurol.* 409, 647–663.

Loula, F., Prasad, S., Harber, K., and Shiffrar, M. (2005). Recognizing people from their movement. *J. Exp. Psychol. Hum. Percept. Perform.* 31, 210–220.

Maes, P. (1991). "A bottom-up mechanism for behaviour selection in an artificial creature," in *From Animals to Animats*, eds J. A. Meyer and S. Wilson (Cambridge, MA: MIT Press), 238–246.

Makin, T. R., Holmes, N. P., and Ehrsson, H. H. (2008). On the other hand: dummy hands and peripersonal space. *Behav. Brain Res.* 191, 1–10.

Maravita, A., and Iriki, A. (2004). Tools for the body (schema). *Trends Cogn. Sci. (Regul. Ed.)* 8, 79–86.

Matheson, M., and Dürr, V. (2003). Load compensation in targeted limb movements of an insect. *J. Exp. Biol.* 206, 3175–3186.

McFarland, D., and Bösser, T. (1993). *Intelligent Behavior in Animals and Robots.* Cambridge: MIT Press.

Miall, R., Weir, D., Wolpert, D. and Stein, J. (1993). Is the cerebellum a smith predictor? *J. Mot. Behav.* 25, 203–216.

Möller, R., and Schenck, W. (2008). Bootstrapping cognition from behavior – a computerized thought experiment. *Cogn. Sci.* 32, 504–542.

Morasso, P., and Sanguineti, V. (1994). "Self-organizing topographic maps and motor planning," in *SAB94: Proceedings of the Third International Conference on Simulation of Adaptive Behavior: From Animals to Animats 3* (Cambridge, MA: MIT Press), 214–220.

Muller, C. M. P., Brenner, E., and Smeets, J. B. J. (2009). Maybe they are all circles: clues and cues. *J. Vis.* 9, 10.1–10.5.

Niven, J. E., Buckingham, C. J., Lumley, S., Cuttle, M. F., and Laughlin, S. B. (2010). Visual targeting of forelimbs in ladder-walking locusts. *Curr. Biol.* 20, 86–91.

Page, K. L., Zakotnik, J., Dürr, V., and Matheson, T. (2008). Motor control of aimed limb movements in an insect. *J. Neurophysiol.* 99, 484–499.

Paillard, J. (1999). "Body schema and body image – a double dissociation in deafferented patients," in *Motor Control: Today and Tomorrow*, eds G. Gantchev, S. Mori, and J. Massion (Sofia: Academic Publishing House), 197–214.

Pezzulo, G., Butz, M. V., Castelfranchi, C., and Falcone, R. (eds.) (2008). "The challenge of anticipation, a unifying framework for the analysis and design of artificial cognitive systems," in *Lecture Notes in Computer Science*, Vol. 5225 (Berlin: Springer).

Prinz, W. (1997). Perception and action planning. *Eur. J. Cogn. Psychol.* 9, 129–154.

Rosenbaum, D. A., Engelbrecht, S., Bushe, M., and Loukopoulos, L. (1993). Knowledge model for selecting and producing reaching movements. *J. Mot. Behav.* 25, 217–227.

Rosenbaum, D. A., Meulenbroek, R. J., Vaughan, J., and Jansen, C. (2001). Posture-based motion planning: applications to grasping. *Psychol. Rev.* 108, 709–734.

Schilling, M. (2009). "Dynamic equations in MMC networks: construction of a dynamic body model," in *Proceedings of the 12th International Conference on Climbing and Walking Robots and the Support Technologies for Mobile Machines (CLAWAR)*, Istanbul.

Schilling, M. (2011a). Universally manipulable body models – dual quaternion representations in layered and dynamic MMCs. *Auton. Robots* 30, 399–425.

Schilling, M. (2011b). "Learning by seeing – associative learning of visual features through mental simulation of observed action," in *Proceedings of the ECAL 2011*, ed. R. Doursat (Paris: MIT Press), 731–738.

Schilling, M., and Cruse, H. (2008). "The evolution of cognition – from first order to second order embodiment," in *Modeling Communication with Robots and Virtual Humans*, eds I. Wachsmuth and G. Knoblich (Berlin: Springer), 77–108.

Schilling, M., Paskarbeit, J., Schneider, A., and Cruse, H. (2012). "Flexible internal body models for motor control: on the convergence of constrained dual quaternion mean of multiple computation networks," in *Proceedings of the International Joint Conference on Neural Networks 2012*, Brisbane.

Schilling, M., Paskarbeit, J., Schmitz, J., Schneider, A., and Cruse, H. (in press). *Grounding an Internal Body Model of a Hexapod Walker – Control of Curve Walking in a Biological Inspired Robot – Control of Curve Walking in a Biological Inspired Robot.* Accepted for publication at IROS 2012.

Shadmehr, R., and Wise, S. P. (2005). *The Computational Neurobiology of Reaching and Pointing: A Foundation for Motor Learning (Computational Neuroscience).* Cambridge: MIT Press.

Steels, L. (2003). Intelligence with representation. *Philos. Transact. A Math. Phys. Eng. Sci.* 361, 2381–2395.

Steinkühler, U., and Cruse, H. (1998). A holistic model for an internal representation to control the movement of a manipulator with redundant degrees of freedom. *Biol. Cybern.* 79, 457–466.

Strauss, R., and Pichler, J. (1998). Persistence of orientation toward a temporarily invisible landmark in Drosophila melanogaster. *J. Comp. Physiol. A* 182, 411–423.

Sturm, J., Plagemann, C., and Burgard, W. (2009). Body schema learning for robotic manipulators from visual self-perception. *J. Physiol.* 103, 220–231.

Thoroughman, K. A., and Shadmehr, R. (2000). Learning of action through adaptive combination of motor primitives. *Nature* 407, 742–747.

Toussaint, M. (2006). A sensorimotor map: modulating lateral interactions for anticipation and planning. *Neural Comput.* 18, 1132–1155.

van Beers, R., Wolpert, D., and Haggard, P. (2002). When feeling is more important than seeing in sensorimotor adaptation. *Curr. Biol.* 12, 834–837.

Verschure, P., Voegtlin, T., and Douglas, R. (2003). Environmentally mediated synergy between perception and behaviour in mobile robots. *Nature* 425, 620–624.

Webb, B. (2004). Neural mechanisms for prediction: do insects have forward models? *Trends Neurosci.* 27, 278–282.

Wessnitzer, J., and Webb, B. (2006). Multimodal sensory integration in insects – towards insect brain control architectures. *Bioinspir. Biomim.* 1, 63–75.

Wolpert, D., Ghahramani, Z., and Jordan, M. (1995). An internal model for sensorimotor integration. *Science* 269, 1880–1882.

Wolpert, D. M., and Flanagan, J. (2001). Motor prediction. *Curr. Biol.* 11, 729–732.

Wolpert, D. M., and Ghahramani, Z. (2000). Computational principles of movement neuroscience. *Nat. Neurosci.* 3, 1212–1217.

Wolpert, D. M., and Kawato, M. (1998). Multiple paired forward and inverse models for motor control. *Neural Netw.* 11, 1317–1329.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 30 June 2012; accepted: 14 September 2012; published online: 08 October 2012.

Citation: Schilling M and Cruse H (2012) What's next: recruitment of a grounded predictive body model for planning a robot's actions. Front. Psychology ***3****:383. doi: 10.3389/fpsyg.2012.00383*

This article was submitted to Frontiers in Cognition, a specialty of Frontiers in Psychology.

APPENDIX

CLASSICAL MMC NETWORK DESCRIBING A THREE-SEGMENTED ARM WORKING IN A TWO-DIMENSIONAL PLANE

Figure 1 shows the manipulator that consists of three segments, upper arm L_1, lower arm L_2, and hand L_3, controlled by three joints: shoulder, elbow, and wrist joint. The shoulder is situated at the origin of the x–y coordinate system. In addition, we introduce diagonal vectors D_1 and D_2 and an end-effector vector R that points to the tip of the arm. In the following we explain how this vector graph (as shown in **Figure 1**) can be used to derive equations. The resulting set of equations constitutes a weight matrix for a RNN. First, we determine all equations formed by all possible combinations of vectors forming vector triangles: The complete graph consists of several such triangles. Each triangle is a closed polygon chain which means that the three vectors complement each other to zero.

$$\begin{aligned} L_1 + D_2 - R &= 0 \qquad \text{(A1)}\\ L_1 + L_2 - D_1 &= 0\\ D_1 + L_3 - R &= 0\\ L_2 + L_3 - D_2 &= 0 \end{aligned}$$

Any of these equations can be solved for each of the contained variables. Next, all equations determining a given variable are used to form a set of equations. In this simple example, each variable can be found in two equations. In (2) this is shown for L_1 as an example.

$$\begin{aligned} L_1 &= R - D_2 \qquad \text{(A2)}\\ L_1 &= D_1 - L_2 \end{aligned}$$

In this way, we obtain six systems of two equations each. This procedure is called Multiple Computation of the same variable.

As we are considering a dynamic system that is expressed with respect to time, all variables depend on the time. The MMC principle is an iterative procedure to calculate new values for the next time step depending on the current state. For each variable the two equations are simply integrated through calculation of the Mean value (therefore the name – MMC).

$$L_1(t+1) = \frac{1}{2}\left(R(t) - D_2(t)\right) + \frac{1}{2}\left((D_1(t) - L_2(t)\right) \qquad \text{(A3)}$$

The result is one equation describing the new value for a variable depending on a weighted sum of the current values of the other variables. These equations can be directly understood as a weight matrix for a neural network. To establish the weight matrix, the vectors have to be decomposed into their x- and y-components. This leads to a set of corresponding linear equations. In the 2D example, we get two identical nets (one for each component) and for an extension to three dimensions we only have to introduce a third network representing the z-component. The network is shown in **Figure A1**.

The introduction of recurrent connections, i.e., feeding back the current value of the variable weighted by a damping factor, leads to smoother transitions in the network. The network becomes more stable and oscillations are prevented.

$$\begin{aligned} L_1(t+1) = {} & \frac{d}{d+2} L_1(t) + \frac{1}{d+2}\left(R(t) - D_2(t)\right)\\ & + \frac{1}{d+2}\left((D_1(t) - L_2(t)\right) \qquad \text{(A4)} \end{aligned}$$

Until now, we described the simple linear version of the classical MMC approach. All variables are allowed to freely change. For the rigid segments of the arm (the upper and lower arm as well the hand) we usually want to constrain the changes of variables, e.g., the segments shall not change length or the joint movements shall be restricted. This can be easily done through introducing constraints and applying the constraints after each calculation (Steinkühler and Cruse, 1998). To evade the introduction of non-linear constraints one can also use other kinds of representation. When using joint angle representation, it is not necessary to normalize the segments length after each time step. We have shown such a solution for general movements of a nine DoF arm in three dimensions using dual quaternion representations (Schilling, 2011a). In the following, we want to derive a simpler network for the special case of the insect leg.

ANGULAR MMC NETWORK REPRESENTING AN INSECT LEG

The leg of a stick insect only consists of three DoF (see **Figure 3**). Therefore, it is possible to derive a simple MMC network using redundant trigonometric relationships. The leg model can compute inverse and forward kinematics of the manipulator.

Due to the kinematic structure we can derive a specific solution for the insect leg. As the second and third joint act in a plane and their rotation axes are parallel, we can use basic trigonometric function to come up with a solution for these joint angles (see in **Figure 3C**) which hold true in this plane. We can compute the forward kinematics for the leg. The height value directly corresponds to the z value given in the leg coordinate system:

$$z = s_1 \sin\beta + s_2 \sin(\beta + \gamma) \qquad \text{(A5)}$$

The width (as a numeric value) is given as the projection of the leg onto the leg plane.:

$$l = s_o + s_1 \cos\beta + s_2 \cos(\beta + \gamma) \qquad \text{(A6)}$$

Both values are computed as a summation of the single segment portions. From this we can derive multiple computations related to the joint angles of the second and third joint. We end up with sine and cosine expression for the angles.

$$\begin{aligned} \sin\beta &= \frac{z - s_2 \sin(\beta + \gamma)}{s_1} \qquad \text{(A7)}\\ \cos\beta &= \frac{l - s_0 - s_2 \cos(\beta + \gamma)}{s_1} \end{aligned}$$

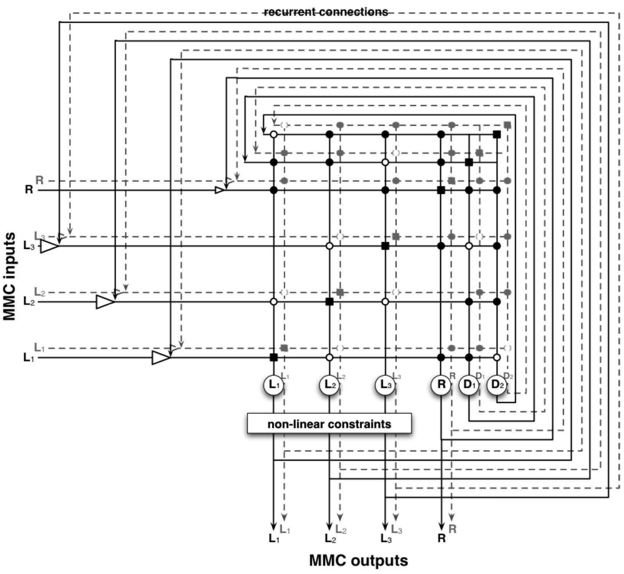

FIGURE A1 | The recurrent neural network consisting of 2 × 6 units. The network is composed of two identical networks (The one for the *x*-components is shown with black lines and the other for the *y*-components of the vectors is shown with the dashed lines). The units represent the components of the six vectors L_1, L_2, L_3, D_1, D_2, and R of the manipulator. When an input is given, the corresponding recurrent channel is suppressed (symbolized by the open arrow heads).

As sine and cosine functions are only surjective, the inverse is ambiguous and cannot directly be used to calculate the actual angles. But we can combine these and at the same time integrate the two equations by using the arc tangent function which is the quotient of the two. Again, the variables are time dependent and as we use the arc tangent to integrate the multiple computations the result is the new value for a variable:

$$\beta(t+1) = \arctan \frac{z(t) - s_2 \sin(\beta(t) + \gamma(t))}{l(t) - s_0 - s_2 \cos(\beta(t) + \gamma(t))} \tag{A8}$$

We can derive an equation for representing gamma in the same way.

The first joint is perpendicular to the other two joints. The axis of rotation lies in the leg plane and coincides with the z-axis of the leg coordinate system. Therefore, the rotation can be directly computed from the x- and y-values of the leg vector (see **Figure 3B**), showing a view directly from above).

$$\tan\alpha = \frac{y}{x} \tag{A9}$$

We can also setup an additional equation for the projection of the leg onto the leg plane.

$$\sin\alpha = \frac{y}{l} \tag{A10}$$

The multiple computations can now be used to calculate the different variables which are then integrated. On the one hand, as described above, we integrate several of the trigonometric relations into one equation through application of the arc tangent function. On the other hand, we integrate multiple computations for one variable as the computation of the mean value of the – possibly – different solutions. Here, we also include the preceding value of the variable weighted by a damping factor in order to avoid oscillations. Again, the resulting set of equations can be directly interpreted as a RNN weight matrix.

This network has several advantages compared to an explicit computation. First, the network is able to solve forward, inverse, and any mixed kinematic problems in a few iteration steps. Second, as explicit computations involve the application of the inverse of sine and cosine functions, these require case distinctions. Third, for cases where no solution is possible (e.g., the target point is too far away), the net still converges to a stable and geometric valid solution which is minimizing the error (in the example this would be the leg pointing into the direction of a far away target).

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 09 August 2012
doi: 10.3389/fpsyg.2012.00289

Perceiving one's own limb movements with conflicting sensory feedback: the role of mode of movement control and age

Lei Wang[1]*, Christine Sutter[1], Jochen Müsseler[1], Ronald Josef Zvonimir Dangel[2] and Catherine Disselhorst-Klug[2]

[1] *Department of Work and Cognitive Psychology, Rheinisch-Westfälische Technische Hochschule Aachen University, Aachen, Germany*
[2] *Department of Rehabilitation and Prevention Engineering, Helmholtz Institute of Applied Medical Engineering, Rheinisch-Westfälische Technische Hochschule Aachen University, Aachen, Germany*

Edited by:
Dorit Wenke, Humboldt University at Berlin, Germany

Reviewed by:
Martina Rieger, University for Health Sciences, Medical Informatics and Technology, Austria
Sandra Sülzenbrück, IfADo – Leibniz Research Centre for Working Environment and Human Factors, Germany

****Correspondence:***
Lei Wang, Department of Work and Cognitive Psychology, Rheinisch-Westfälische Technische Hochschule Aachen University, Jägerstrasse 17-19, D-52066 Aachen, Germany.
e-mail: lei.wang@psych.rwth-aachen.de

Previous studies have demonstrated a great uncertainty in evaluating one's own voluntary actions when visual feedback is suspended. We now compare these limitations in younger and older adults during active or passive limb movements. Participants put their dominant hand on a robot arm and performed movements actively or the relaxed limb was moved passively. Either a distorted visual feedback or no visual feedback at all was provided during the movement. Perception of limb movements was attenuated through visual feedback. This effect was more pronounced in older adults. However, no difference between active and passive movements was found. The results provide evidence for the limited awareness of body effects, even in the absence of voluntary actions.

Keywords: aging, visuomotor transformation, tool use, perception, action control, active and passive movement control, proprioception, vision

INTRODUCTION

Intentional actions commonly generate bimodal sensory effects: on the one hand the proximal, body-related action effects like the proprioceptive sensation from the required joints, and on the other hand the distal action effects, for example, the displacement of the cursor on the monitor. These sensory inputs must be monitored and integrated for online action control and error-based learning, especially in case of tool use, as demonstrated in a dual-feedback model (**Figure 1**). The execution of motor commands produces spatial displacements of the body effector (e.g., the hand) and the tool (e.g., the mouse cursor on the computer screen) controlled by the body effector. Sensory feedbacks of proximal and distal movement effects will be used to update the actual spatial configuration of the body effector and the tool. Based on these updates, new motor commands will be generated to continue the action in a modified way.

The bimodal sensory inputs are not necessarily congruent. For instance, the hand movement controlling a computer mouse causes usually larger displacements of the cursor on the monitor. Since in most circumstances the distal goals of intentional actions are represented visually, visual information should be predominant. Direct evidence of visual predominance was first provided by Hay et al. (1965). In their study a wedge prism perturbed actual hand positions. As a result perceived hand positions shifted toward the visually displayed hand positions. This effect is one example of visual capture and verified through later investigation (e.g., Pavani et al., 2000). Further evidence was found in studies focusing on adaptive movement control (e.g., Bedford, 1993). The implementation of a visual distortion is one example for establishing a novel action environment in motor control. Exposure to such distortions, for example by introducing prism goggles (e.g., Bedford, 1993; Redding and Wallace, 1997, 2006), changes in visuomotor gain (e.g., Heuer and Hegele, 2007), or visuomotor rotation (e.g., Krakauer et al., 2000), lead to visuomotor adaptation. This reflects the flexible nature of the motor control system. Perceptual processes underlying such flexibility rely on the compliance of proprioceptive sensation. For instance, spatial re-alignment in prism adaptation is based on transformation of the proprioceptive mapping to match the changed visual mapping (Bedford, 1993; Redding and Wallace, 2006). Similar perceptual processes were observed by Ghahramani et al. (1996). In a pointing task participants adapted to perturbed visual feedback of the finger, so that actual finger positions arising from the proprioceptive sensation were remapped to the visually perturbed positions. Consequently, visual dominance and compliance of proprioception are fundamental for adaptive movement control in such cases. Furthermore, proprioception is even dispensable for adaptive control (Bernier et al., 2006) as demonstrated by a deafferented patient, who adapted to a novel kinematic environment in the same way healthy subjects did. Apparently, distal representations of the movement's

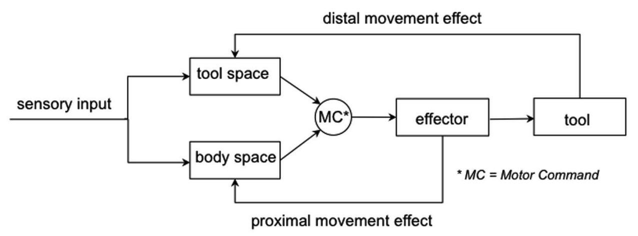

FIGURE 1 | The dual action-feedback loop of goal-oriented actions when using a tool. The motor commands launch at first a movement of the body effector, which then causes the desired displacement of the tool. Thereby the actor receives the movement feedback from the own body (proximal movement effect) and from the tool (distal movement effect).

goal controls actions. In order to maintain the flexibility of the human information processing system (visual) distal action effects are predominant while proximal action effects are attenuated.

In addition, the proprioceptive sensation *per se* may be not as precise as the visual perception. In the experiment of Van Beers et al. (1998), participants were seated at a table and had to perform position-matching tasks relying either on visual or proprioceptive information. The precision of the visual localization was between 0.2° and 0.6°, whereas the proprioceptive position sense showed a larger variance ranging from 0.6° to 1.1°. Other studies also demonstrated a great uncertainty in perceiving one's own voluntary actions when visual feedback was perturbed (Fourneret and Jeannerod, 1998; Slachevsky et al., 2001; Knoblich and Kircher, 2004; Müsseler and Sutter, 2009) or prohibited (Ghilardi et al., 1995). Taken together, the proprioceptive sensation of limb movements seems to be highly susceptible and less reliable than the visual sensation. Empirical evidence shows that humans are able to integrate multisensory signals in an optimized fashion to maximize the reliability of the perception (Ernst and Banks, 2002; Drewing and Ernst, 2006). Considering a motor action as an object of perception, integration of sensory feedback from visual and proprioceptive senses should follow the same principle. Therefore, in connection with the aforementioned lack of reliability of proprioception, it makes perfect sense that vision dominates action control, since the variance of the visual estimation is lower than that of the proprioceptive estimation.

The major question addressed in the current study is if there are any factors that moderate the bimodal integration, and consequently, affect the predominance of the visual feedback. We focused on two potential factors. A process-related factor could be the presence of motor commands. These can be understood as neural signals generated as exclusive sources of voluntary actions. Since motor commands build a link between distal and proximal action effects the movement mode should play an important role in information processing. The study by Zwickel et al. (2010) investigated whether producing active movements in a specified direction with a hand-held stylus or passive movements with the hand being transported by a robot affected the direction estimation of a concurrently presented stimulus motion. Judgments were significantly biased in the direction of the produced movement when movements were performed actively, whereas no such effect was observed for passive movements. Accordingly, we assume that the motor commands could enhance sensory integration and consequently strengthen the impact of distal feedback on proximal movement perception.

A subject-related factor could be age. Mounting evidence suggests that declines in proprioceptive function represent a fundamental aspect of the aging process (Adamo et al., 2007; Ribeiro and Oliveira, 2007; Goble et al., 2009). A variety of age-related neurophysiological changes may account for the declines in proprioception. Changes in the peripheral nervous system as potential cause are for example decreased spindle diameter, decreased sensitivity of muscle spindles, decreased number of intrafusal fibers, and a decline in the number of joint mechanoreceptors (for a comprehensive review see Goble et al., 2009). Declines in proprioceptive functions are also thought to be a result of changes in the central nervous systems, since increased proprioceptive processing demands were found to significantly impact the assessment of proprioceptive acuity in the elderly (Stelmach et al., 1990; Teasdale and Simoneau, 2001; Adamo et al., 2007). Based on these findings, we assume that the elderly would be more dependent on the visual feedback, which would then unfold its dominance more intensively.

Finally, the following hypotheses were proposed: (a) Distorted visual feedback makes movement perception more difficult. (b) Compared with younger people the older participants should show a poorer performance in perceiving their own body movements. (c) The impact of distorted visual feedback should unfold more intensively for older people. (d) Active movements should enhance the impact of the distorted visual feedback and cause poorer performance in both age groups. To examine these hypotheses, the current study compared the performance in limb movement perception of older and younger adults in various feedback (distorted visual feedback vs. no visual feedback) and movement control (active vs. passive movement execution) conditions.

MATERIALS AND METHODS

PARTICIPANTS

In total eight younger participants (five male), aged between 22 and 29 years (mean: 25 years; SD: 2.7 years) and eight older participants (four male), aged between 61 and 70 years (mean: 66.5 years; SD: 4 years) voluntarily participated. The younger participants were students of the RWTH Aachen University. The older participants were recruited from the senior-college of the RWTH Aachen University via phone calls. All of them were right-handed and had normal or corrected-to-normal vision. Participants were all neurologically intact and had no known history of neuromotor disorders. Prior to the experiments, participants signed an informed consent statement.

APPARATUS AND STIMULI

The experiment was carried out in a movement analysis laboratory using a lightweight robot LBR-IV. It belongs to a new generation of robots developed first by the German Aerospace Center (DLR). The robot presents redundant kinematics with seven degrees of freedom, allowing more complexity in the execution of the movements. Sensors evaluating the torque in each joint in real-time provide several useful features, for instance the compensation of

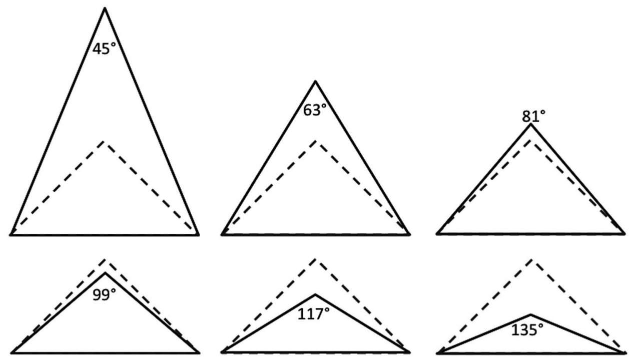

FIGURE 2 | The six standardized movement trajectories of the hand (solid lines). For visual feedback an equal-sided right-angled triangle (with dashed lines) was constantly displayed.

the gravity and accelerated reaction when the robot is submitted to external forces. The robot LBR-IV was deployed to define six standardized trajectories (**Figure 2**, solid lines) that formed either an acute ($g = 45^{\circ}$, 63°, or 81°) or an obtuse triangle ($g = 99^{\circ}$, 117°, or 135°). All triangles were isosceles with a constant horizontal base of 26 cm.

The participants sat on a chair in front of the robot arm (**Figure 3**). The chair and the robot arm stood immovable through the experiment. The right shoulder of the participant and the resting robot arm on its start location were on the same sagittal plane. The distance between the shoulder of the participant and the robot arm was approximately 70 cm. Participants put their dominant hand on the robot arm and either performed the movement actively or the relaxed limb was moved passively. Every movement was constrained within one of the six pre-defined trajectories. Short audio signals (pure tone with 840 Hz for 100 ms) were provided to mark the beginning and the end of each movement. The audio signals were clearly audible to the participants, despite the ear protection they were wearing throughout the experiment. The actual limb movement was covered by a curtain (2 m × 1.6 m) and thus, invisible to the participants. During the movement the participants either received distorted visual feedback on a LCD monitor (Eizo FlexScan L768, 19″, 75 Hz refresh rate, 1024 × 768 pixel resolution), which was positioned approximately 110 cm away and 30° left in front of the participants, or no visual feedback at all. The distorted visual feedback consisted of a cursor (a blue dot with a diameter of 3 mm) moving along the sides of a static equilateral right-angled triangle with a base of 26 cm (**Figure 2**, dashed lines), which was presented centrally on the display.

Communication between the robot arm and the feedback monitor was facilitated by a MatLab (R2009a) program on a Windows computer. The movement of the robot arm was proportionally transferred into cursor movement, depending on the ratio between the total length of the actual limb trajectory and the feedback triangle, so that the cursor appeared to be completely synchronized with the robot arm. The passive movements have pre-defined acceleration profiles and a constant duration of approximately 6 s. In order to ensure that participants in the visual feedback condition were really tracking the cursor as instructed, 10% of the trials were constructed as so-called catch trials. In catch trials the blue cursor indicating the movement brightened shortly (yellow), which should be detected and reported by the participants as a secondary task. Immediately after the completion of the movement they had to give a verbal response according to their estimation of the shape of the hand trajectory by saying "spitz" (="acute") or "stumpf" (="obtuse"). The experimenter registered the responses manually.

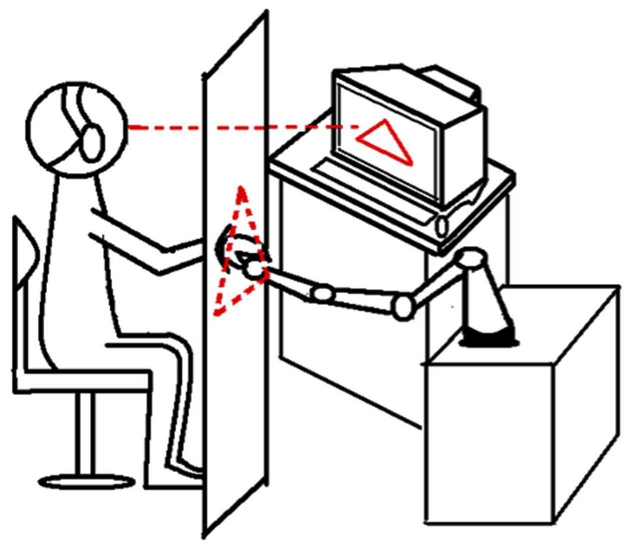

FIGURE 3 | Schematic view of the set-up. The participant is sitting in front of a curtain, putting her/his hand through it on a robot arm. With help of the robot, limb movements could be carried out passively or actively. A distorted visual feedback about the limb movement could be presented on a LCD monitor, left alongside the robot in front of the participant.

PROCEDURE

The experiment was carried out in two consecutive sessions. In Session I, a trial started with an audio signal, after which the robot arm began to move. Starting from the vertex down left, the robot led the relaxed right hand of the participant to complete one of the six standardized trajectories. After another audio signal indicated the end of the movement, participants instantaneously estimated the shape of their unseen hand trajectory.

Session I contained two blocks differing in feedback conditions: one block contained only trials with distorted visual feedback, while the other block contained only trials without visual feedback. The sequence of the blocks was counterbalanced across all participants. Every pre-defined trajectory was presented 15 times resulting in 90 trials per block. Prior to experimental trials 15 practice trials were provided to familiarize the participants with the task and its requirements. The whole session took about 60–70 min. At the end of the session participants were given a short questionnaire, in which they were asked about the strategy for making their estimations.

Session II followed the same procedure as Session I, except for the movement mode. Instead of being passively moved by the robot arm (Session I), participants had to accomplish the movement actively by pushing the robot along the pre-defined trajectories. Session II was carried out at least 6 weeks later than Session I. This quite long interval was introduced to avoid transfer effects from the preceding session.

DATA ANALYSIS

Hit rates were computed by coding the binary judgments as either correct or incorrect and calculating the percentage of correct answers. For hit rates (percentage) a 2 × 2 × 2 × 6 mixed ANOVA with the between-subject factors age (young vs. older) and movement mode (passive vs. active), and the within-subject factors feedback (distorted visual feedback vs. no visual feedback) and shape (45°, 63°, 81°, 99°, 117°, and 135°) was conducted. The second dependent variable was the area under the curve (AUC). Given the binary nature of the behavioral data and perceptual sensitivity as the underlying ability dimension, we computed a direct indicator for the perceptual sensitivity relying on receiver operating characteristic (ROC). This method is based on the signal-detection theory (for a review see Macmillan and Creelman, 2005). It provides a possibility to estimate the true sensibility of the participants, which is independent of their individual and often varying decision criteria. AUC reaches the maximum of 1, when judgments are perfect and without any error; AUC has the minimal value of 0, when the judgments are made completely randomly. Based on aggregated judgments across all stimuli, the mean AUC of ROC was calculated for each participant in each feedback and movement condition. For mean AUCs a 2 × 2 × 2 mixed ANOVA with the between-subject factors age (young vs. older) and movement mode (passive vs. active), and the within-subject factor feedback (distorted visual feedback vs. no visual feedback) was conducted.

RESULTS

Results regarding the hit rates (percentage of correct answers) showed that the accuracy of participants systematically varied with the shape of the trajectories. The stronger a movement trajectory deviated from a right-angled triangle, the easier it was for the participants to judge the movement correctly (**Figure 4**). Overall performance across all stimuli indicated that participants were remarkably uncertain about their own hand movement, especially when the distorted visual feedback was presented. The average hit rate in this condition did not exceed 77% across all stimuli, and was 10% lower than the hit rate without visual feedback. In accordance with the aforementioned comparison between both feedback conditions, a significant main effect of the factor feedback was found [$F_{(1,14)} = 19.68$, $p < 0.001$, $\eta^2 = 0.58$]. The trajectory shape (different triangles) also influenced the hit rate significantly [$F_{(5,70)} = 21.17$, $p < 0.001$, $\eta^2 = 0.60$]. And more importantly, a trend of the feedback by age interaction was observed [$F_{(1,14)} = 3.09$, $p < 0.10$, $\eta^2 = 0.173$], which was caused by a stronger decline in performance of older adults due to the distorted visual feedback. No other discernable effects were found in the ANOVA, which means that the expected main effects of age and movement mode were not observed.

Based on our hypotheses, the sensitivity of the participants, and therefore the AUCs should be influenced by feedback, movement mode, and age of the observer. The disturbance through visual feedback was statistically significant [$F_{(1,14)} = 21.18$, $p < 0.001$, $\eta^2 = 0.60$], indicating a poorer sensibility with distorted feedback ($M = 0.78$) than with no feedback ($M = 0.88$). As depicted in **Figure 5**, the impact of distorted visual feedback was tendentially more manifest in older adults than in younger [feedback by age interaction: $F_{(1,14)} = 3.45$, $p < 0.084$, $\eta^2 = 0.20$]. This finding is corroborated through independent sample t-tests (with Bonferroni correction, $\alpha_{\text{adjust}} = 0.025$), yielding a tendency for a difference between younger and older participants [$t_{(14)} = 1.83$, $p < 0.045$, one tailed], when distorted visual feedback was given. All other main effects (including the expected main effects of age and movement mode) and interactions were not significant ($p > 0.10$).

DISCUSSION

The present study aimed to examine the predominance of visual perception over proprioceptive perception of limb movement in different conditions. Binary judgments regarding movement shape turned out to be less accurate when distorted visual feedback was presented during movement execution. This impact on performance was tendentially stronger in older participants than in the young ones. No difference was observed between active and passive movement execution. The issue of age-related changes and results regarding the factor movement execution will be discussed.

AGE-RELATED CHANGES IN ACTION CONTROL AND PERCEPTION

In our study, a significant influence of distorted visual feedback on movement perception was observed. The attenuation of proximal action effects was in accordance with our assumption. This influence was obtained for both age groups and had tendentially a greater impact on the older participants. This is in accordance with our assumption that in case of distorted visual feedback the older participants should rely more on visual information, which provides apparently more reliable information and causes stronger visual capture. However, the absence of group differences in the condition without visual feedback suggests that age-related degeneration in peripheral neural structures alone cannot account for the result. As mentioned earlier, attentional processes may have played a crucial role as well. Age-related deficits in position sense, motion sense, and dynamic position sense would increase the demand for proprioceptive movement monitoring (Seidler-Dobrin and Stelmach, 1998). Consequently, the interplay of increased demand to process proprioceptive information, the decreased attentional resources in elderly (e.g., Doumas et al., 2008), and attentional distraction through visual feedback may have resulted in tendentially poorer performance of the older participants. This finding could be a possible explanation for the lack of explicit strategic action control in elderly (McNay and Willingham, 1998; Hegele and Heuer, 2010), since knowledge about discrepancy between visual and proprioceptive information about the movement is indispensable to generate appropriate control strategies.

In the absence of visual feedback, the older participants showed nearly identical performance to the young ones. This finding was not in line with a multitude of previous studies investigating the relationship between aging and motor ability (Darling et al., 1989; e.g., Cooke et al., 1989; Boisgontier et al., 2012). These studies indicated a clear decline of proprioceptive acuity in the elderly.

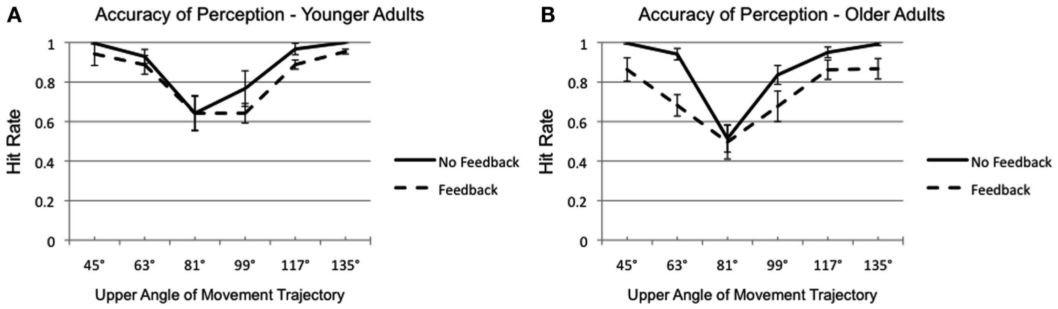

FIGURE 4 | Hit rate with (dashed line) and without (solid line) visual feedback of younger (A) and older (B) participants. Each data point represents the average of all repetitions of a certain movement trajectory. The error bars represent the standard errors.

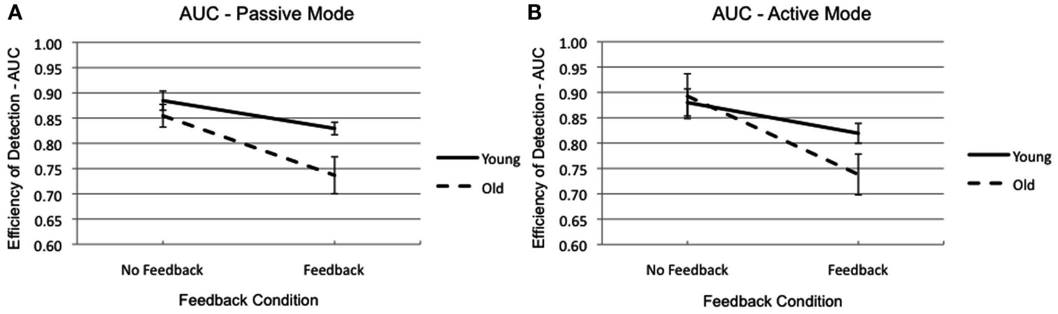

FIGURE 5 | The AUC of younger (solid line) and older (dashed line) participants depending on feedback condition and movement mode [(A) passive and (B) active]. The error bars represent the standard errors.

For example, Adamo et al. (2007) demonstrated that reproducing elbow joint positions relying only on the proprioceptive sense resulted in significantly poorer performance in older adults than in younger adults. There are several methodological reasons that could account for the absence of the expected effect related to age.

First, the task used in the current study differed substantially from those of previous studies (e.g., Stelmach and Sirica, 1986; Pickard et al., 2003; Adamo et al., 2007, 2009) examining the sense of limb position across the lifespan. These studies employed typically single joint matching tasks where the participants were required to match a memorized target joint angle in the absence of vision or to match a concurrently held limb position with the contralateral limb. In the current study the task required a binary judgment rather than a position match. The task required participants to monitor and to reconstruct the movement trajectory based on crucial movement segments, which concurrently recruited multiple joints (shoulder, elbow, and wrist). Reproducing a position may be a much more sensitive measure than giving a binary judgment. Additionally, the older participants in the current study were 66 years on average. Hence, they belong to adults at late working age. It has been argued that age-related changes in proprioception, especially in upper limb position sense are more pronounced in individuals exhibiting a sedentary lifestyle (Adamo et al., 2007), which is apparently not the case for our older participants who were students at the senior-college. Indeed, all older participants in our study reported in a pre-experimental survey that they frequently use a computer and can handle a computer mouse skillfully. It can be assumed that declines in proprioceptive functions may generally represent a fundamental aspect of the aging process, however, behavioral decline will not manifest strongly in adults at late working age, especially when they practice an active lifestyle. Taken together, the task used in the current study was probably not sensitive enough to detect age-related differences. Therefore, it remains interesting to replicate the study with participants of higher seniority and to measure additional behavioral indicators like movement reproduction.

Second, due to the small sample size the current study may have a lacked power. Evaluation of the short questionnaire, to inspect the individual judgment behavior, revealed a noticeable diversity of strategies. The participants seemed to have used very different movement cues to inform their judgments. These cues could be simple, e.g., "the height" and "the side length" of the triangle, or they could be more complex, e.g., "the ratio between height and base." Some cues were even dynamic, for instance "the acceleration at the first ascent." And some participants seemed to switch between strategies in different conditions. The large variety of strategies could have increased the variance in judgments. As depicted in **Figure 4**, the data regarding hit rates showed large variances across participants. This could have covered the age-related effects. Taking the factor age for example, the ANOVA reported in the early section yielded a $p = 0.139$ and an observed power $(1 - \beta)$ of 311, which apparently had a substantial scope for improvement with a larger sample size. Due to technical restrictions an increase of test sample was not possible for the current study. Nevertheless, we believe that the preliminary data of the current study will be confirmed by a future work with an optimized sample size.

ACTIVE vs. PASSIVE MOVEMENT

The dual action-feedback loop (**Figure 1**) suggests the necessity for the motor system to integrate bimodal feedbacks in order to control voluntary actions. Consequently, the perception of one's own limb movement is attenuated by the distal action-feedback. More importantly, the stronger the integration is, the larger the influence of visual feedback could be. Since Zwickel et al. (2010) showed that active movements could substantially enhance bimodal integration, we assumed that active movements should strengthen the impact of distorted visual feedback and cause poorer judgment performance compared to passive movements.

Contrary to our prediction, the mode of movement execution did not show any influence on the judgment. In this context, it is important to take a more comprehensive view on potential effects of active movements. On the one hand the efference copy of motor commands can directly contribute to the human-position sense (Winter et al., 2005; Gandevia et al., 2006; Gritsenko et al., 2007), and on the other hand active movement control can contribute to human-position sense by improving proprioception (Laufer et al., 2001). These findings would however lead to a contradictory prediction as we have originally made, namely improved judgment performance with active movement execution. Since the variation of movement mode did not cause any changes in the performance, it is not clear whether the mechanisms canceled each other, or rather there was no effect of the movement mode at all. The latter possibility could be due to the particular feature of the active movements in the current study. The active movement mode allowed the participants to move their dominant hand actively, however, these active movements differ from real goal-directed actions in at least two aspects. Firstly, the control of the own movement was limited to velocity and acceleration. Secondly, instead of one smooth aiming movement there were three single movement segments, one segment along each side of the triangle. Thus, constrained active movement represents only an intermediate level of motor control between pure passive movement without any control and pure voluntary action and therefore could be insufficient to enhance bimodal integration and the crosstalk of visual and proprioceptive sense. This speculation could be examined in a future study by comparing different movement modes regarding both visual capture of the proprioceptive position sense, and conversely, the repulsion effect of actual body movement on the visual perception (Zwickel et al., 2010).

CONCLUSION

The purpose of the present study was to examine the predominant role of distal feedback in both active and passive movement modes and in younger and older adults. The results supported previous observations about the limited awareness of the proprioceptive sense, and more importantly, evidence for those limitations even in the absence of voluntary actions was provided through the present study as well. Although there was a slightly stronger interference from distal action-feedback for our older participants, it is worth stressing that they, despite of expected age-related functional declines, did not show any noticeable difference in their performance compared to the younger adults, at least if there was no distracting visual feedback.

Since the coordination of perception and action is a major function in human information processing and a pre-requisite for successful interactions with our environment, it is substantial to understand how humans integrate all the information from various senses to perceive their own actions and to act adequately. Further investigations based on our findings could provide an empirical basis for various applied fields, especially for the design of tools and working environments, in which sensorimotor transformations are essential.

ACKNOWLEDGMENTS

We wish to thank Sarah Lodomez and Martin Henning for technical implementation of the experiments. We also wish to thank Miriam Tamm for advice on the data analysis. This research was supported by a grant from the RWTH Aachen Pathfinder Fund to Christine Sutter, Jochen Müsseler, and Catherine Disselhorst-Klug (MTPa16).

REFERENCES

Adamo, D. E., Alexander, N. B., and Brown, S. H. (2009). The influence of age and physical activity on upper limb proprioceptive ability. *J. Aging Phys. Act.* 17, 272–293.

Adamo, D. E., Martin, B. J., and Brown, S. H. (2007). Age-related differences in upper limb proprioceptive acuity. *Percept. Mot. Skills* 104, 1297–1309.

Bedford, F. L. (1993). Perceptual and cognitive spatial learning. *J. Exp. Psychol. Hum. Percept. Perform.* 19, 517–530.

Bernier, P. M., Chua, R., Bard, C., and Franks, I. M. (2006). Updating of an internal model without proprioception: a deafferentation study. *Neuroreport* 17, 1421–1425.

Boisgontier, M. P., Olivier, I., Chenu, O., and Nougier, V. (2012). Presbypropria: the effects of physiological ageing on proprioceptive control. *Age.* doi:10.1007/s11357-011-9300-y. [Epub ahead of print].

Cooke, J. D., Brown, S. H., and Cunningham, D. A. (1989). Kinematics of arm movements in elderly humans. *Neurobiol. Aging* 10159–10165.

Darling, W. G., Cooke, J. D., and Brown, S. H. (1989). Control of simple arm movements in elderly humans. *Neurobiol. Aging* 10, 149–157.

Doumas, M., Smolders, C., and Krampe, R. T. (2008). Task prioritization in aging: effects of sensory information on concurrent posture and memory performance. *Exp. Brain Res.* 187, 275–281.

Drewing, K., and Ernst, M. O. (2006). Integration of force and position cues for shape perception through active touch. *Brain Res.* 1078, 92–100.

Ernst, M. O., and Banks, M. A. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* 415, 429–433.

Fourneret, P., and Jeannerod, M. (1998). Limited conscious monitoring of motor performance in normal subjects. *Neuropsychologia* 36, 1133–1140.

Gandevia, S. C., Smith, J. L., Crawford, M., Proske, U., and Taylor, J. L. (2006). Motor commands contribute to human position sense. *J. Physiol. (Lond.)* 571, 703–710.

Ghahramani, Z., Wolpert, D. M., and Jordan, M. I. (1996). Generalization to local remappings of the visuomotor coordinate transformation. *J. Neurosci.* 16, 7085–7096.

Ghilardi, M. F., Gordon, J., and Ghez, C. (1995). Learning a visuomotor transformation in a local area of work space produces directional biases in other areas. *J. Neurophysiol.* 73, 2535–2539.

Goble, D. J., Coxon, J. P., Wenderoth, N., Van Impe, A., and Swinnen, S. P. (2009). Proprioceptive sensibility in the elderly: degeneration, functional consequences and plastic-adaptive processes. *Neurosci. Biobehav. Rev.* 33, 271–278.

Gritsenko, V., Krouchev, N. I., and Kalaska, J. F. (2007). Afferent input, efference copy, signal noise, and biases in perception of joint angle during active versus passive elbow movements. *J. Neurophysiol.* 98, 1140–1154.

Hay, J. C., Pick, H. L., and Ikeda, K. (1965). Visual capture produced by prism spectacles. *Psychon. Sci.* 2, 215–216.

Hegele, M., and Heuer, H. (2010). Implicit and explicit components of dual adaptation to visuomotor rotations. *Conscious. Cogn.* 19, 906–917.

Heuer, H., and Hegele, M. (2007). Learning new visuo-motor gains at early and late working age. *Ergonomics* 50, 979–1003.

Knoblich, G., and Kircher, T. T. J. (2004). Deceiving oneself about being in control: conscious detection of changes in visuomotor coupling. *J. Exp. Psychol. Hum. Percept. Perform.* 30, 657–666.

Krakauer, J. W., Pine, Z. M., Ghilardi, M. F., and Ghez, C. (2000). Learning of visuomotor transformations for vectorial planning of reaching trajectories. *J. Neurosci.* 20, 8916–8924.

Laufer, Y., Hocherman, S., and Dickstein, R. (2001). Accuracy of reproducing hand position when using active compared with passive movement. *Physiother. Res. Int.* 6, 65–75.

Macmillan, N. A., and Creelman, C. D. (2005). *Detection Theory: A User's Guide*, 2nd Edn. Mahwah, NJ: Lawrence Erlbaum Associates.

McNay, E. C., and Willingham, D. B. (1998). Deficit in learning of a motor skill requiring strategy, but not of perceptuomotor recalibration, with aging. *Learn. Mem.* 4, 411–420.

Müsseler, J., and Sutter, C. (2009). Perceiving one's own movements when using a tool. *Conscious. Cogn.* 18, 359–365.

Pavani, F., Spence, C., and Driver, J. (2000). Visual capture of touch: Out-of-the-body experiences with rubber gloves. *Psychol. Sci.* 11, 353–359.

Pickard, C. M., Sullivan, P. E., Allison, G. T., and Singer, K. P. (2003). Is there a difference in hip joint position sense between young and older groups? *J. Gerontol. A Biol. Sci. Med. Sci.* 58, 631–635.

Redding, G. M., and Wallace, B. (1997). Prism adaptation during target pointing from visible and nonvisible starting locations. *J. Mot. Behav.* 29, 119–130.

Redding, G. M., and Wallace, B. (2006). Generalization of prism adaptation. *J. Exp. Psychol. Hum. Percept. Perform.* 32, 1006–1022.

Ribeiro, F., and Oliveira, J. (2007). Aging effects on joint proprioception: the role of physical activity in proprioception preservation. *Eur. Rev. Aging Phys. Act.* 4, 71–76.

Seidler-Dobrin, R. D., and Stelmach, G. E. (1998). Persistence in visual feedback control by the elderly. *Exp. Brain Res.* 119 467–474.

Slachevsky, A., Pillon, B., Fourneret, P., Pradat-Diehl, P., Jeannerod, M., and Dubois, B. (2001). Preserved adjustment but impaired awareness in a sensory-motor conflict following prefrontal lesions. *J. Cogn. Neurosci.* 13, 332–340.

Stelmach, G. E., and Sirica, A. (1986). Aging and proprioception. *Age (Omaha)* 9, 99–103.

Stelmach, G. E., Zelaznik, H. N., and Lowe, D. (1990). The influence of aging and attentional demands on recovery from postural instability. *Aging (Milano)* 2, 155–161.

Teasdale, N., and Simoneau, M. (2001). Attentional demands for postural control: the effects of aging and sensory reintegration. *Gait Posture* 14, 203–210.

Van Beers, R. J., Sittig, A. C., and Denier van der Gon, J. J. (1998). The precision of proprioceptive position sense. *Exp. Brain Res.* 122, 367–377.

Winter, J. A., Allen, T. J., and Proske, U. (2005). Muscle spindle signals combine with the sense of effort to indicate limb position. *J. Physiol. (Lond.)* 568, 1035–1046.

Zwickel, J., Grosjean, M., and Prinz, W. (2010). What part of an action interferes with ongoing perception? *Acta. Psychol. (Amst.)* 134, 403–409.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 16 April 2012; accepted: 24 July 2012; published online: 09 August 2012.

Citation: Wang L, Sutter C, Müsseler J, Dangel RJZ and Disselhorst-Klug C (2012) Perceiving one's own limb movements with conflicting sensory feedback: the role of mode of movement control and age. Front. Psychology ***3****:289. doi: 10.3389/fpsyg.2012.00289*

This article was submitted to Frontiers in Cognition, a specialty of Frontiers in Psychology.

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 08 October 2012
doi: 10.3389/fpsyg.2012.00393

Mine is earlier than yours: causal beliefs influence the perceived time of action effects

Carola Haering* and Andrea Kiesel

Cognitive Psychology Unit, Department of Psychology, Institut fuer Psychologie, Julius-Maximilians-Universität Wuerzburg, Wuerzburg, Germany

Edited by:
Dorit Wenke, Humboldt-Universität zu Berlin, Germany

Reviewed by:
Sukhvinder Obhi, Wilfrid Laurier University, Canada
Andrea Desantis, Université Paris Descartes, France

****Correspondence:***
Carola Haering, Cognitive Psychology Unit, Department of Psychology, Institut fuer Psychologie, Julius-Maximilians-Universität Wuerzburg, Roentgenring 11, 97070 Wuerzburg, Germany.
e-mail: haering@psychologie.uni-wuerzburg.de

When a key press causes a stimulus, the key press is perceived later and the stimulus earlier than key presses and stimuli presented independently. This bias in time perception has been linked to the intention to produce the effect and thus been called intentional binding (IB). In recent studies it has been shown that the IB effect is stronger when participants believed that they caused the effect stimulus compared to when they believed that another person caused the effect (Desantis et al., 2011). In this experiment we ask whether causal beliefs influence the perceived time of an effect when the putative effect occurs temporally close to another stimulus that is also an effect. In our study two participants performed the same task on connected computers with separate screens. Each trial started synchronously on both computers. When a participant pressed a key, a red and a yellow stimulus appeared as action effects simultaneously or with a slight delay of up to 50 ms. The participants' task was to judge the temporal order of these two effect stimuli. Participants were either told that one participant caused one of the two stimuli while the other participant seated at the other computer caused the other stimulus, or each participant was told that he/she caused both stimuli. The different causal beliefs changed the perceived time of the effects' appearance relative to each other. When participants believed they each caused one effect, their "own" effect was perceived earlier than the other participant's effect. When the participants believed each caused both effects, no difference in the perceived temporal order of the red and yellow effect was found. These results confirm that higher order causal beliefs change the perceived time of an action effect even in a setting in which the occurrence of the putative effect can be directly compared to a reference stimulus.

Keywords: intentional binding, causal belief, causality, temporal order judgments, TOJ, agency

INTRODUCTION

When an action triggers an effect stimulus, the action and the effect are perceived to be closer to each other in time. For example, when the time of an operant action causing a tone is estimated in relation to a revolving clock hand, the action is perceived later than a non-operant action that does not cause an effect (Haggard et al., 2002a,b; Haggard, 2005). Additionally, tone effects in the operant condition are perceived earlier than tones presented in isolation. Thus, in the operant condition action and effect tone are perceived to be closer in time than actions and tones alone.

This bias in perceived time has been termed intentional binding (IB) because the bias is restricted to conditions in which participants intentionally perform actions. Recent studies demonstrated that key presses and subsequent stimuli are perceived to be closer to each other in time when freely chosen actions produced the stimuli as their effects. However, when the participant's finger was moved by the key (Wohlschläger et al., 2003a) or the movement of the finger was triggered by a TMS signal (Haggard et al., 2002b) instead of the movement being initialized by the participant him/herself, key presses were perceived earlier and/or tomes were perceived later in these "unintentional" movement conditions, i.e., a reversed pattern of results compared to intentional movement conditions was observed.

Interestingly, IB is not restricted to own actions, but also occurs with observed actions performed by another person (Wohlschläger et al., 2003b). In their intentional observation condition participants judged the time when another person pressed a key. In the unintentional condition participants watched how a key with a rubber hand lying on the key moved downward. The action was perceived to be later in the intentional conditions than in the unintentional rubber hand condition. Thus, the perceived time of the action as a measure of IB is restricted to intentional conditions in which the observer attributes the key presses to an intentional action, even if it is only observed (see also Wohlschläger et al., 2003a). Similarly, the perceived times of actions and effects of a co-actor are closer to each other to a similar degree as those of own actions (Strother et al., 2010).

However, these results are in contrast to a study of Engbert et al. (2007) where no difference in the perceived duration of intervals was found between observed actions of the experimenter and observed key movements with a rubber hand resting on the key. This difference could arise from the different methods used, namely duration estimation and the estimation of the points in time of action and effect. It has been suggested that those methods focus differently on diverging aspects of IB (Humphreys and Buehner, 2009). The estimation of duration relies more on inferential postdictive processes while methods focusing on points in time of action and effect rely on shorter-lived predictive processes. However, this explanation is speculative and has not been directly tested.

To conclude, IB in terms of a shift in the perceived time of action and/or effect occurs for intentional movements, that is for movements that aim at producing a specific effect. IB is not restricted to own intentional actions, but it also occurs for actions of other people that the observer believes to be intentional behavior.

Given that the bias in time perception for actions and contingently following effects depends on own intentional behavior or the belief that a person behaved intentionally, one may assume that IB is stronger for own actions compared to other persons' actions. For another person's actions the intention of the actor has to be inferred, while for own actions the intention to act is an inherent predecessor (if not the ultimate cause) of the action. If IB is stronger for own actions compared to another person's actions, own action effects should be perceived earlier than action effects that are caused or at least *believed* to be caused by someone else.

In line with this reasoning, Desantis et al. (2011) showed that a tone that was caused by a participant was perceived earlier when the participant believed he/she had caused the tone than when the participant believed that the tone had been caused by a key press of another person in the room, a confederate of the experimenters. In their experiment each trial started with the presentation of either the name of the participant or the name of a confederate to inform the participant which one would allegedly be causing a tone effect in this trial. Then, the participant and the confederate pressed a key at approximately the same time. In reality it was always the participant who caused the tone to appear 350, 550, or 750 ms after the key press. After each trial, the participant indicated when he/she had perceived the tone by reporting the position of a revolving clock hand at the moment he/she had perceived the tone. As predicted, participants perceived effects earlier when they believed they had caused the effect compared to the situation when participants believed the confederate had caused the effect, demonstrating that causal belief influenced the perceived time of the effect.

In the current study we aimed at finding further support for the notion that allegedly "own" action effects are perceived earlier than allegedly "another person's" action effects by using a new design and a psychophysical method to assess time judgments instead of the clock method. In our study two participants performed the experiment simultaneously. Participants were asked to imagine that they were the security officer of a ship and had to save a passenger who fell overboard. The participants' task was to release either one or two life buoys into the water by pressing a key as quickly as possible. After pressing the key, a red and a yellow life buoy appeared. The temporal order of the two stimuli varied slightly, with a delay of up to 50 ms (varied in 10 ms steps from −50 ms to +50 ms). Half of the participants believed they took part in a shared task and that they caused an "own" single effect, e.g., the red buoy appearing, while the other participant caused the other effect, e.g., the yellow buoy appearing (single effect group). The other half of participants believed as a control group that they always caused one compound effect consisting of the two effect stimuli, i.e., the red and the yellow buoy (compound effect group). That is, in both groups each participant in reality triggered both the red and the yellow life buoy with his/her key press. However, only the compound effect group was veridically instructed that each participant would cause both effects as a compound effect with his/her key press in each trial. The single effect group believed that each participant caused one specific stimulus of the two effects.

To assess the perceived time of action effects, participants performed a temporal order judgment (TOJ) task. That is, they indicated which effect (i.e., the yellow or the red one) they perceived first in each trial. With this design we could directly compare the influence of causal belief on the temporal perception. In the single effect group one effect (e.g., the yellow one) was believed to be the "own" effect while the other effect was believed to be the "other participant's" effect. So for this group the temporal order of the red and the yellow effect directly represents the order of the "own" and the "other's" effect. The compound effect group serves as a control group to ensure that not generally the effect of one specific color is preferred regarding temporal order.

Temporal order judgments allow us to estimate IB effects with a psychophysical method. Choosing a psychophysical method also offered the opportunity to analyze not only the perceived time of the effect stimuli relative to each other, but also to compare the temporal resolution of time judgments (Nolden et al., 2012). By using this method we could test not only if the TOJs were biased by the causal belief, but also if participants were less able to distinguish the perceived temporal order of events due to this bias. Regarding the perceived time of effects, we expected that putative "own" effects are perceived earlier than effects that were believed to be caused by another person. When a participant believes he/she caused both effects, the perceived time of those effects should lie in between.

MATERIALS AND METHODS

PARTICIPANTS

Forty-eight students of the University of Wuerzburg (18 male, all right-handed) participated in the experiment due to course requirements. Participants were between 18 and 28 years old (mean 20 years). The data of one additional participant were replaced as he/she did not believe that the experimental computers were actually connected. As the experiment could only be conducted with two participants at a time, the replacing participant took part together with one further participant, whose data were discarded to maintain counter balanced conditions.

APPARATUS AND STIMULI

The experiment was run on two standard PCs equipped with 17″ CRT screens. The PCs were connected via the computers' parallel ports to synchronize the beginning of each trial. Stimulus presentation and data collection were accomplished with the software package E-Prime2 (Schneider et al., 2002).

The experimental setup is shown in **Figure 1**. The two screens were placed side by side (distance ca. 1.5 m) with a divider wall in between to ensure that participants only saw their own screen. During the main part of the experiment participants sat in front of their screen. To avoid the participant hearing the other participant pressing the key, both participants wore Vic Firth SIH1 isolation

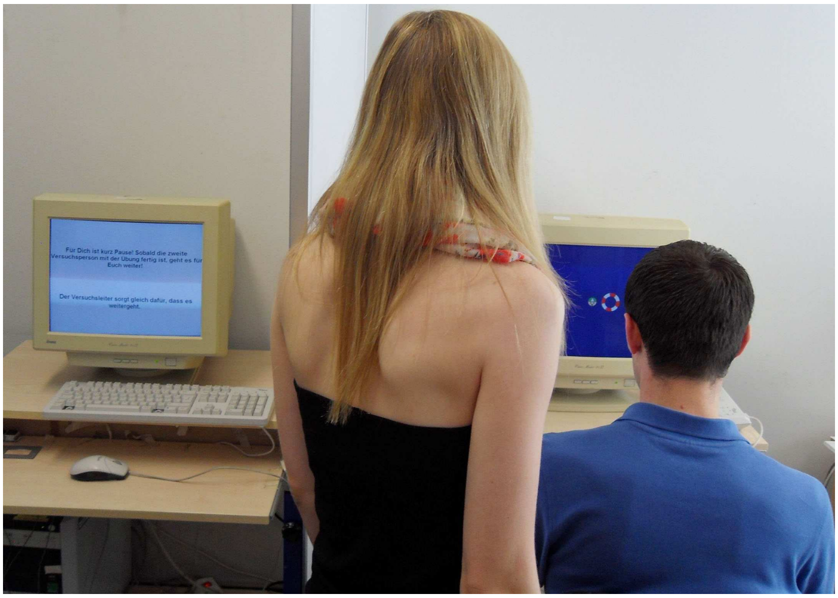

FIGURE 1 | Experimental layout. The two computers were connected via parallel port. The situation depicts an instruction block in the single effect group: The right participant performs key presses always triggering his "own" effect, here the red life buoy, while the left participant watches. Earlier the right participant had watched the left participant triggering her "own" effect, the yellow life buoy. After the instruction phase each participant performed the task on his/her computer separated by the divider wall. Throughout experimental trials both participants wore headphones delivering white noise to ensure they did not hear button presses.

headphones. The experimenter stayed in the room throughout the entire experiment to ensure that participants did not communicate with each other.

Stimuli were presented on an avy blue background. All messages were printed in white. We used a white fixation cross extending 0.7 cm. The imperative stimulus was the head of a person wearing a swim cap (diameter 1.9 cm) that appeared in the middle of the screen, described as a passenger who fell overboard. The targets for the temporal order task were a red and a yellow life buoy (diameter 3.9 cm) appearing 2 cm left or right of the center of the screen.

PROCEDURE

In each session two participants took part and were either both assigned to the single effect group or both to the compound effect group. Where not stated otherwise, the procedure was the same for both groups. All participants were asked to imagine they were a security officer of a ship who has to save a passenger who repeatedly falls overboard. The participants' task was to release one (single effect group) or two life buoys (compound effect group) by pressing the left mouse button as fast as possible. After pressing the key, the two life buoys appeared with a slight temporal delay (see trial structure described below).

In the single effect condition, each participant was told that he/she was in control over one of the two life buoys on both computers while the other participant controlled the other life buoy via the connecting cable. One participant was told to control the red life buoy (single red effect condition) and the other participant was told to control the yellow life buoy (single yellow effect condition). In the compound effect group, both participants were told to control both buoys on their own computer. The connection of the computers was explained to ensure that the experiment ran synchronously for both participants in the compound effect group. As each participant caused both effects on his screen, actually in both groups only the starting time of each trial was synchronized.

To improve the credibility of the group-specific instructions regarding who caused which life buoy to occur, participants performed an instruction phase before the main experiment. First, each participant was informed by written instructions that he/she caused either the red buoy, the yellow buoy, or both buoys to appear. Each of the two participants then performed an instruction block while the other participant stood behind and watched (see **Figure 1**). In the single effect group only the participant's "own" effect appeared randomly on the left or right side of the "drowning" passenger. In the compound effect group always both buoys appeared with the assignment of color to side of the screen counterbalanced within participants. After the first participant had completed the instruction block, it was his/her turn to watch the second participant accomplishing the instruction block.

During the rest of the experiment participants wore isolation headphones and heard constant white noise. The volume of the white noise was adjusted so that participants did not hear the sounds caused by the mouse clicks. Each trial started with the fixation cross presented centrally for 100 ms (for a schematic sketch of experimental trials see **Figure 2**). After a blank of 500 ms the imperative stimulus, the passenger, followed. The participants' task was to press the left mouse button as quickly as possible in response to the passenger's appearance. When the participant pressed the button within the time limit of 750 ms the first of the two effects (life buoys) appeared after a variable interval of between 400 and 610 ms after response onset. The second effect appeared either at

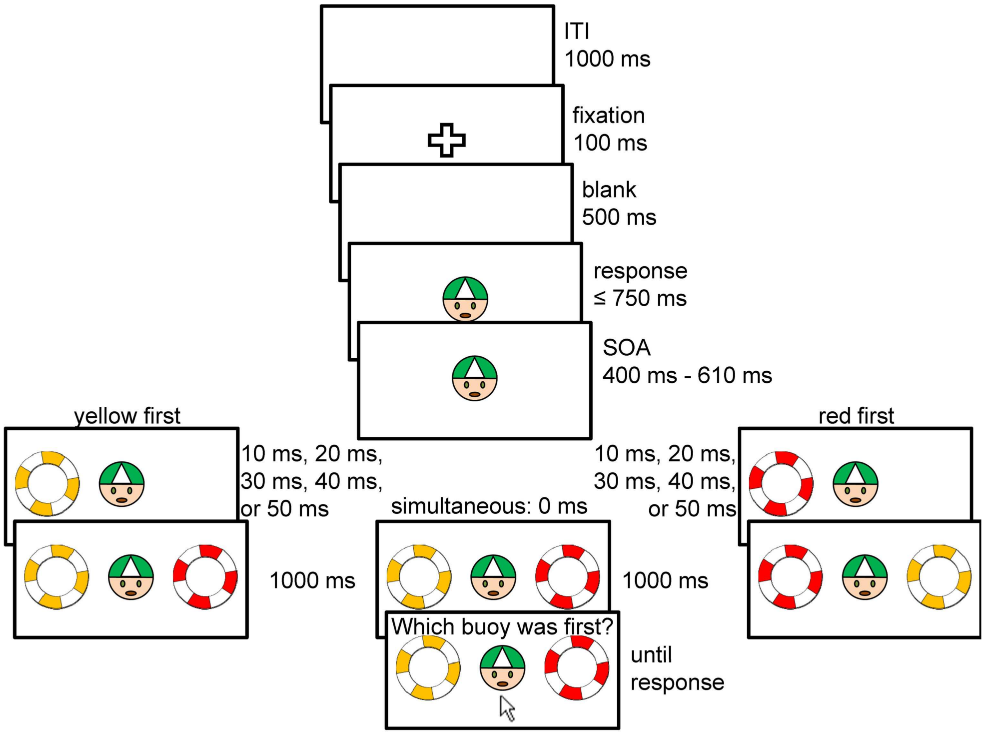

FIGURE 2 | Schematic trial procedure of the temporal order judgment (TOJ) task. Effect stimuli (the life buoys) were presented simultaneously or separated by a delay of 10–50 ms. TOJs were given by clicking on the life buoy the participant judged to have appeared earlier. The location of red and yellow stimuli on the screen was counterbalanced within participants. For better readability in the figure the background color is shown in white instead of blue in the experiment and texts are printed in black instead of white.

the same time (i.e., separated by a delay of 0 ms) or after 10, 20, 30, 40, or 50 ms. We will here after refer to these delays as temporal distance of the yellow effect in relation to the red one, that is, positive delays (10, 20, 30, 40, and 50 ms) indicate that the red effect appeared first, and negative delays (−50, −40, −30, −20, and −10 ms) indicate that the yellow effect appeared first.

After both effects were visible for 1000 ms, participants were asked to indicate which buoy appeared first (the German words "Welcher Rettungsring war zuerst?" appeared above the stimuli). For this TOJ task, the mouse cursor appeared 4.5 cm below the passenger and the participant had to click on the buoy he/she believed appeared first. After clicking on a buoy, all stimuli and the mouse cursor disappeared. The next trial started 1000 ms after both participants had clicked on a buoy.

If a participant did not press the mouse button within 750 ms after the imperative stimulus (the passenger) appeared, the passenger disappeared, and an error message reminded the participant to respond as quickly as possible to save the passenger from drowning ("Bittereagierenach Erscheinen des Passagiersim Wasser so schnellwiemöglich, sonstertrinkter!"). Participants had to acknowledge this message by clicking on a check box labeled "Ok!" to end the trial. This time limit was introduced to avoid very slow responses, because participants could easily realize that very slow responses did not, contrary to the instructions, always cause the "own" effect to appear later than the "other's" effect.

We included some reminder trials without TOJs in which only one effect occurred to remind participants who controlled which effect. In the single effect groups the "own" effect was accompanied by the message "Diesmal war der andereim Vergleichzulangsam. Du hast den Passagiergerettet!" (German for "This time the other participant was in comparison too slow. You saved the passenger!"). The allegedly "other's" effect was accompanied by the message "Diesmalwarst Du im Vergleichzulangsam. Der andere hat den Passagiergerettet!" (German for "This time you were slower. The other participant saved the passenger!"). In the respective trials in the compound effect group the message always read "Aufgrundeinertechnischen Fehlfunktionistnure in Rettungsring ins Wassergefallen!" (German for "Due to a technical fault only one life buoy fell into the water"). In both groups the message had to be acknowledged with a click on a check box labeled "Ok!"

The two instruction blocks comprised 20 trials each, resulting in an instruction phase comprising of 20 self-performed and 20 observed instruction trials. After the instruction phase, participants performed 26 practice trials that included all trial types that would be in the main experimental blocks to ensure that participants understood all tasks. Six experimental blocks with 48 trials each followed. In each block, each delay (−50, −40, −30, −20, −10, 0, 10, 20, 30, 40, and 50 ms) was repeated four times. In addition, there were four reminder trials per block in which only one buoy appeared (red or yellow presented at the left or right side). The temporal and spatial order of the effects' appearance was counterbalanced within participants. Each effect appeared equally as often on the left side as on the right side of the screen.

Before debriefing at the end of the experiments we asked participants separately to describe their task and why the computers' were connected. All but one (the excluded participant) described the experiment as instructed and did not suspect the instructions to be false.

DATA ANALYSIS

In each experimental trial participants indicated whether the yellow or the red effect appeared first. To analyze whether the "own" effect is perceived earlier than the "other's" effect and to compare TOJs with the compound effect group, we made the arbitrary decision to analyze how often the yellow buoy was perceived earlier than the red buoy[1], i.e., for each participant and delay we calculated the proportion of "yellow first" responses. Based on this analysis, we expected that participants who believed they caused the yellow effect to perceive the yellow effect earlier than participants who believed they caused the red effect. Furthermore, participants in the compound effect group were expected to perceive the yellow effect later than participants who believed they caused the yellow effect, but earlier than participants who believed they caused the red effect. About 3.9% of all planned TOJ trials were stopped before any effect appeared because the participants did not respond within 750 ms.

We fitted logistic functions to the "yellow first" responses using the psignifit toolbox (Wichmann and Hill, 2001) for MATLAB. From each fitted function we calculated the 50%-value of the function, the Point of Subjective Simultaneity (PSS). This value represents the temporal delay between the yellow and the red effect that results in the participant not being able to discriminate the order of the two stimuli and thus has to guess, resulting in 50% "red first" and 50% "yellow first" responses. When the yellow buoy is perceived earlier than the red buoy, the PSS is larger than zero, indicating that a yellow buoy that occurs x ms after a red buoy is perceived as occurring simultaneously with the red buoy. In contrast, when the red buoy is perceived earlier than the yellow buoy, the PSS is smaller than zero because the yellow buoy that appears × ms before the red buoy is perceived as occurring simultaneously with the red buoy.

We also calculated the difference limen (DL) as the difference between the 75% and the 25% score of the function divided by two. The DL is a measure for the steepness of the function and indicates the temporal resolution of the judgments of each participant. The higher the temporal resolution of judgments, the more consistent a participant is in his/her judgments regarding each delay, resulting in a steeper function and thus a smaller DL.

RESULTS

We conducted ANOVAs on the PSS and the DLs including the between-subjects factor type of causal belief (single yellow effect, single red effect, compound effect).

The ANOVA on the PSS revealed differences between causal belief conditions, $F(2{,}45) = 6.86$, $p = 0.003$, $\eta_p^2 = 0.234$ (see **Figure 3**). When participants believed they caused the yellow effect (single yellow effect), the yellow effect would have to appear 7.4 ms after the red effect for them to be perceived simultaneously (i.e., the PSS was 7.4 ms). When participants believed they caused the red effect (single red effect), the red effect would have to appear 5.4 ms after the yellow effect for them to be perceived simultaneously (i.e., the PSS was −5.5 ms). When participants believed they caused both effects (compound effect), the yellow effect would have to appear 1.2 ms after the red effect for them to be perceived simultaneously.

Post hoc t-tests revealed that each single comparison was significant, i.e., PSS for the single yellow effect group was larger than the PSS in the compound effect group (7.4 vs. 1.2 ms), $t(34) = -2.06$, $p = 0.047$, and it was larger than the PSS in the single red effect group (7.4 vs. −5.5 ms), $t(22) = -3.04$, $p = 0.006$. In the single red effect group the PSS was smaller than in the compound group, $t(34) = 2.64$, $p = 0.012$.

The ANOVA on the DLs revealed no difference between groups, $F(2{,}45) = 0.48$, $p = 0.622$, $\eta_p^2 = 0.021$. DLs amounted to 25.9 ms in the single yellow effect group, to 23.8 ms in single red effect group, and to 22.0 ms in the compound effect group.

DISCUSSION

In this study, we aimed at investigating whether different causal beliefs about who causes one of two effect stimuli influence the perceived temporal order of these effect stimuli. Participants who believed they caused only one of two effect stimuli perceived their "own" effect earlier than the "other's" effect. This confirms the assumption of stronger IB for allegedly "own" effects than for effects that are believed to be caused by another person. The "own" effect was also perceived earlier than the effect of the respective color in a group of participants who believed they caused both effects as compound effect. Thus, we can rule out that one of the effects was generally perceived earlier than the other due to any stimulus features.

To measure time perception for "own" effects, we applied TOJs. TOJs have recently been shown to be a useful method to measure the IB effect using a psychophysical method (Cravo et al., 2011). Choosing this psychophysical method has at least two advantages. First, in contrast to the clock paradigm, TOJs allow us to directly compare the temporal order of a putative effect and another stimulus within one trial. This let us directly assess the time perception of the "own" and the "other's" effect instead of deducing temporal order from time estimations in different trials in relation to the revolving clock hand.

Second, TOJs enable us to analyze not only the perceived time (the PSS) of the effect stimuli, but also to compare the DLs of time judgments as a measure of the temporal resolution of time judgments (Nolden et al., 2012). Importantly, DLs did not differ between the single effect group and the compound effect group. Applying this method enabled us to rule out the possibility that the manipulation of the causal belief influenced temporal resolution because, for example, of changing difficulty level of the task. Instead, the belief manipulation added a constant difference to temporal estimations, but left the overall consistency in TOJs

---

[1] As in the compound effect group, because no effect is assigned as "own" effect we cannot analyze according to the "own" and the "other's" effect. Choosing the "yellow first" responses was arbitrary, but because each response was a discrete decision between "yellow first" and "red first," the results would have been equal with the opposite sign if we had chosen "red first" responses.

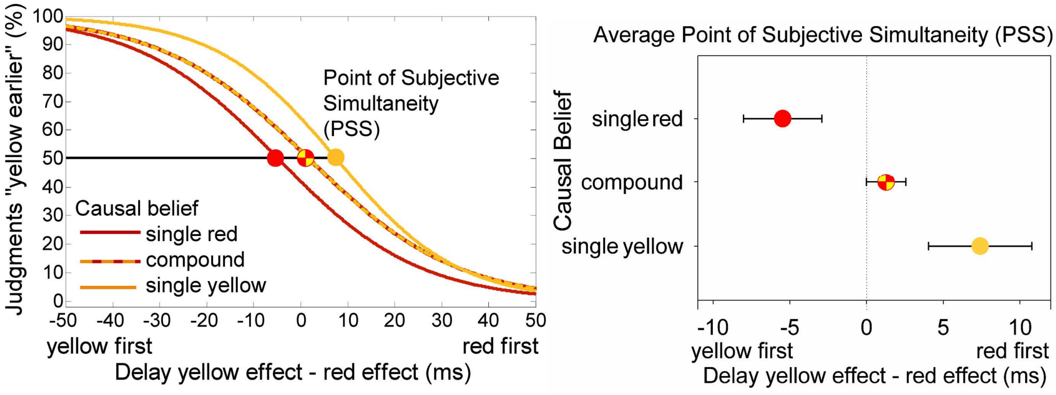

FIGURE 3 | Results of prototypical participants (left) and group means for the PSS (right). On the left fitted functions of three prototypical participants from the single effect yellow and red group (colors according to group name) and the compound effect group (dashed line). The group means of the points of subjective simultaneity are shown in the bar chart on the right. Error bars depict standard errors.

unchanged. Thus, TOJs were biased by the causal belief, but this bias did not affect the reliability of the TOJ.

Taken together, our result that "own" effects are perceived earlier strengthens and extends the recent finding of Desantis et al. (2011), who showed that an effect tone is perceived earlier in trials in which the participant believed he/she caused the tone compared to other trials when the participant believed that another person caused the tone. Here the information who would cause the effect tone in the next trial participants could have changed the level of participants' arousal or motivation in trials in which they knew they would cause a tone with their key press compared to trials in which they knew they would press a key, but hear another person's effect. In our study, participants compared the perceived time of the "own" and the "other's" effect relative to each other within each trial. So the participants' belief that they produced one specific effect remained constant throughout the experiment. This enables us to exclude any possible explanation based on trial-by-trial differences for differing time judgments between "own" and "other's" effects. Instead causal belief influences the perceived time of action effects on a stimulus-specific level. In addition, assessing the DLs of time judgments enables us to rule out that the temporal resolution differs depending on the instruction to cause one or two effects. To sum up, our study fosters the conclusion that IB is stronger for allegedly "own" action effects than for action effects that are attributed to another person's action.

The influence of causality and causal beliefs on IB has been discussed from the time the IB effect was first described (see Moore and Obhi, 2012 for a recent review). Eagleman and Holcombe (2002), for example, discussed whether the temporal attraction between action and effect was the counterpart of larger perceived causality between cause and effects the closer the effect appears after the cause (Hume, 1739; Michotte, 1963). Similarly, IB has been discussed as a process that supports the feeling of agency, i.e., the perceived causal control over one's action effects. Interestingly, agency and IB have been found to be correlated only when both measures are collected within one trial, but not when they are measured in different trials (Ebert and Wegner, 2010; see also Obhi and Hall, 2011). Nevertheless, there is evidence that IB depends on causal beliefs because IB effects occur for action effects, but not for effects caused by observed non-agentic sources (Wohlschläger et al., 2003a,b; Cravo et al., 2009). Furthermore, IB in terms of a later perception of the action is restricted to cases where the causal relation between action and effect is highly reliable in terms where the effect follows the action with high contingency (Moore and Haggard, 2008; Moore et al., 2009).

Recently, the impact of causality on IB has been demonstrated even more convincingly. Dogge et al. (2012) observed IB even in the absence of a voluntary action. In that study, the effect of an involuntary passive key press was perceived shifted toward the key press when participants believed that the passive key press caused the effect. In contrast, when no causal belief instruction was given about a causal relation between key press and effect tone, there was no shift in the perceived time of the effect. The authors assume that the predictive thought of the effect (cf. Wegner and Wheatley, 1999) before the passively induced key press leads to an increased level of perceived control and thus to a shift in the perceived time of the effect when the movement was believed to cause the effect, even in the absence of a voluntary movement. This shift in the perceived time of the effect is smaller after involuntary compared to voluntary key presses, but it shows that even in the absence of an intended movement the causal relation between the movement and the effect is sufficient to induce a certain degree of IB.

Further support of a relation between causal belief and IB is evidenced in a study of Buehner and Humphreys (2009). Their participants heard two tones and were asked synchronize two key presses to the two tones. In a non-causal condition the second tone followed the first after a fixed interval. In a "causal condition" the second tone was caused by the first key press and thus occurred after a fixed interval after the action (the same interval as in the non-causal condition). Actually, participants timed the two key presses in relation to the times of the tones differently in the two conditions, suggesting that they perceived the action and effect to be closer in time in the causal condition as suggested by IB (see Buehner and Humphreys, 2010 for similar results on causal relations between spatial stimuli). However, in this experiment causality was manipulated in that there were physical differences such as different time intervals in the causal and non-causal conditions. In our study there were no physical differences between "own" effects and "other's" effects across participants confirming that it is actually the causal belief alone that changed the perceived time of action effects.

Interestingly, the conclusion that "own" effects are perceived earlier than others' effects seems to be contradicted by recent results reported by Obhi and Hall, 2011; for similar results see also Strother et al., 2010). They investigated IB in a social situation, in which two participants performed a task together on one computer with one shared key. In each trial, one participant triggered a tone by pressing the key and the other participant was to respond by pressing the same key as quickly as possible after the key was pressed. Two-hundred milliseconds after the first key press a tone effect occurred. Each participant then judged who they believed had caused the effect (the actor) and the time of the actor's key press. In this study, the IB effect was not reduced when the participant was the responder and thus judged the time of the actor's, i.e., another person's, key press and effect than when the participant was the actor himself/herself. That is, time judgments for the action and the effect were the same, regardless of whether the participant believed that the other participant caused the effect (and thus judged the observed action of the actor) or whether the participant believed himself/herself to be the actor (and thus judged the time of his/her own action).

To resolve this contradiction we suggest that there is a critical difference between the experimental setting of Desantis et al. (2011) and our setting on the one hand, and between the experimental setting of Obhi and Hall (2011) on the other hand. In Obhi and Hall's study participants were instructed to cooperate on the experimental task. In contrast, in the study of Desantis et al. participants performed the task on their own, and in our study no cooperation was needed because one life buoy would be sufficient to save a swimmer's life. This fits well with Obhi and Hall's, 2011, p. 655) suggestion that participants might form "a 'we' identity" in the shared task. Even if not directly expressed in the instructions, our task implied a competitive rather than a cooperative situation as only one participant, probably the faster, will complete the task.

The instruction to cooperate on a single task might be the reason that participants showed a similar amount of IB for own and observed actions and effects in Obhi and Hall's (2011) study. Another line of research, the so-called "social Simon-effect," demonstrates that in cooperative settings, participants integrate the intention of another person into their own task set (Sebanz et al., 2003; Knoblich and Sebanz, 2006; Dolk et al., 2011; Liepelt et al., 2011). For example, when two participants share a Simon task, that is, one participant responds to green targets by pressing a left key and another participants responds to red targets by pressing a right key, performance is influenced by the location of the target. A participant who responds with the left key responds more slowly when the target stimuli occurs on the right (incompatible) side of the screen than when the target occurs on the left (compatible) side of the screen. In contrast, when one participant performs his/her half of the task alone (which is actually a Go-NoGo task, e.g., respond to green targets, do not respond to red targets), the compatibility effect regarding the location of the target and the response key is heavily reduced. Based on this evidence, we assume that participants adopt the intention of the other participant more strongly when participants cooperate on a task than when they infer from information given on screen that the effect they perceive will be caused by another person performing the same task at the same time (as in the study of Desantis et al., 2011) or when they compete on a task (as in our study). This assumption could explain why on the one hand Obhi and Hall (2011) found in a collaborative situation the size of IB in terms of the perceptual shift of actions and effects toward each other was independent of whether actions and effects are attributed to the own action or the action of an observed participant. On the other hand the assumption would also explain why in non-collaborative situations both Desantis et al. (2011) and we observed stronger IB in terms of an earlier perception of the "own" compared to "another person's" effect. However, this is a *post hoc* hypothesis and future studies are needed to investigate how cooperation vs. competition changes the perceived time of another person's action effects.

ACKNOWLEDGMENTS

We thank Richmond Kinney for proofreading. This research was supported by a grant from the German Science Foundation (DFG, Grant Ki 1388/3-1). This publication was funded by the German Research Foundation (DFG) and the University of Wuerzburg in the funding programme Open Access Publishing.

REFERENCES

Buehner, M. J., and Humphreys, G. R. (2009). Causal binding of actions to their effects. *Psychol. Sci.* 20, 1221–1228.

Buehner, M. J., and Humphreys, G. R. (2010). Causal contraction: spatial binding in the perception of collision events. *Psychol. Sci.* 21, 44–48.

Cravo, A. M., Claessens, P. M. E., and Baldo, M. V. C. (2009). Voluntary action and causality in temporal binding. *Exp. Brain Res.* 199, 95–99.

Cravo, A. M., Claessens, P. M. E., and Baldo, M. V. C. (2011). The relation between action, predictability and temporal contiguity in temporal binding. *Acta Psychol. (Amst)* 136, 157–166.

Desantis, A., Roussel, C., and Waszak, F. (2011). On the influence of causal beliefs on the feeling of agency. *Conscious. Cogn.* 20, 1211–1220.

Dogge, M., Schaap, M., Custers, R., Wegner, D. M., and Aarts, H. (2012). When moving without volition: implied self-causation enhances binding strength between involuntary actions and effects. *Conscious. Cogn.* 21, 501–506.

Dolk, T., Hommel, B., Colzato, L. S., Schütz-Bosbach, S., Prinz, W., and Liepelt, R. (2011). How "social" is the social Simon effect? *Front. Psychol.* 2:84. doi:10.3389/fpsyg.2011.00084

Eagleman, D. M., and Holcombe, A. O. (2002). Causality and the perception of time. *Trends Cogn. Sci. (Regul. Ed.)* 6, 323–325.

Ebert, J. P., and Wegner, D. M. (2010). Time warp: authorship shapes the perceived timing of actions and events. *Conscious. Cogn.* 19, 481–489.

Engbert, K., Wohlschläger, A., Thomas, R., and Haggard, P. (2007). Agency, subjective time, and other minds. *J. Exp. Psychol. Hum. Percept. Perform.* 33, 1261–1268.

Haggard, P. (2005). Conscious intention and motor cognition. *Trends Cogn. Sci. (Regul. Ed.)* 9, 290–295.

Haggard, P., Aschersleben, G., Gehrke, J., and Prinz, W. (2002a). "Action, binding, and awareness," in *Common Mechanisms in Perception and Action: Attention and Performance*, XIX, eds W. Prinz and B. Hommel (Oxford: Oxford University Press), 266–285.

Haggard, P., Clark, S., and Kalogeras, J. (2002b). Voluntary action and conscious awareness. *Nat. Neurosci.* 5, 382–385.

Hume, D. (1739). *A Treatise of Human Nature: Being an Attempt to*

Introduce the Experimental Method of Reasoning into Moral Subjects. London: John Noon.

Humphreys, G. R., and Buehner, M. J. (2009). Magnitude estimation reveals temporal binding at super-second intervals. *J. Exp. Psychol. Hum. Percept. Perform.* 35, 1542–1549.

Knoblich, G., and Sebanz, N. (2006). The social nature of perception and action. *Curr. Dir. Psychol. Sci.* 15, 99–104.

Liepelt, R., Wenke, D., Fischer, R., and Prinz, W. (2011). Trial-to-trial sequential dependencies in a social and non-social Simon task. *Psychol. Res.* 75, 366–375.

Michotte, A. E. (1963). *The Perception of Causality*. London: Methuen.

Moore, J. W., and Haggard, P. (2008). Awareness of action: inference and prediction. *Conscious. Cogn.* 17, 136–144.

Moore, J. W., Lagnado, D. A., Deal, D. C., and Haggard, P. (2009). Feelings of control: contingency determines experience of action. *Cognition* 110, 279–283.

Moore, J. W., and Obhi, S. S. (2012). Intentional binding and the sense of agency: a review. *Conscious. Cogn.* 21, 546–561.

Nolden, S., Haering, C., and Kiesel, A. (2012). Assessing intentional binding with the method of constant stimuli. *Conscious. Cogn.* 21, 1176–1185.

Obhi, S., and Hall, P. (2011). Sense of agency and intentional binding in joint action. *Exp. Brain Res.* 211, 655–662.

Schneider, W., Eschman, A., and Zuccolotto, A. (2002). *E-Prime User's Guide*. Pittsburgh: Psychology Software Tools.

Sebanz, N., Knoblich, G., and Prinz, W. (2003). Representing others' actions: just like one's own? *Cognition* 88, B11–B21.

Strother, L., House, K. A., and Obhi, S. S. (2010). Subjective agency and awareness of shared actions. *Conscious. Cogn.* 19, 12–20.

Wegner, D. M., and Wheatley, T. (1999). Apparent mental causation – sources of the experience of will. *Am. Psychol.* 54, 480–492.

Wichmann, F. A., and Hill, N. J. (2001). The psychometric function: I. Fitting, sampling, and goodness of fit. *Percept. Psychophys.* 63, 1293–1313.

Wohlschläger, A., Engbert, K., and Haggard, P. (2003a). Intentionality as a constituting condition for the own self-and other selves. *Conscious. Cogn.* 12, 708–716.

Wohlschläger, A., Haggard, P., Gesierich, B., and Prinz, W. (2003b). The perceived onset time of self-and other-generated actions. *Psychol. Sci.* 14, 586–591.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 13 June 2012; accepted: 20 September 2012; published online: 08 October 2012.

Citation: Haering C and Kiesel A (2012) Mine is earlier than yours: causal beliefs influence the perceived time of action effects. Front. Psychology ***3****:393. doi: 10.3389/fpsyg.2012.00393*

This article was submitted to Frontiers in Cognition, a specialty of Frontiers in Psychology.

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 27 December 2012
doi: 10.3389/fpsyg.2012.00581

Binding success and failure: evidence for the spontaneous integration of perceptual features and object evaluations

Bernhard Hommel[1]* **and André W. Keizer**[1,2]

[1] Cognitive Psychology Unit, Leiden Institute for Brain and Cognition, Leiden University, Leiden, Netherlands
[2] Department of Psychology, Cognitive Neuroscience Group, University of Amsterdam, Amsterdam, Netherlands

Edited by:
Dorit Wenke, Humboldt University at Berlin, Germany

Reviewed by:
Mike Wendt, University of Armed Forces, Germany
Christian Frings, University of Trier, Germany

***Correspondence:**
Bernhard Hommel, Department of Psychology, Cognitive Psychology Unit, Leiden University, Wassenaarseweg 52, 2333 AK Leiden, Netherlands.
e-mail: hommel@fsw.leidenuniv.nl

Humans represent perceptual events in a distributed, feature-specific fashion, which calls for some sort of feature integration. It has been suggested that processing an event leads to the creation of a temporary binding of the corresponding feature codes – an object file. Here we show that object files do not only comprise of perceptual feature codes but also include codes that reflect evaluations of the perceptual event.

Keywords: feature binding, affective binding, object file, event integration, emotion

INTRODUCTION

Humans represent the events they perceive in a distributed fashion, which calls for some sort of feature integration. Kahneman et al. (1992) have argued that people bind the cognitive codes of event features into temporary object files. They demonstrated that participants respond particularly fast and accurately to repeated stimuli if these also appear in the same location, suggesting that the first encounter led to the binding of shape and location codes. Moreover, repeating one or more features of a stimulus but alternating others impairs performance (Hommel, 1998), suggesting that feature-repetition leads to the automatic retrieval of the just-created binding, which interferes with processing the present feature combination if it differs from the previous one (Hommel, 2004). Indeed, repeating one of two features of a visual stimulus reactivates the cortical area coding for the non-repeated feature (Keizer et al., 2008). Research on object files has mainly focused on the binding of perceptual feature codes. However, feature codes are no copies of external events but brain responses to those events, which raises the question whether other, non-perceptual responses become part of an object file as well. Here we investigated whether object files also contain information about people's evaluative responses to a stimulus.

As in Keizer et al. (2008), we presented participants with pairs of stimuli in each trial, a prime (S1) followed by a probe (S2; see **Figure 1A**). Both stimuli consisted of blends of a face and a house, and either the face or the house moved diagonally up and down. Participants did not respond to S1 but categorized the moving object's motion direction (top-left/bottom-right or top-right/bottom-left). This task allowed for the orthogonal repetition and alternation of the moving object and the direction in which it moved. The integration of moving object and motion upon processing of S1 was expected to yield an interaction of the two repetition effects, with a pattern that indicates worse performance if one of the two features repeats while the other alternates (Keizer et al., 2008).

Evaluative responses to the two stimuli were induced by making participants believe that the sequence of the objects that moved followed a consistent pattern across trials and having them predict whether a face or house would move on S1 and on S2. Encountering a stimulus that meets the prediction was considered to evoke a positive evaluation (success), while stimuli not meeting the prediction were thought to evoke a negative evaluation (failure). The question was whether the effect of repeating the type of evaluative response (in trials where the two predictions happened to be both correct or both incorrect) vs. alternating the evaluative response (in trials where just one prediction happened to be correct) would interact with the effect of repeating vs. alternating the type of object being moved and/or the type of motion. If so, this would suggest that codes related to visual event features are integrated with codes representing event-related success and failure.

MATERIALS AND METHODS

Twenty-four students participated for course credit or pay. They judged the motion direction shown in S2 by pressing a left vs. right key of a computer keyboard. After predictions were made, a trial would start with a 1000-ms blank interval. Then S1 appeared for 675 ms, followed by a 1000-ms blank interval. Thereafter, S2 appeared for 675 ms, followed by another 1000-ms blank interval.

As in Keizer et al. (2008), visual stimuli were composed by superimposing luminance-matched grayscale front-view photographs of male and female faces and of houses. The house-face combinations for the 240 trials were constructed by randomly drawing from eight possible houses and faces, except that all four combinations of repeating vs. alternating the moving object (face or house) were equally likely. The face image and the house images for a given trial were randomly selected from the set of eight face

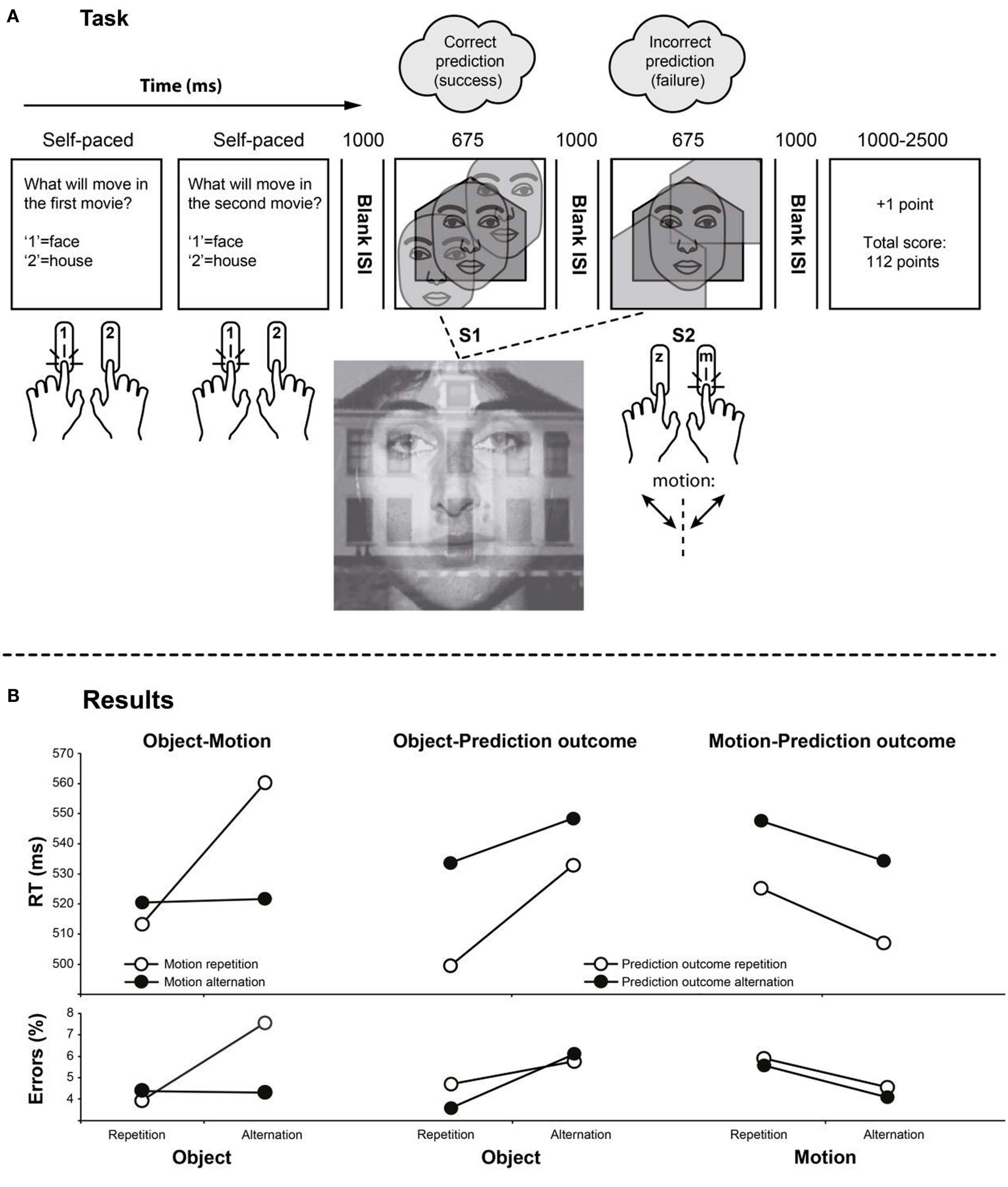

FIGURE 1 | (A) Sequence of events, from left to right. **(B)** Reaction times (RTs) and error percentages as a function of object repetition/alternation and motion repetition/alternation (leftmost panels), object repetition/alternation and prediction-outcome repetition/alternation (middle panels), and motion repetition/alternation and prediction-outcome repetition/alternation (rightmost panels).

images and eight house images, and they were always the same for S1 and S2.

Before the stimuli were presented, participants predicted whether a face or a house would move on S1 and S2 by pressing the "1" or "2" key on a computer keyboard (counterbalanced), respectively. They received points for each correct prediction (i.e., 0–2 per trial) and the number of earned points, together with a running total, was presented at the end of each trial. The three highest-scoring participants received an extra of 5€ after the experiment was completed.

RESULTS

The RTs and error rates for the response to S2 were analyzed. Trials in which RTs deviated more than 2 SDs from the mean were excluded. The remaining data were aggregated according to whether the moving object was the same in S1 and S2 (object repetition) or different (object alternation), the direction of the movement was the same for S1 and S2 (movement repetition) or different (movement alternation), and whether the outcomes of the two predictions in each trial were the same (two times success or two times failure; a prediction outcome repetition) or different (one failure and one success; a prediction outcome alternation). RTs and error rates from the resulting eight design cells were entered into ANOVAs with three corresponding factors:repetition (vs. alternation) of moving object, motion, and prediction outcome (see **Figure 1B**; **Table 1**).

In RTs, responses were faster upon the repetition of moving object, $F(23,1) = 53.2$, $\text{MSE} = 523.9$, $p < 0.001$, and prediction outcome, $F(23,1) = 66.2$, $\text{MSE} = 448.0$, $p < 0.001$, and slower if the type of motion was repeated, $F(23,1) = 12.8$, $\text{MSE} = 928.5$, $p < 0.005$. More importantly for our purposes, interactions were obtained for moving object and motion, $F(1,23) = 106.4$, $\text{MSE} = 236.2$, $p < 0.001$, and moving object and prediction outcome, $F(1,23) = 7.0$, $\text{MSE} = 599.1$, $p < 0.05$; separate analyses confirmed that the latter did not depend on whether the outcomes were positive or negative.

Error rates followed the same pattern: performance was more accurate if the moving object repeated, $F(1,23) = 10.4$, $\text{MSE} = 14.6$, $p < 0.005$, an effect that interacted with the repetition of motion, $F(1,23) = 7.5$, $\text{MSE} = 22.1$, $p < 0.05$. The only exception was the interaction of moving object and prediction outcome, which showed the opposite pattern of the RTs – a more pronounced object repetition-alternation effect with prediction alternation. However, the corresponding interaction was far from significance, $F(1,23) = 1.35$, $\text{MSE} = 24.15.6$, $p > 0.25$.

DISCUSSION

We were able to replicate the well-known observation of worse performance if a visual feature is repeated while another alternates (Hommel, 1998), suggesting that participants spontaneously integrated the codes of these features – the type of moving object and motion direction in our case. In addition to the more interesting interactions, we also obtained main effects of all three experimental factors. In the cases of object repetition and prediction outcome repetition, the underlying pattern is rather straightforward: alternations of features can be suspected to create neural conflict between the present and the previous feature values (Kühn et al., 2011), which slows down object identification. In the case of motion, however, alternations produced faster, rather than slower responses. Even though the interaction with object repetition (see **Figure 1B**) makes the interpretation difficult, we speculate that the exposure to a repeated motion pattern over 675 ms might have resulted in motion adaptation (Ölveczky et al., 2007), which impaired the processing of objects moving into the same direction. This need not have prevented the standard feature-repetition benefit but it might have overshadowed this effect in the data.

More importantly, however, our findings demonstrate that participants coded their successes and failures in predicting the motion direction of the two visual stimuli, and integrated these codes with codes representing the moving object. Interestingly, this integration seemed to be selective for the object feature that the prediction was referring to, while there was no evidence that motion direction interacted with prediction outcomes. This might suggest that object files do not simply lump together all information that relates to a given object but, rather, consist of a complex, multi-level representational structure (Hommel, 2004). However, it is also possible that all available information was actually integrated but only partially retrieved while processing S2. Hence, it might be that control processes modulate stimulus-driven retrieval of information in such a way that only relevant feature codes are retrieved to a degree that allows affecting behavior (e.g., Keizer et al., 2010).

Also of importance for the purpose of the present study, the interaction between the repetition of object features and the repetition of the outcomes of object-related judgments suggests that codes referring to the physical features of object are integrated with codes referring to the evaluation of objects or object-related aspects. In other words, object files seem to allow for the evaluative "tagging" of visual feature codes. It is interesting to consider how "affective" or "emotional" these evaluative codes actually are. On the one hand, one might consider them some kind of "somatic markers" (Damasio, 1994) that relate to and represent the emotional experience one had when creating them. If so, reactivating an evaluative code in the process of a stimulus-induced retrieval of object information might lead to the recall or simulation of the emotional state one was in when having experienced success or failure with regard to this particular object. In this case, evaluative codes may actually be considered "affective markers." On the other hand, however, it is also possible that evaluative codes only indicate successes and failures without necessarily revoking any related emotional state. In a recent study, Eder et al. (submitted) provided evidence that actions are not only integrated with representations

Table 1 | Means of mean reaction times and SD for responses (RT; in ms) and percentages of errors (PE) for responses to stimulus 2, as a function of the repetition vs. alternation of motion direction, moving object, and reward.

Motion	**Repeated**				**Alternated**			
Moving object	**Repeated**		**Alternated**		**Repeated**		**Alternated**	
Prediction outcome	**RT (SD)**	**PE (SD)**	**RT (SD)**	**PE (SD)**	**RT (SD)**	**PE (SD)**	**RT (SD)**	**PE (SD)**
Repeated	497 (53)	5 (6)	553 (70)	7 (6)	502 (56)	5 (7)	513 (60)	5 (6)
Alternated	529 (65)	3 (4)	567 (56)	8 (6)	539 (62)	4 (6)	530 (61)	4 (5)

of their affective consequences (Eder and Hommel, in press) but that these representations have two different kinds of effect on action control: a directive function when selecting responses in a stimulus-driven forced-choice task and (in addition) an incentive function when selecting freely chosen actions. It is possible that the former is based on the mere information whether a particular response will or will not produce positive outcomes while the latter relies on a simulation of the expected affective state in the sense of Damasio (1994). If so, the present study might be taken to speak more to the representations underlying the directive function of evaluation-related outcome representations.

ACKNOWLEDGMENTS

This study and its publication were made possible through funding received from the Netherlands Organization for Scientific Research (NWO; 433-09-243).

REFERENCES

Damasio, A. R. (1994). *Descartes' Error: Emotion, Reason, and the Human Brain.* New York: Grosset/Putnam.

Eder, A. B., and Hommel, B. (in press). Anticipatory control of approach and avoidance: an ideomotor approach. *Emotion Review.*

Hommel, B. (1998). Event files: evidence for automatic integration of stimulus-response episodes. *Vis. cogn.* 5, 183–216.

Hommel, B. (2004). Event files: feature binding in and across perception and action. *Trends Cogn. Sci. (Regul. Ed.)* 8, 494–500.

Kahneman, D., Treisman, A., and Gibbs, B. J. (1992). The reviewing of object files: object-specific integration of information. *Cogn. Psychol.* 24, 175–219.

Keizer, A. W., Nieuwenhuis, S., Colzato, L. S., Theeuwisse, W., Rombouts, S. A. R. B., and Hommel, B. (2008). When moving faces activate the house area: an fMRI study of object file retrieval. *Behav. Brain Funct.* 4, 50.

Keizer, A. W., Verment, R., and Hommel, B. (2010). Enhancing cognitive control through neurofeedback: a role of gamma-band activity in managing episodic retrieval. *Neuroimage* 49, 3404–3413.

Kühn, S., Keizer, A., Colzato, L. S., Rombouts, S. A. R. B., and Hommel, B. (2011). The neural underpinnings of event-file management: evidence for stimulus-induced activation of, and competition among stimulus-response bindings. *J. Cogn. Neurosci.* 23, 896–904.

Ölveczky, B. P., Baccus, S. A., and Meister, M. (2007). Retinal adaptation to object motion. *Neuron* 56, 689–700.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 18 June 2012; accepted: 10 December 2012; published online: 27 December 2012.

Citation: Hommel B and Keizer AW (2012) Binding success and failure: evidence for the spontaneous integration of perceptual features and object evaluations. Front. Psychology ***3****:581. doi: 10.3389/fpsyg.2012.00581*

This article was submitted to Frontiers in Cognition, a specialty of Frontiers in Psychology.

frontiers in
PSYCHOLOGY

REVIEW ARTICLE
published: 28 September 2012
doi: 10.3389/fpsyg.2012.00369

Adaptive skeletal muscle action requires anticipation and "conscious broadcasting"

T. Andrew Poehlman[1]*, Tiffany K. Jantz[2] and Ezequiel Morsella[2,3]

[1] Marketing Department, Cox School of Business, Southern Methodist University, Dallas, TX, USA
[2] Department of Psychology, San Francisco State University, San Francisco, CA, USA
[3] Department of Neurology, University of California, San Francisco, CA, USA

Edited by:
Dorit Wenke, Humboldt University at Berlin, Germany

Reviewed by:
Bernhard Hommel, Leiden University, Netherlands
Robert Gaschler, Humboldt-Universität Berlin, Germany

***Correspondence:**
T. Andrew Poehlman, Marketing Department, Cox School of Business, Southern Methodist University, Fincher Building, PO Box 750333 Dallas, TX 75275, USA.
e-mail: poehlman@cox.smu.edu

Historically, the conscious and anticipatory processes involved in voluntary action have been associated with the loftiest heights of nervous function. Concepts like mental time travel, "theory of mind," and the formation of "the self" have been at the center of many attempts to determine the purpose of consciousness. Eventually, more reductionistic accounts of consciousness emerged, proposing rather that conscious states play a much more basic role in nervous function. Though the widely held *integration consensus* proposes that conscious states integrate information-processing structures and events that would otherwise be independent, *Supramodular Interaction Theory* (SIT) argues that conscious states are necessary for the integration of only certain kinds of information. As revealed in this selective review, this integration is related to what is casually referred to as "voluntary" action, which is intimately related to the skeletal muscle output system. Through a peculiar form of broadcasting, conscious integration often controls and guides action via "ideomotor" mechanisms, where anticipatory processes play a central role. Our selective review covers evidence (including findings from anesthesia research) for the integration consensus, SIT, and ideomotor theory.

Keywords: consciousness, skeletal muscle, anticipation, ideomotor action, voluntary action

Understanding how consciousness arises from the brain is a far greater task than what the average person might surmise. The unfortunate truth is that, at the present stage of understanding, not only do scientists not have a clue regarding how conscious states emerge from the human nervous system, but they do not even possess the smallest inkling regarding how something like consciousness could emerge from any set of real or hypothetical circumstances (Levine, 1983; Banks, 1995; Godwin et al., in press). As Shallice (1972, p. 383) concludes, "The problem of consciousness occupies an analogous position for cognitive psychology as the problem of language behavior does for behaviorism, namely, an unsolved anomaly within the domain of the approach."

In this selective review, we discuss a subset of findings revealing some humble progress regarding this puzzle. This progress stems primarily from observations of everyday action planning, anticipatory processing, and the voluntary control of overt action through the skeletal muscle system. Examination of these interconnections reveals why, for every voluntary action, the actor can self-report conscious content responsible for that action. As explained below, the implications of this often overlooked but reliable observation (that voluntary actions are connected to conscious content) are important, even if self-reports on the causes of these actions by actors are often inaccurate (Nisbett and Wilson, 1977; Wegner, 2002, 2003). By integrating various disparate literatures, we put a non-traditional frame on the connections among anticipation, conscious states, and skeletal muscle action ("skeletomotor action," for short). For instance, instead of studying consciousness by focusing on perception (the dominant approach; Crick and Koch, 2003), we examine consciousness by working backward from overt action to trace the central processes responsible for action (Morsella and Bargh, 2010). We also find relevant clues about the nature of consciousness from research on anesthesia.

Prior to discussing the interconnections among conscious states, anticipation, and skeletomotor action, it is important to explain what we mean by the generally ethereal concept of consciousness. "Consciousness," which is also sometimes referred to as "sentience" (Pinker, 1997), a "phenomenal state" (Jackson, 1982; Tye, 1999), "qualia" (Gray, 2004), or "subjective experience," has been perhaps best defined by the philosopher Nagel (1974), who proposed that an organism possesses subjective experiences if there is *something it is like* to be that organism – something it is like, for example, to be human and experience warmth, love, yellowness, or breathlessness. Similarly, Block (1995, p. 227) says, "the phenomenally conscious aspect of a state is what it is like to be in that state." In this article, we are interested in this most basic form of consciousness, a form of consciousness that should be distinguished from higher forms of consciousness (e.g., self-consciousness, consciousness of one's culture, etc.). From our perspective, if any thing has an experience of any kind, then it possesses the kind of consciousness in which we are interested.

While it is true that throughout the history of psychology, "consciousness" has been coupled with outstandingly complex phenomena like "the self" and mental time travel (see review of high-level theories in Morsella, 2005), recently, less lofty accounts of consciousness have emerged, proposing rather that conscious states play a much more basic role in nervous function. One promising direction of this research has been achieved by juxtaposing conscious and unconscious processes in terms of their cognitive and neural correlates (e.g., Shallice, 1972; Baars, 1988, 2002; Logothetis and Schall, 1989; Crick and Koch, 1995; Kinsbourne, 1996; Wegner and Bargh, 1998; Grossberg, 1999; Di Lollo et al., 2000; Dehaene and Naccache, 2001; Gray, 2004; Libet, 2004; Laureys, 2005; Morsella, 2005; Merker, 2007; Doesburg et al., 2009; Damasio, 2010; Boly et al., 2011). This contrastive approach has revealed that many sophisticated processes can, and do, occur unconsciously (cf., Godwin et al., in press). For example, motor programming – which calculates the muscle fibers that should be activated at a given time in order to enact action – falls into the category of processes that can occur unconsciously (James, 1890; Grossberg, 1999; Fecteau et al., 2001; Rossetti, 2001; Rosenbaum, 2002; Goodale and Milner, 2004; Johnson and Haggard, 2005; Heath et al., 2008). Additionally, low-level (or "pre-conscious") perceptual processing also occurs unconsciously (Crick and Koch, 1995; Gray, 2004; Koch, 2004). Other mechanisms linking perception to action can also transpire unconsciously, as with the relatively obvious case of reflexes or in the less common case of automatisms (see review in Morsella and Bargh, 2011). It is important to note, that subliminal stimuli have been shown to reliably elicit motor acts as well (Fehrer and Biederman, 1962; Fehrer and Raab, 1962; Taylor and McCloskey, 1990, 1996; Hallett, 2007).

This contrastive approach has revealed that so much of nervous function is both unconscious and sophisticated. It has lead many researchers to what would have once been an unanticipated question: What do conscious states, in fact, add to brain function?

At present, it seems the answer lies in what has come to be called the *integration consensus* (Tononi and Edelman, 1988; Damasio, 1989; Freeman, 1991; Baars, 1998; Zeki and Bartels, 1999; Dehaene and Naccache, 2001; Llinás and Ribary, 2001; Varela et al., 2001; Clark, 2002; Ortinski and Meador, 2004; Sergent and Dehaene, 2004; Del Cul et al., 2007; Doesburg et al., 2009; Ulhaas et al., 2009; Boly et al., 2011). The integration consensus proposes that conscious states integrate neural activities and information-processing structures that would otherwise be independent (see a review in Baars, 2005). For example, when actions are decoupled from consciousness (e.g., in neurological disorders such as anarchic hand syndrome and utilization behavior syndrome; Lhermitte, 1983; Marchetti and Della Sala, 1998), the actions (e.g., a hand meandering through irrelevant actions like tugging at its owner's shirt) often appear impulsive or inappropriate, as if they are not influenced by the kinds of information by which they should be influenced (Morsella and Bargh, 2011). Most theoretical frameworks in the integration consensus speak of conscious information as being available "globally" in some kind of mental workspace (Baars, 2002; Sergent and Dehaene, 2004).

Separate from this global-reach system of conscious integration, unconscious processes involve smaller networks of brain areas and require less widespread activation than their conscious counterparts (Sergent and Dehaene, 2004; Baars, 2005; Gaillard et al., 2009). (See review in Morsella et al., 2010.) For example, the unconsciously mediated action of reflexive swallowing involves substantially fewer brain regions than volitional swallowing (Kern et al., 2001; Ortinski and Meador, 2004). Additionally, in the unconscious phases of deep sleep, auditory input yields activity that is limited to only the primary auditory cortex (Portas et al., 2000).

It seems that, for consciousness, the mode of interaction among regions is as important as the nature and loci of the regions (Buzsáki, 2006). For instance, the presence or lack of what has been called "interregional synchrony" leads to different cognitive and behavioral outcomes (Hummel and Gerloff, 2005; see review of neuronal communication through "coherence" in Fries, 2005). In binocular rivalry, for example, it is evident that the mode of interaction between areas is important for conscious states. During this phenomenon (Logothetis and Schall, 1989), an observer is presented with different visual stimuli to each eye simultaneously (e.g., an image of a house in one eye and of a face in the other). It might seem reasonable that, faced with such stimuli, an observer would perceive an image combining both objects – a house overlapping a face. Surprisingly, even though both images are always present, an observer experiences seeing only one object at time (i.e., a house and then a face). At any moment, the observer is unaware of the computational processes leading to this outcome; the conflict and mechanism of resolution are unconscious. Neurally, while experiencing binocular rivalry, it is only the conscious percept that is coupled, in terms of interregional synchrony, to both perceptual brain activity and motor-related processes in frontal cortex, thus supporting the view that the mode of interaction between areas, and not just activation of the areas, is important for consciousness (Doesburg et al., 2009).

EVIDENCE FROM RESEARCH ON ANESTHESIA

Supporting the integration consensus, findings in the field of anesthesiology suggest that anesthetic agents work on consciousness in part by halting the integration of information across widespread brain networks (Mashour, 2004; Hudetz, 2006; Alkire et al., 2008; Lee et al., 2009). Anesthetics may inhibit integration by acting on structures that are necessary for widespread cortical broadcasting and by slowing neural responses, thereby affecting synchronization (Munglani et al., 1993; Alkire et al., 2008). Indeed, Flohr's (1995) *information-processing theory*, John and Prichep's (2005) *anesthetic cascade*, Mashour's (2004) *cognitive unbinding paradigm*, and Alkire et al.'s (2000) *unified theory of narcosis* all directly or indirectly support the idea that anesthetics are acting by disrupting integration in the brain (Mashour, 2006).

Regarding thalamic accounts of consciousness (e.g., Penfield and Jasper, 1954; Merker, 2007), the most consistently reported effect of anesthetic agents is the reduction in thalamic blood flow and metabolism during the loss of consciousness (Hudetz, 2006; Alkire et al., 2008; Långsjö et al., 2012). It has also been suggested that thalamic blocking of somatosensory information may be the cause of the anesthetic state (Angel, 1991; Hudetz, 2006). Some anesthetics may work by affecting the posterior lateral corticothalamic complex and perhaps a medial cortical core, either directly or indirectly, thus resulting in unconsciousness (Alkire et al., 2008). Additionally, thalamocortical connectivity is associated with recovery from vegetative states (Laureys et al., 2000a,b; Mashour, 2006). However, not all anesthetics act on the thalamus in the same manner. The anesthetic ketamine, for example, results in increases in thalamic metabolism while sevoflurane sedation decreases such metabolism while the subject remains conscious (cf., Alkire et al., 2008). Additionally, studies using electroencephalography (EEG) have shown that, as soon as a subject loses consciousness, there is a marked change in cortical EEG, while thalamic EEG remains relatively the same for some minutes afterward. This begs the question as to whether the thalamus is inactivated directly or perhaps indirectly following cortical suppression (Alkire et al., 2008).

Some research on anesthesia suggests that frontal cortex alone may not constitute consciousness (Penfield and Jasper, 1954; Merker, 2007; Alkire et al., 2008). For instance, recent investigations into feedforward and feedbackward connectivity while under anesthesia suggest that conscious states are associated with, not only frontal activations, but specific frontoparietal networks (Ku et al., 2011). Additionally, low doses of anesthetics have been shown to slow the feedback stream of cortical processing, while increasing doses slow both the feedforward and feedback streams of cortical processing. These findings suggest that some form of widespread feedback dynamics, or "reentrant" processing (Di Lollo et al., 2000; Fahrenfort et al., 2007), may play an integral part in conscious awareness (see below; Hudetz, 2006; Långsjö et al., 2012). In addition, the notion that frontal cortex is unnecessary for consciousness is consistent with investigations on prefrontal lobe syndromes (Gray, 2004), the phenomenology of action and behavior (Desmurget et al., 2009; Desmurget and Sirigu, 2010), and the psychophysiology of consciousness in dreams, which involves prefrontal deactivations (Muzur et al., 2002). (See evidence for a necessary role of frontal cortex in consciousness in Boly et al., 2011). There are other regions that may be unnecessary for the brain to constitute a basic form of consciousness. For example, although the absence of the spinal cord or cerebellum leads to sensory, motor, cognitive, and affective deficits, the non-participation of these regions does not seem to eliminate basic consciousness (Schmahmann, 1998; Morsella et al., 2010). Similarly, non-participation of the basal ganglia, hippocampus, mammillary bodies, right cerebral cortex, or mediodorsal nucleus of the thalamus does not seem to hinder the ability of the nervous system to generate a basic form of consciousness (see evidence in Morsella et al., 2010; Godwin et al., in press).

SUPRAMODULAR INTERACTION THEORY AND LIMITATIONS OF THE INTEGRATION CONSENSUS

One limitation of the integration consensus is that it fails to specify exactly which kinds of integration require conscious states and which kinds can occur unconsciously. For example, conscious processing is unnecessary for integrations across different sensory modalities (e.g., the binding of features in perceptual objects) or integrations involving smooth muscle effectors (e.g., integrations in the pupillary reflex; Morsella et al., 2009a). In both cases, these integrations/conflicts can transpire unconsciously. In contrast, people tend to be aware of some of the conflicts in their nervous system. When a swimmer holds her breath underwater, for example, she cannot help but be aware of the conflict of restraining an automatic process like breathing. Further, *approach–approach* conflicts also beg for awareness (Lewin, 1935; Miller, 1959). These types of conflicts, *conscious conflicts* (Morsella, 2005), involve competition for control of the skeletal muscle output system and are triggered by incompatible skeletomotor plans, as when one holds one's breath while underwater, suppresses uttering something, or inhibits a prepotent response in a laboratory *response interference paradigm* (e.g., the Stroop and Flanker tasks; Stroop, 1935; Eriksen and Eriksen, 1974). *Supramodular Interaction Theory* (SIT; Morsella, 2005) proposes that, while the primary function of conscious states is to integrate information, only certain kinds of information require conscious integration. Specifically, it is high-level information in the service of curbing skeletomotor action so that such action is adaptive, as in the case of holding one's breath or breathing at a faster rate for some reward. Conscious conflicts are a dramatic case of such interactions. (The theory is called "supramodular," because the integrations occur at a high-level, beyond that of the Fodorian module, which is used for, say, color, and motion detection; the term "interaction" is used in the theory because conscious states permit interactions between high-level systems vying for skeletomotor control; see treatments of modularity in Fodor, 1983; Callebaut and Rasskin-Gutman, 2009.) The actual integration amongst such *response systems* may actually be "post-conscious" (Morsella, 2005). (For a thorough review of the nature of the difference between the *access* of information during conscious states and the *subjectivity* associated with that information, see Atkinson et al., 2000.) From our standpoint, conscious states are necessary, not to integrate perceptual-level processes (like feature binding), but to permit interactions among action goal inclinations that, eventually, influence the skeletal muscle system; this idea is captured in the principle of *Parallel Responses into Skeletal Muscle* (PRISM; Morsella, 2005).

To summarize in different and more concrete terms, SIT proposes that, in the nervous system, there are three distinct kinds of integration or "binding" (Morsella and Bargh, 2011). Perceptual binding (or *afference binding*) is the binding of perceptual processes and representations. This occurs in feature binding (e.g., the binding of shape to color; Zeki and Bartels, 1999) and intersensory binding (McGurk and MacDonald, 1976; Vroomen and de Gelder, 2003), in which disparate senses integrate information across the perceptual field (e.g., visual and auditory inputs regarding the source of a sound interact unconsciously). (See additional evidence for unconscious afference binding in Zmigrod and Hommel, 2011).

The second form of binding (*efference binding*) links perceptual processing to action/motor production (Haggard et al., 2002). (For advanced treatments of the topic of integration across perception and action, see Hommel et al., 2001; Astor-Jack and Haggard, 2005; Magen and Cohen, 2010.) This kind of stimulus-response ($S \rightarrow R$) binding allows for automatic button presses in response to a cue. Research has shown that efference binding can happen unconsciously, as when subjects are able to select the correct motor response (one of two button presses) when confronted with a subliminal cue (Fehrer and Biederman, 1962; Fehrer and Raab, 1962; Taylor and McCloskey, 1990, 1996; Hallett, 2007). (For studies revealing how instructions held in mind can lead to $S \rightarrow R$ mappings that resemble that of reflexes, see Cohen-Kdoshay and Meiran, 2009; Hommel, 2000; Wenke et al., 2007). More commonly, this kind of binding can also be mediated unconsciously in actions such as the pain withdrawal reflex and reflexive swallowing and inhalation. The third form of binding, *efference–efference binding*, occurs when two streams of efference binding are trying to influence skeletomotor action simultaneously (Morsella and Bargh, 2011). Importantly, these streams of efference are "bound" at, at least, the level of overt action. For instance, when a swimmer holds her breath, she experiences the conflict between the two efferent streams (wanting to inhale/wanting to suppress inhalation) and produces an action that is a "binding" of the two inclinations. In this case, the integration at the level of overt behavior is her holding her breath but behaving less comfortably than if provided with oxygen. In such a way, conflicted behavior is overtly different from non-conflicted behavior, as Skinner notes (Skinner, 1953). To him, such behaviors are more perturbable and slower in their execution. In the context of laboratory research, conflicted skeletomotor action is also apparent when a research participant suppresses a prepotent response like word reading in a response interference paradigm such as the classic Stroop task (where participants are asked only to name the color in which a word is presented). Importantly, conflicts involving perceptual processing or smooth muscle do not yield such changes in consciousness (Morsella et al., 2009a).

The tenets of SIT principally concern, not which kinds of interactions do and do not occur with phenomenal mediation, but, in identifying the function of consciousness, which kinds of basic processes cannot occur without phenomenal mediation. Thus, it is not within the scope of SIT to identify all the modular or supramodular outputs that one can be conscious of. Rather, SIT is about which integrative processes require conscious mediation. During conflicts that require conscious mediation, one is aware of the conflicting components (e.g., pain and hunger) that are brought together to influence action. Interestingly, however, one is unaware of the computational products of conscious interaction, which, should they exist, are observable only in the form of expressed behavior (e.g., breathing faster for some reward; Morsella, 2005). In other words, one is unconscious of the representations reflecting the resolution of the conflict (if such representations exist). Consciousness is necessary for the integration, but the integration is best represented, not in consciousness, but in overt behavior. Hence, our theoretical approach is named "SIT" and not "supramodular *integration* theory," because, for the reasons just outlined, the term *integration* is a loaded term. One must consider that "to combine" does not necessarily imply "to resolve."

It should be reiterated that this survey comprises a selective review of research findings, a review based on one specific vantage point (for other accounts of information integration, see Tononi and Edelman, 1988; Logan et al., 1999; Baars, 2002; Miller and Ulrich, 2003; Goodale and Milner, 2004; Dijksterhuis and Nordgren, 2006; Ulrich et al., 2007). From the present standpoint, consciousness can be construed as a "crosstalk" medium that allows conflicting efference streams to influence action collectively, leading to *integrated actions* (Morsella and Bargh, 2011) such as our swimmer holding her breath. Absent consciousness, behavior can be influenced by only one of the efference streams, leading to *un-integrated actions* (Morsella and Bargh, 2011) such as unconsciously inhaling while underwater, or, in another common example, reflexively dropping a carelessly made latte at Starbucks, because it feels too hot. As mentioned above, the integration afforded by consciousness involves high-level information that can be polysensory, and occurs at a stage of processing "beyond" that of the traditional Fodorian module (Fodor, 1983). The information that is represented consciously (or, in the "conscious field"; Morsella, 2005) can be considered the output of systems that are usually consciously impenetrable: In this sense, one may be able to suppress dropping the latte, but one cannot suppress the subjective urge to perform the act. From this standpoint, conscious crosstalk permits important information (or outputs) to be broadcasted to the systems responsible for skeletomotor action.

In summary, the difference between unconscious action (i.e., reflexes and the like) and conscious action is that the former is always a case of un-integrated action, and the latter *can* be a case of "integrated action." Our central claim here is that integrated action occurs when two (or more) action plans – that might normally influence behavior on their own – simultaneously co-activate and try to influence the same skeletal muscle effector at the same moment in time (Morsella and Bargh, 2011). It follows then that integrated action in every day life occurs when one: *holds* one's breath, *refrains* from dropping a hot latte, *does not* scratch an itch, or breathes *faster than normal* on purpose (e.g., for some reward). In ours and others' academic studies, integrated actions occur when participants are asked to do things like suppress a prepotent response in a laboratory paradigm such as the Stroop Task. (See Morsella et al., 2011, for a quantitative review of laboratory evidence supporting SIT.)

THE SKELETAL MUSCLE EFFECTOR SYSTEM

The skeletal muscle effector system differs substantively from most effector systems in the body (e.g., smooth muscle) in that distinct brain regions and brain systems try to control it in different – and often opposing – ways. From this standpoint, skeletal muscle is like a single steering wheel controlled simultaneously by multiple agentic systems. Each of these agentic systems has its own particular operating principles, phylogenetic origins, and concerns. While motor programs are instantiated by unconscious algorithms (Rosenbaum, 2002), the selection of higher level action *goals* happens because conscious states are able to crosstalk, which in turn leads to constraint and curbing of skeletomotor output. For example, one system in a chef's body "protests" when she accidentally touches a hot pot in her kitchen, but another system reinforces another act just as accidental when she mindlessly brings sugar to her lips in a moment of thought. As in the case of our chef, people are conscious of the tendencies (e.g., the urges and cravings) of these systems, but not necessarily of the factors engendering the tendencies themselves (tissue damage versus the relative rarity of sugar in nature; Nisbett and Wilson, 1977; Baker et al., 2004).

It has been known since at least the nineteenth century that skeletal muscle (or "striated muscle") is the only bodily effector system that *can be* (though often it is not) controlled consciously. However, why this is so has never been addressed theoretically. SIT is – in essence – a systematic reinterpretation of this age-old fact: *Skeletomotor actions are at times "consciously mediated" because these actions are directed by multiple, encapsulated systems that require conscious states to crosstalk and yield adaptive action, especially when the systems are in conflict* (Morsella, 2005). Although identifying still higher level systems is beyond the present purview of SIT, PRISM has correctly predicted that certain aspects of emotional behaviors, reproductive behaviors, parental care, and addiction-related behaviors should be coupled with conscious states, because they all exert influence over skeletal muscle plans.

It should be emphasized that there is nothing intrinsically special about skeletal muscle that causes it to be related to conscious states. Conscious processing distinguishes itself from unconscious processing not simply because it involves skeletal muscle, but because of the particular *way* conscious processing involves skeletal muscle: encapsulated systems in the brain vie to implement their own concerns over the organism in the arena of skeletomotor action planning. Yet, it is important to keep in mind that skeletal muscle is often controlled without conscious mediation, like when a person reading an academic paper shifts his posture, blinks, breathes, or yawns.

MARRYING CONSCIOUSNESS TO THE ANTICIPATORY, PHYSIOLOGICAL SYSTEM IT SUBSERVES

A primary strength of this approach is that, instead of trying to reverse engineer the purpose of consciousness by examining all that consciousness is capable of doing, it integrates consciousness with most basic of physiological processes it evolved to subserve. From this view, consciousness is one of many processes in the service of adaptive skeletomotor control, which is not surprising given that the primary function of the entire nervous system is to activate the right muscles at the right time. Richard Dawkins notes this succinctly, "The main way in which brains actually contribute to the success of survival machines is by controlling and coordinating the contraction of muscles" (Dawkins, 1976, p. 49). And Roe and Simpson (1958) propose that, in evolutionary history, overt action is *the* critical product of a nervous system, because natural selection can operate only on overt action.

Our approach outlines how consciousness is a phenomenon falling squarely within the *somatic nervous system* (**Figure 1**); it is within the somatic system that instrumental actions (e.g., holding a hot cup of coffee) are achieved through the mysterious phenomenon of *direct cognitive control* (Morsella et al., 2009c). Interestingly, direct cognitive control is probably best exemplified by one's ability to immediately control the direction of thought or the movements of a finger or arm (or any other skeletal muscle effectors). Further, when *direct control* is unavailable, indirect forms of control can be implemented. For example, while it is clear that one may not be able to directly influence one's affective/incentive states at will (Öhman and Mineka, 2001), a nurse can watch her favorite comedy to cheer herself up after a trying day watching people suffer. In other words, regarding direct cognitive control, no one can make oneself intentionally become frightened, happy, angry, sad, or become hungry if the adequate conditions are absent. Yet, people use indirect cognitive control to seek and even pay for certain experiences (e.g., going to movies or comedy clubs) to put themselves in a desired state that cannot be instantiated through an act of will.

While instrumental use of the skeletomotor system involves direct cognitive control, an additional component is often required: a mental representation of the instrumental consequences of action. For instance, we must have the idea of what a cup of coffee looks or feels like in our hands in order to perform this type of instrumental action. Because of this, the skeletomotor system is – by nature – highly anticipatory (Frith et al., 2000; Berthoz, 2002; Llinás, 2002). The operating principles of the directed actions of this system are perhaps best understood in terms of the historical notion of *ideomotor processing* (Greenwald, 1970; Hommel et al., 2001; Hommel, 2009; Hommel and Elsner, 2009). Ideomotor theory holds that the mental image of an instrumental action tends to lead to the execution of that action (Lotze, 1852; Harleß, 1861; James, 1890), with the motor programming involved being unconscious (James, 1890). Simply imagining moving your right arm to reach out for that coffee cup makes the action more likely to occur (for a treatment of why the motor programs involved are unconscious, see Gray, 1995, 2004; Grossberg, 1999; Prinz, 2003). Originating in the times of Lotze (1852), Harleß (1861), and Carpenter (1874), the hypothesis states that action guidance and action knowledge are limited to perceptual-like representations (or, *event codes*; cf., Hommel et al., 2001) of action outcomes (e.g., the "image" of one's finger flexing; Gray, 1995, 2004; Rossetti, 2001; Rosenbaum, 2002; Jeannerod, 2006). (See neuroimaging evidence for the ideomotor principle in Melcher et al., 2008.) From this standpoint, conscious contents regarding ongoing action are primarily of the perceptual consequences of action (Jeannerod, 2006).

Ideomotor processing is evident in the following anecdote (mentioned in Berger et al., 2012). The television program *60 Minutes* presented a story about how, with today's technological developments, patients can control robotic arm/limb prostheses. In the episode, the *60 Minutes* interviewer was surprised to learn that a soldier who had tragically lost his lower arm in combat could, in just a few trials, control the grasping motions of a robotic hand. The robot hand was connected to an array of electrodes attached to the muscles of the intact part of the soldier's upper arm. The interviewer asked the soldier how, when operating the prostheses for only a few trials, it was possible to know which muscles to activate in order to have the robot enact a particular action. The soldier replied to the effect that he had no idea regarding which muscles to activate, nor what the muscles were actually doing. Rather, the soldier claimed that, to enact any action on the part of the robotic arm, all that had to be done was imagine the grasping action. This image, what Harleß in the nineteenth century called in German the *Effektbild* (in English, "the picture or image of the effect"), was somehow translated (unconsciously) into the kind of muscular activation that would normally result in a grasping action. (Additional evidence for ideomotor theory stems from *response-effect compatibility* paradigms; Kunde, 2001, in which anticipated action consequences influence how quickly one executes a given action; cf., Hubbard et al., 2011.)

These images (or mental representations) tend to mirror the real-world perceptual aspects of their outcomes (i.e., the mental representation of holding a coffee cup involves haptic and visual information, etc., Hommel, 2009). This is obvious in the case of subvocalizing (i.e., talking in one's head). The imagery of the act is isomorphic in some sense to the act (Morsella and Bargh, 2010). Once an action outcome (e.g., grasping a cup) is selected, unconscious motor efference streams enact the action by activating the right muscles at the right time.

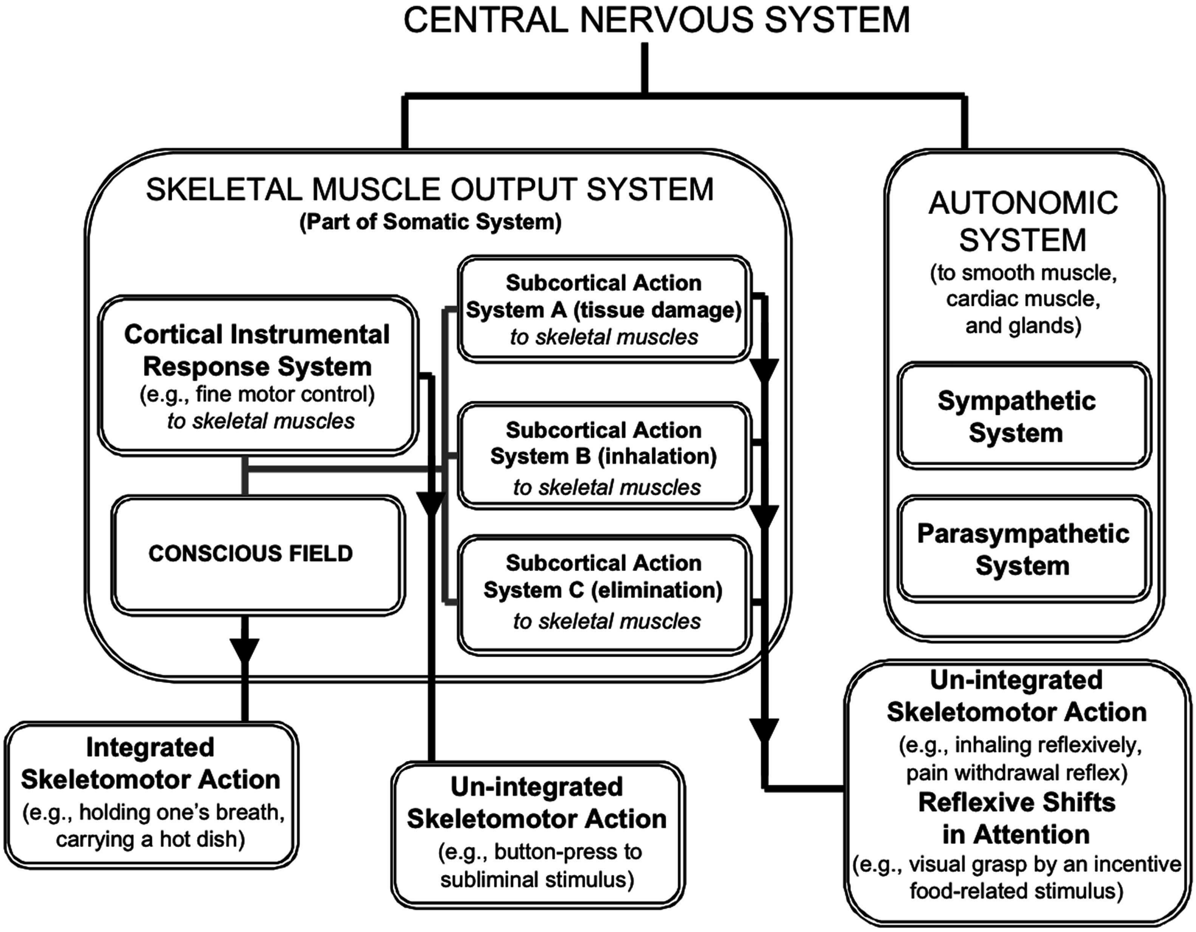

FIGURE 1 | The major divisions of the nervous system and the circumscribed province of conscious processing within the system. The major divisions include the Somatic and Autonomic systems. Within the former, Fodorian modules operate within a few multimodal, supramodular response systems (the cortical *instrumental system* and subcortical action systems), each defined by its concern (e.g., tissue damage and elimination). The instrumental system can control fine motor acts through ideomotor processing. Afference binding within systems can be unconscious. Although response systems can influence action directly, as in the case of un-integrated actions, only in virtue of conscious states can multiple response systems interact and influence action collectively, as when one holds one's breath.

There are, of course, also times when mental representations inhibit or cannot lead to the performance of instrumental goals. For example, though the system can represent leaping over a tall building in a single bound, limitations of the body prevent the action from occurring. Conversely, a woman holding an overheated latte can be prevented from sipping it because the "incentive systems" that are concerned with bodily needs curb her against inflicting tissue damage through one's skeletomotor actions (Morsella, 2005; Morsella et al., 2009b). Interestingly, because of the anticipatory nature of ideomotor processing, the same stimulus (a cup) can elicit different action tendencies, with each tendency serving the same overarching goal (grasping a cup and bringing it to the mouth; see Lashley, 1942). In motor control, this is known as "motor equivalence" (Lashley, 1942). (For Lashley's conceptualization of the role of consciousness in behavior, see Lashley, 1923.) Thus, while adaptive skeletomotor action requires integration and anticipation, it also – in some cases – requires a more elaborate form of anticipation: mental simulation/representation (Schacter and Addis, 2007).

THE ECHO HYPOTHESIS AND NATURE OF CONSCIOUS BROADCASTING

Regarding the nuts and bolts of conscious broadcasting, one intriguing hypothesis stemming from observations of phenomena involving backward masking (Breitmeyer and Ögmen, 2006) and other forms of masking (e.g., object-substitution masking; Di Lollo et al., 2000) is that, for a representation to be a conscious representation, the initial modules that constructed the representation must then, in turn, receive feedback activation about that representation. An interesting aspect of consciousness is that these representations are broadcasted and available globally (Baars, 2002). Perhaps, if visual modules X and Y construct a representation for broadcast, that representation becomes conscious only after feedback activation from the broadcast returns to these two modules, much like an echo (Di Lollo et al., 2000; Fahrenfort et al., 2007). This may be because (a) this echoic (or, "reentrant";

Fahrenfort et al., 2007) processing is a necessary ingredient for the generation of consciousness (Di Lollo et al., 2000), or (b) simply because consciousness requires involvement of frontal cortex, which, after receiving the broadcast, must send top-down activation back to the modules for the representation to be conscious (Boly et al., 2011). It may also be for other, less interesting reasons, such as (c) conscious representations require a substantial amount of activation (Kinsbourne, 1996), and reentrant feedback results in this necessary increase in activation, or that (d), for a representation to be conscious, it must be activated for a long time (Lau, 2009), something that can be achieved through feedback and sustained reverberation (Hebb, 1949). Regardless of the mechanism by which feedback may be necessary for turning an unconscious representation into a conscious one, the echo hypothesis is a falsifiable proposal that can further illuminate the component processes giving rise to conscious states.

Because of the broadcasting in the conscious field, a representation is then available to more systems than just the one that produced it. Critical for a successful broadcast of any kind of information (in any system) is that there be "receivers" capable of detecting and processing the information. The nature of such receivers remains mysterious, but one can surmise that, regarding the representations at play and with respect to such receivers, these representations must possess properties that make them communicable across a wide range of brain systems, including those concerned with action control. Indeed, there is independent evidence for the notion that the representations involved in consciousness happen to be highly broadcastable (Fodor, 1983; see treatment in Godwin et al., in press). In a cyclical manner, after each broadcast, each concerned system evaluates the outputs in the field and then generates its own output, which then influences the content of the field (Baumeister and Masicampo, 2010; Morsella and Bargh, 2010). In this way, the field changes in a self-evolving manner. This is perhaps best illustrated by way of example. Imagine a student in a lecture who suddenly gets an incessant tickle in his throat and wants to cough (i.e., an "action goal" of coughing enters the conscious field). This want, however, also leads to another action goal – to not make noise in the class during the lecture. In turn, this most recent goal could lead to the willful activation of a memory of a hysterical moment from a movie to distract him from coughing, but ironically this can also lead to the action goal of suppressing a chuckle. In this way, the contents of the conscious field change over time in a multi-determined manner, with conscious contents entering it and exiting it while influencing subsequent contents, all while unconscious systems evaluate contents and contribute their own contents (Morsella and Bargh, 2010), all in the service of constraining skeletomotor action. Thus, the Jamesian stream of consciousness involves not only one conscious thought – broadcast to a plethora of receivers and leading to another conscious thought – but (a) conscious thoughts triggering unconscious processes which lead to the introduction of other conscious thoughts into the field, and (b) unconscious processes spawning their own conscious outputs, independent of field contents. Hence, the function of conscious states is not to *observe* outputs, but to allow continuous interactions among outputs and the systems that gave rise to them. Hence, perhaps it is better to compare the phenomenal field, not to a surveillance system, but to a senate (Morsella, 2005).

What we refer to as "voluntary action" occurs with all of these processes at play. The voluntary action is believed by the actor to be a function of these conscious representations (which remains possible), but it may well be that the act and the conscious representations are both determined by some other, unconscious factor (Wegner, 2002). Regardless, as mentioned above, for every voluntary act, the actor can provide through self-report an identification of a conscious content that he or she believes gave rise to the act, regardless of whether these introspections are incorrect. In the case of voluntary action, these contents tend to be anticipatory and isomorphic with action outcomes (Morsella and Bargh, 2010).

With all this in mind, it could be said that the voluntary act is, in a sense, a "loaded" action, with a heavy load of information-processing, conscious representations, and anticipatory mechanisms. This standpoint defines a voluntary action in ways more informative than the common "homuncular" definition of voluntary action – that an action is voluntary if the organism intended to do it. Our approach reveals that, unlike involuntary actions (e.g., dropping an overheated latte because of the pain withdrawal reflex), voluntary actions can be construed as a form of integrated action, which occurs when multiple action plans are co-activated and trying to influence the same skeletomotor effector. As noted by Passingham (1995), voluntary actions are special in that they can be suppressed; from present standpoint, the act of suppression (like our student suppressing his cough) is an archetypal integrated action.

SYNTHESIS

Building on the integration consensus, SIT proposes that conscious states integrate information-processing structures and nervous events that would otherwise be independent. According to SIT, the integration involved is primarily related to the skeletal muscle output system – where anticipatory processes play a central role – and, through a form of broadcasting, this integration controls and guides voluntary action, often via ideomotor mechanisms. Importantly, SIT is unique in its ability to explain subjective data from (a) intersensory conflicts, (b) smooth muscle conflicts, and (c) conflicts from skeletomotor conflicts (e.g., holding one's breath and Stroop-like interference). SIT also explains why skeletal muscle is "voluntary" muscle.

Throughout the process of evolution, there has been a trend toward increased compartmentalization of function in the nervous system (Allman, 2000). In phylogeny, the introduction of new structures such as organs and tissues involves complex, often competitive interactions with previously existing ones. This problem, known as the "struggle of parts" problem (cf., Mayr, 2001), may have been a particularly formidable challenge during the evolution of something as complex as the human nervous system and could have led to various forms of "integrative solutions," including unconscious reflexes (Sherrington, 1906; Campbell, 1993) and neural convergence (Damasio, 1989).

A fundamental assumption of our approach is that, although crosstalk between high-level action systems *could* conceivably occur without something like conscious states, such a solution was not selected in our evolutionary history. Instead, for reasons that only the happenstance process of evolution could explain (Simpson, 1949; Gould, 1977), these specific physical adaptations seem to have been selected to solve this large-scale, crosstalk problem (Morsella, 2005). Certainly, it is easy to imagine integrated actions (e.g., suppressing a chuckle) occurring without anything like conscious states, but, then again, there are many solutions to phylogenic problems that the human body did not arrive at by way of evolution. SIT aims to take an inductive and descriptive approach at understanding nervous function "as is," and not as it (perhaps) should be. This makes SIT a *descriptive* rather than *normative* theory; and intuitions regarding how the nervous system *should* work (to be "optimal") take a back seat to actual data revealing the manner in which it actually works (even if it is suboptimal). Hence, while some theorists have proposed that consciousness is "epiphenomenal," (i.e., serving no function), it seems premature to arrive at such a conclusion until there is a sufficient scientific understanding about the place of consciousness in nature.

CRITICISMS AND ALTERNATIVE EXPLANATIONS FOR THE CURRENT APPROACH

SIT contrasts the actions of the skeletal muscle effector system with the actions of smooth muscle (e.g., the pupillary reflex), but it is possible that this juxtaposition could be criticized *a priori* as a false comparison because the behaviors of smooth muscle are not seen as a veritable form of action. From the point of view of such a critique, processes including the pupillary reflex, peristalsis, digestion, breathing, and other "vegetative" organismic actions should not be compared to what is commonly regarded as a typical form of action (e.g., blinking voluntarily). While we have to allow that these smooth muscle actions do not *feel* like actions, any denial of these phenomena under the category of action would exclude them, not on the basis of how an agnostic observer might see them, but because of the intuitions humans hold about the sources of their actions. If we imagine an intelligent non-human observer (e.g., an imaginary, extraterrestrial ethologist) studying the every actions humans are capable of, events such as the pupillary reflex would be worthy of being "coded" as an action just as certainly as a voluntary closing of the eyes (e.g., a wink) or an involuntary closing of the eyes (e.g., a reflexive blink; Skinner, 1953).

A second – intuitively intriguing – criticism could be that there are many aspects of conscious experience that have little or no connection to skeletal muscle plans. This criticism is rightly stated, indeed. However, in response to this criticism, it is important to distinguish the *primary* role of evolutionary adaptations from their secondary roles and current uses (Lorenz, 1963; Gould, 1977). A scientist could argue, for example, that color perception evolved for selecting fruits and detecting camouflaged prey and no sophisticated observer would counter that color perception could also be used to appreciate a painting. In fact, most people easily appreciate the idea that the color harmony of a painting is beautiful to us – at least in part – because it involves the kinds of stimuli that are of adaptive significance in another context. Similarly, SIT proposes that the *original* and *primary* function of conscious states was (and is) to integrate conflicting action plans involving skeletal muscle, not that all future and possible benefits of consciousness will be encapsulated in this single benefit.

Supramodular interaction theory proposes that conscious states involve broadcasts of the "outputs" of response systems that may conflict with the tendencies of other systems and that the outputs from response systems incessantly modulate one's consciousness, regardless of whether there is inter-system conflict or not. Hence, there is *chronic engagement* among the systems (Morsella, 2005), assuring that no resources, time, or "intelligent homunculus" are required to decide which outputs should participate in the conscious field at a given time. That rich intelligence is embedded in the inherent structure of the apparatus, as in the case of many evolutionary products (Simpson, 1949).

It is easy to imagine a more efficient arrangement, such one that invokes conscious states only under conditions of conflict. However, chronic engagement solves the problem at a more parsimonious level. Consider that traffic lights, pool filters, and ball-return machines at bowling alleys operate and expend energy continuously, regardless of whether their function is presently needed. These systems were chosen because the cost (in this case technologically, and hence monetarily) of adding an additional detection mechanism that activates the apparatus when it is needed is greater than the benefit of what would likely be a very complex and intricate system. In this way, chronic engagement is "efficiently inefficient" in the sense that it does not require additional mechanisms to determine whether channels of crosstalk should be open or closed (Morsella, 2005). Such deceptively "inefficient" solutions can be observed in biological functions outside the nervous system, as in most biological filters (e.g., the kidneys) which continuously filter a substrate regardless of the status of the substrate.

Chronic engagement also gives rise to the oft-mentioned monitoring role of the conscious field (e.g., Angell, 1907; Norman and Shallice, 1980). However, it is misleading to characterize the field as merely supervising the outputs of response systems because the function of the field is not to *observe* outputs, but to *allow continuous interactions* among them. To build on the analogy of a senate, the senators (systems) must always be in attendance, regardless of whether they should sit quietly or debate (Morsella, 2005). Because the outputs of all the systems are always phenomenally represented (whether they are helpful or not), one experiences the subjective experience of pain even when feeling the pain is at the moment not conducive to adaptive action. And there is no way for an actor to "tell himself" that, because he needs to lose 30 pounds for a movie role, he will not experience hunger, even though the ultimately adaptive behavior for him in the modern context is to secure a role in the film.

A seeming mystery that engenders a third criticism is as follows, if conscious states are primarily for skeletomotor action, then why do conscious states continue to exist even when the skeletal muscle system is deactivated because of, for example, damage to the nervous system or a congenital disorder? In response to this criticism, one should consider the following analogy. Consider that many of today's automobiles contain navigational systems whose primary function is to help navigate the car to one's desired destination. With this in mind, it is conceivable that the navigational system would continue to function despite problems with, say, the transmission of the car. In a similar way, central conscious processes, whose primary function was serving skeletomotor action, can continue to function even after the peripheral structures that they are intended to serve are non-operational. It is often the case in situations where the body has been rendered ineffective, though effectors or efference generators are compromised, that the central processes that subserve the efferent processes remain intact (similar decoupling of central conscious processing from peripheral events occurs in *phantom limb*; Ramachandran, 1999). In short, consciousness is a system meant to integrate actions of systems that influence skeletal muscle, but it is not dependent on the current capacity for skeletomotor action.

We now take a moment to address alternative explanations about the phenomena we sought to explain in the selective review. First, one may argue that, in attempting to describe the function of consciousness in the nervous system, instead of proposing a framework such as SIT, it is more parsimonious to simply hypothesize that the primary role of consciousness is to suppress actions, for holding one's breath, carrying a hot plate of food, or performing response interference tasks (e.g., the Stroop task) involves response suppression. However, this fails to account for the role of conscious states in integrated actions such as breathing faster for some reward, which requires inter-system crosstalk but no suppression.

Second, because novel skeletomotor actions tend to be executed consciously, one may argue that the function of consciousness is, not one of establishment of crosstalk for the purpose of integration, but instantiating stimulus-response relationships that are "arbitrary." One problem with this intriguing hypothesis is that (a) it is difficult to define what constitutes an "arbitrary" mapping, (b) there are countless cases of unconscious processes that seem to involve arbitrary mappings, as in the case of motor programming (Grossberg, 1999; Rosenbaum, 2002), and (c) some non-arbitrary mappings (e.g., holding one's breath leads to a negative subjective state) never become unconscious, despite extensive training and an inordinate amount of rehearsing the stimulus-response mappings. Moreover, unlike SIT, this hypothesis fails to explain why smooth muscle actions and intersensory conflicts are mediated unconsciously.

CONCLUSION

Supramodular Interaction Theory is a framework marrying the central advancements in knowledge from the integration consensus (chiefly that consciousness is for some type of information integration) with an explanation of why "voluntary" action is described as such. This marriage leads to an explanation of the primary function (but not the only possible function) of consciousness and gets past the tautology of calling voluntary action "voluntary" because it is able to be willed. According to SIT, the integration achieved through conscious states is primarily related to the skeletal muscle output system, where anticipatory processes play a central role as in the case of ideomotor control. SIT is unique in that, while marrying consciousness to the physiological processes it subserves, it explains subjective data from (a) intersensory conflicts, (b) smooth muscle conflicts, and (c) conflict from skeletomotor conflicts (e.g., holding one's breath). An obvious limitation of the current approach is that it sheds no light on why "subjectivity" is associated with the integrative functions these states appear to subserve. Thus, more than 40 years later, Shallice's (1972) conclusion that consciousness is an unsolved anomaly within the scientific approach still rings true. Nevertheless, the findings presented above reveal some conceptual progress regarding the nature of consciousness in the brain. Today, one can perhaps propose that, if the heart can be conceptualized as a pump and the kidney as a filter, then consciousness can be conceptualized as a form of information broadcasting (or, more precisely, information integration). The new findings showing that subjective awareness requires reentrant (or echoic) processing present a promising direction in understanding the nature the broadcast/binding that consciousness seems to instantiate. The physical basis of the broadcasting associated with consciousness is most likely unlike anything else we currently understand.

ACKNOWLEDGMENT

This research was supported by the Center for Human Culture and Behavior at San Francisco State University.

REFERENCES

Alkire, M. T., Haier, R. J., and Fallon, J. H. (2000). Toward a unified theory of narcosis: brain imaging evidence for a thalamocortical switch as the neurophysiologic basis of anesthetic-induced unconsciousness. *Conscious. Cogn.* 9, 370–386.

Alkire, M. T., Hudetz, A. G., and Tononi, G. (2008). Consciousness and anesthesia. *Science* 322, 876.

Allman, J. M. (2000). *Evolving Brains.* New York: Scientific American Library.

Angel, A. (1991). The G. L. Brown lecture. Adventures in anaesthesia. *Exp. Physiol.* 76, 1–38.

Angell, J. R. (1907). The province of functional psychology. *Psychol. Rev.* 14, 61–91.

Astor-Jack, T., and Haggard, P. (2005). "Intention and reactivity," in *Attention in Action: Advances from Cognitive Neuroscience*, eds G. W. Humphreys and M. J. Riddoch (Hove: Psychology Press), 109–130.

Atkinson, A., Thomas, M., and Cleeremans, A. (2000). Consciousness: mapping the theoretical landscape. *Trends Cogn. Sci. (Regul. Ed.)* 4, 372–382.

Baars, B. J. (1988). *A Cognitive Theory of Consciousness.* Cambridge: Cambridge University Press.

Baars, B. J. (1998). The function of consciousness: reply. *Trends Neurosci.* 21, 201.

Baars, B. J. (2002). The conscious access hypothesis: origins and recent evidence. *Trends Cogn. Sci. (Regul. Ed.)* 6, 47–52.

Baars, B. J. (2005). Global workspace theory of consciousness: toward a cognitive neuroscience of human experience. *Prog. Brain Res.* 150, 45–53.

Baker, T. B., Piper, M. E., McCarthy, D. E., Majeskie, M. R., and Fiore, M. C. (2004). Addiction motivation reformulated: an affective processing model of negative reinforcement. *Psychol. Rev.* 111, 33–51.

Banks, W. P. (1995). Evidence for consciousness. *Conscious. Cogn.* 4, 270–272.

Baumeister, R. F., and Masicampo, E. J. (2010). Conscious thought is for facilitating social and cultural interactions: how simulations serve the animal-culture interface. *Psychol. Rev.* 117, 945–971.

Berger, C. C., Bargh, J. A., and Morsella, E. (2012). "The 'what' of doing: introspection-based evidence for James's ideomotor principle," in *Psychology of Self-Control*, eds A. Durante and C. Mammoliti (New York: Nova), 145–160.

Berthoz, A. (2002). *The Brain's Sense of Movement.* Cambridge, MA: Harvard University Press.

Block, N. (1995). On a confusion about a function of consciousness. *Behav. Brain Sci.* 18, 227–287.

Boly, M., Garrido, M. I., Gosseries, O., Bruno, M. A., Boveroux, P., Schnakers, C., Massimini, M., Litvak, V., Laureys, S., and Friston, K. (2011). Preserved feedforward but impaired top-down processes in the vegetative state. *Science* 332, 858–862.

Breitmeyer, B. G., and Ögmen, H. (2006). *Visual Masking: Time Slices Through Conscious and Unconscious Vision.* Oxford: Oxford University Press.

Buzsáki, G. (2006). *Rhythms of the Brain.* New York: Oxford University Press.

Callebaut, W., and Rasskin-Gutman, D. (2009). *Modularity: Understanding the Development and Evolution of Natural Complex Systems.* Cambridge, MA: MIT Press.

Campbell, N. A. (1993). *Biology*, 3rd Edn. Redwood City, CA: Benjamin/Cummings Publishing.

Carpenter, W. B. (1874). *Principles of Mental Physiology*. New York: Appleton.

Clark, A. (2002). Is seeing all it seems? Action, reason and the grand illusion. *J. Conscious. Stud.* 9, 181–202.

Cohen-Kdoshay, O., and Meiran, N. (2009). The representation of instructions operates like a prepared reflex: flanker compatibility effects found in the first trial following S–R instructions. *Exp. Psychol.* 56, 128–133.

Crick, F., and Koch, C. (1995). Are we aware of neural activity in primary visual cortex? *Nature* 375, 121–123.

Crick, F., and Koch, C. (2003). A framework for consciousness. *Nat. Neurosci.* 6, 1–8.

Damasio, A. R. (1989). Time-locked multiregional retroactivation: a systems-level proposal for the neural substrates of recall and recognition. *Cognition* 33, 25–62.

Damasio, A. R. (2010). *Self Comes to Mind: Constructing the Conscious Brain*. New York: Pantheon.

Dawkins, R. (1976). *The Selfish Gene*. New York: Oxford University Press.

Dehaene, S., and Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition* 79, 1–37.

Del Cul, A., Baillet, S., and Dehaene, S. (2007). Brain dynamics underlying the nonlinear threshold for access to consciousness. *PLoS Biol.* 5, e260. doi:10.1371/journal.pbio.0050260

Desmurget, M., Reilly, K. T., Richard, N., Szathmari, A., Mottolese, C., and Sirigu, A. (2009). Movement intention after parietal cortex stimulation in humans. *Science* 324, 811–813.

Desmurget, M., and Sirigu, A. (2010). A parietal-premotor network for movement intention and motor awareness. *Trends Cogn. Sci. (Regul. Ed.)* 13, 411–419.

Di Lollo, V., Enns, J. T., and Rensink, R. A. (2000). Competition for consciousness among visual events: the psychophysics of reentrant visual pathways. *J. Exp. Psychol. Gen.* 129, 481–507.

Dijksterhuis, A., and Nordgren, L. F. (2006). A theory of unconscious thought. *Perspect. Psychol. Sci.* 1, 95–109.

Doesburg, S. M., Green, J. L., McDonald, J. J., and Ward, L. M. (2009). Rhythms of consciousness: binocular rivalry reveals large-scale oscillatory network dynamics mediating visual perception. *PLoS ONE* 4, e6142. doi:10.1371/journal.pone.0006142

Eriksen, B. A., and Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Percept. Psychophys.* 16, 143–149.

Fahrenfort, J. J., Scholte, H. S., and Lamme, V. A. (2007). Masking disrupts reentrant processing in human visual cortex. *J. Cogn. Neurosci.* 19, 1488–1497.

Fecteau, J. H., Chua, R., Franks, I., and Enns, J. T. (2001). Visual awareness and the online modification of action. *Can. J. Exp. Psychol.* 55, 104–110.

Fehrer, E., and Biederman, I. (1962). A comparison of reaction time and verbal report in the detection of masked stimuli. *J. Exp. Psychol.* 64, 126–130.

Fehrer, E., and Raab, D. (1962). Reaction time to stimuli masked by metacontrast. *J. Exp. Psychol.* 63, 143–147.

Flohr, H. (1995). An information processing theory of anaesthesia. *Neuropsychologia* 9, 1169–1180.

Fodor, J. A. (1983). *Modularity of Mind: An Essay on Faculty Psychology*. Cambridge, MA: MIT Press.

Freeman, W. J. (1991). The physiology of perception. *Sci. Am.* 264, 78–85.

Fries, P. (2005). A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends Cogn. Sci. (Regul. Ed.)* 9, 474–480.

Frith, C. D., Blakemore, S. J., and Wolpert, D. M. (2000). Abnormalities in the awareness and control of action. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 355, 1771–1788.

Gaillard, R., Dehaene, S., Adam, C., Clémenceau, S., Hasboun, D., Baulac, M., Cohen, L., and Naccache, L. (2009). Converging intracranial markers of conscious access. *PLoS Biol.* 7, e1000061. doi:10.1371/journal.pbio.1000061

Godwin, C. A., Gazzaley, A., and Morsella, E. (in press). "Homing in on the brain mechanisms linked to consciousness: buffer of the perception-and-action interface," in *The Unity of Mind, Brain and World: Current Perspectives on a Science of Consciousness*, eds A. Pereira and D. Lehmann's (Cambridge: Cambridge University Press).

Goodale, M., and Milner, D. (2004). *Sight Unseen: An Exploration of Conscious and Unconscious Vision*. New York: Oxford University Press.

Gould, S. J. (1977). *Ever Since Darwin: Reflections in Natural History*. New York: Norton.

Gray, J. A. (1995). The contents of consciousness: a neuropsychological conjecture. *Behav. Brain Sci.* 18, 659–676.

Gray, J. A. (2004). *Consciousness: Creeping Up on the Hard Problem*. New York: Oxford University Press.

Greenwald, A. G. (1970). Sensory feedback mechanisms in performance control: with special reference to the ideomotor mechanism. *Psychol. Rev.* 77, 73–99.

Grossberg, S. (1999). The link between brain learning, attention, and consciousness. *Conscious. Cogn.* 8, 1–44.

Haggard, P., Aschersleben, G., Gehrke, J., and Prinz, W. (2002). "Action, binding and awareness," in *Common Mechanisms in Perception and Action: Attention and Performance*, Vol. XIX, eds W. Prinz and B. Hommel (Oxford: Oxford University Press), 266–285.

Hallett, M. (2007). Volitional control of movement: the physiology of free will. *Clin. Neurophysiol.* 117, 1179–1192.

Harleß, E. (1861). Der apparat des willens [The apparatus of the will]. *Zeitshrift für Philosophie und philosophische Kritik* 38, 499–507.

Heath, M., Neely, K. A., Yakimishyn, J., and Binsted, G. (2008). Visuomotor memory is independent of conscious awareness of target features. *Exp. Brain Res.* 188, 517–527.

Hebb, D. O. (1949). *The Organization of Behavior: A Neuropsychological Theory*. New York: Wiley.

Hommel, B. (2000). "The prepared reflex: automaticity and control in stimulus-response translation," in *Control of Cognitive Processes: Attention and Performance XVII*, eds S. Monsell and J. Driver (Cambridge, MA: MIT Press), 247–273.

Hommel, B. (2009). Action control according to TEC (theory of event coding). *Psychol. Res.* 73, 512–526.

Hommel, B., and Elsner, B. (2009). "Acquisition, representation, and control of action," in *Oxford Handbook of Human Action*, eds E. Morsella, J. A. Bargh, and P. M. Gollwitzer (New York: Oxford University Press), 371–398.

Hommel, B., Müsseler, J., Aschersleben, G., and Prinz, W. (2001). The theory of event coding: a framework for perception and action planning. *Behav. Brain Sci.* 24, 849–937.

Hubbard, J., Gazzaley, A., and Morsella, E. (2011). Traditional response interference effects from anticipated action outcomes: a response-effect compatibility paradigm. *Acta Psychol. (Amst.)* 138, 106–110.

Hudetz, A. G. (2006). Suppressing consciousness: mechanisms of general anesthesia. *Semin. Anesth.* 25, 196.

Hummel, F., and Gerloff, C. (2005). Larger interregional synchrony is associated with greater behavioral success in a complex sensory integration task in humans. *Cereb. Cortex* 15, 670–678.

Jackson, F. (1982). Epiphenomenal qualia. *Philos. Q.* 32, 127–136.

James, W. (1890). *Principles of Psychology*. New York: Holt.

Jeannerod, M. (2006). *Motor Cognition: What Action Tells the Self*. New York: Oxford University Press.

John, E. R., and Prichep, L. S. (2005). The anesthetic cascade: a theory of how anesthesia suppresses consciousness. *Anesthesiology* 102, 447.

Johnson, H., and Haggard, P. (2005). Motor awareness without perceptual awareness. *Neuropsychologia* 43, 227–237.

Kern, M. K., Safwan, J., Arndorfer, R. C., and Shaker, R. (2001). Cerebral cortical representation of reflexive and volitional swallowing in humans. *Am. J. Physiol. Gastrointest. Liver Physiol.* 280, G354–G360.

Kinsbourne, M. (1996). "What qualifies a representation for a role in consciousness?" in *Scientific Approaches to Consciousness*, eds J. D. Cohen and J. W. Schooler (Hillsdale, NJ: Erlbaum), 335–355.

Koch, C. (2004). *The Quest for Consciousness: A Neurobiological Approach*. Colorado: Roberts & Company.

Ku, S., Lee, U., Noh, G., Jun, I., and Mashour, G. A. (2011). Preferential inhibition of frontal-to-parietal feedback connectivity is a neurophysiologic correlate of general anesthesia in surgical patients. *PLoS ONE* 6, e25155. doi:10.1371/journal.pone.0025155

Kunde, W. (2001). Response-effect compatibility in manual choice reaction tasks. *J. Exp. Psychol. Hum. Percept. Perform.* 27, 387–394.

Långsjö, J. W., Alkire, M. T., Kaskinoro, K., Hayama, H., Maksimow, A., Kaisti, K. K., Aalto, S., Aantaa, R., Jääskeläinen, S. K., Revonsuo, A., and Scheinin, H. (2012). Returning from oblivion: imaging the neural core of consciousness. *J. Neurosci.* 32, 4935–4943.

Lashley, K. S. (1942). "The problem of cerebral organization in vision," in *Visual Mechanisms. Biological Symposia*, Vol. 7, ed. H. Kluver (Lancaster, PA: Cattell Press), 301–322.

Lashley, Karl S. (1923). The behavioristic interpretation of consciousness.

Psychol. Bull. 30, 237–272 (Part 1), 329–353 (Part 2).

Lau, H. C. (2009). "Volition and the functions of consciousness," in *The Cognitive Neurosciences*, Vol. 4, ed. M. Gazzaniga (Cambridge, MA: MIT Press), 1191–1200.

Laureys, S. (2005). The neural correlate of (un)awareness: lessons from the vegetative state. *Trends Cogn. Sci. (Regul. Ed.)* 12, 556–559.

Laureys, S., Faymonville, M. E., Degueldre, C., Del Fiore, G., Damas, P., Lambermont, B., Janssens, N., Aerts, J., Franck, G., Luxen, A., Moonen, G., Lamy, M., and Maquet, P. (2000a). Auditory processing in the vegetative state. *Brain* 123, 1589–1601.

Laureys, S., Faymonville, M. E., Luxen, A., Lamy, M., Franck, G., and Maquet, P. (2000b). Restoration of thalamocortical connectivity after recovery from persistent vegetative state. *Lancet* 355, 1790–1791.

Lee, U., Mashour, G. A., Kim, S., Noh, G., and Choi, B. (2009). Propofol induction reduces the capacity for neural information integration: implications for the mechanism of consciousness and general anesthesia. *Conscious. Cogn.* 18, 56–64.

Levine, J. (1983). Materialism and qualia: the explanatory gap. *Pac. Philos. Q.* 64, 354–361.

Lewin, K. (1935). *A Dynamic Theory of Personality*. New York: McGraw-Hill.

Lhermitte, F. (1983). 'Utilization behavior' and its relation to lesions of the frontal lobes. *Brain* 106, 237–255.

Libet, B. (2004). *Mind Time: The Temporal Factor in Consciousness*. Cambridge, MA: Harvard University Press.

Llinás, R. R. (2002). *I of the Vortex: From Neurons to Self*. Cambridge, MA: MIT Press.

Llinás, R. R., and Ribary, U. (2001). Consciousness and the brain: the thalamocortical dialogue in health and disease. *Ann. N. Y. Acad. Sci.* 929, 166–175.

Logan, G. D., Taylor, S. E., and Etherton, J. L. (1999). Attention and automaticity: toward a theoretical integration. *Psychol. Res.* 62, 165–181.

Logothetis, N. K., and Schall, J. D. (1989). Neuronal correlates of subjective visual perception. *Science* 245, 761–762.

Lorenz, K. (1963). *On Aggression*. New York: Harcourt, Brace, & World.

Lotze, R. H. (1852). *Medizinische Psychologie oder Physiologie der Seele*. Leipzig: Weidmann'sche Buchhandlung.

Magen, H., and Cohen, A. (2010). Modularity beyond perception: evidence from the PRP paradigm. *J. Exp. Psychol. Hum. Percept. Perform.* 36, 395–414.

Marchetti, C., and Della Sala, S. (1998). Disentangling the alien and anarchic hand. *Cogn. Neuropsychiatry* 3, 191–207.

Mashour, G. A. (2004). Consciousness unbound: toward a paradigm of general anesthesia. *Anesthesiology* 100, 428–433.

Mashour, G. A. (2006). Integrating the science of consciousness and anesthesia. *Anesth. Analg.* 103, 975–982.

Mayr, E. (2001). *What Evolution Is*. London: Weidenfeld & Nicolson.

McGurk, H., and MacDonald, J. (1976). Hearing lips and seeing voices. *Nature* 264, 746–748.

Melcher, T., Weidema, M., Eenshuistra, R. M., Hommel, B., and Gruber, O. (2008). The neural substrate of the ideomotor principle: an event-related fMRI analysis. *NeuroImage* 39, 1274–1288.

Merker, B. (2007). Consciousness without a cerebral cortex: a challenge for neuroscience and medicine. *Behav. Brain Sci.* 30, 63–134.

Miller, J., and Ulrich, R. (2003). Simple reaction time and statistical facilitation: a parallel grains model. *Cogn. Psychol.* 46, 101–151.

Miller, N. E. (1959). "Liberalization of basic S-R concepts: extensions to conflict behavior, motivation, and social learning," in *Psychology: A Study of Science*, Vol. 2, ed. S. Koch (New York: McGraw-Hill), 196–292.

Morsella, E. (2005). The function of phenomenal states: supramodular interaction theory. *Psychol. Rev.* 112, 1000–1021.

Morsella, E., and Bargh, J. A. (2010). What is an output? *Psychol. Inq.* 21, 354–370.

Morsella, E., and Bargh, J. A. (2011). "Unconscious action tendencies: sources of 'un-integrated' action," in *The Handbook of Social Neuroscience*, eds J. T. Cacioppo and J. Decety (New York: Oxford University Press), 335–347.

Morsella, E., Berger, C. C., and Krieger, S. C. (2011). Cognitive and neural components of the phenomenology of agency. *Neurocase* 17, 209–230.

Morsella, E., Gray, J. R., Krieger, S. C., and Bargh, J. A. (2009a). The essence of conscious conflict: subjective effects of sustaining incompatible intentions. *Emotion* 9, 717–728.

Morsella, E., Krieger, S. C., and Bargh, J. A. (2009b). "The function of consciousness: why skeletal muscles are "voluntary" muscles," in *Oxford Handbook of Human Action*, eds E. Morsella, J. A. Bargh, and P. M. Gollwitzer (New York: Oxford University Press), 625–634.

Morsella, E., Lanska, M., Berger, C. C., and Gazzaley, A. (2009c). Indirect cognitive control through top-down activation of perceptual symbols. *Eur. J. Soc. Psychol.* 39, 1173–1177.

Morsella, E., Krieger, S. C., and Bargh, J. A. (2010). Minimal neuroanatomy for a conscious brain: homing in on the networks constituting consciousness. *Neural Netw.* 23, 14–15.

Munglani, R., Andrade, J., Sapsford, D. J., Baddeley, A., and Jones, J. G. (1993). A measure of consciousness and memory during isoflurane administration: the coherent frequency. *Br. J. Anasth.* 71, 633–641.

Muzur, A., Pace-Schott, E. F., and Hobson, J. A. (2002). The prefrontal cortex in sleep. *Trends Cogn. Sci. (Regul. Ed.)* 6, 475–481.

Nagel, T. (1974). What is it like to be a bat? *Philos. Rev.* 83, 435–450.

Nisbett, R. E., and Wilson, T. D. (1977). Telling more than we can know: verbal reports on mental processes. *Psychol. Rev.* 84, 231–259.

Norman, D. A., and Shallice, T. (1980). "Attention to action: willed and automatic control of behavior," in *Consciousness and Self-Regulation*, eds R. J. Davidson, G. E. Schwartz, and D. Shapiro (New York: Plenum Press), 1–18.

Öhman, A., and Mineka, S. (2001). Fears, phobias, and preparedness: toward an evolved module of fear and fear learning. *Psychol. Rev.* 108, 483–522.

Ortinski, P., and Meador, K. J. (2004). Neuronal mechanisms of conscious awareness. *Arch. Neurol.* 61, 1017–1020.

Passingham, R. (1995). *The Frontal Lobes and Voluntary Action*. New York: Oxford University Press.

Penfield, W., and Jasper, H. H. (1954). *Epilepsy and the Functional Anatomy of the Human Brain*. New York: Little, Brown.

Pinker, S. (1997). *How the Mind Works*. New York: Norton.

Portas, C. M., Krakow, K., Allen, P., Josephs, O., Armony, J. L., and Frith, C. D. (2000). Auditory processing across the sleep–wake cycle: simultaneous EEG and fMRI monitoring in humans. *Neuron* 28, 991–999.

Prinz, W. (2003). "How do we know about our own actions?" in *Voluntary Action: Brains, minds, and Sociality*, eds S. Maasen, W. Prinz, and Roth (London: Oxford University Press), 21–33.

Ramachandran, V. S. (1999). *Phantoms in the Brain: Probing the Mysteries of the Human Mind*. New York: Harper Perennial.

Roe, A., and Simpson, G. G. (1958). *Behavior and Evolution*. New Haven, CT: Yale University Press.

Rosenbaum, D. A. (2002). "Motor control," in *Stevens' Handbook of Experimental Psychology, Vol. 1, Sensation and Perception*, 2nd Edn, eds H. Pashler (Series Ed.) and S. Yantis (Vol. Ed.) (New York: Wiley), 315–339.

Rossetti, Y. (2001). "Implicit perception in action: short-lived motor representation of space," in *Finding Consciousness in the Brain: A Neurocognitive Approach*, ed. P. G. Grossenbacher (Netherlands: John Benjamins Publishing), 133–181.

Schacter, D. L., and Addis, D. R. (2007). The cognitive neuroscience of constructive memory: remembering the past and imagining the future. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 362, 773–786.

Schmahmann, J. D. (1998). Dysmetria of thought: clinical consequences of cerebellar dysfunction on cognition and affect. *Trends Cogn. Sci. (Regul. Ed.)* 2, 362–371.

Sergent, C., and Dehaene, S. (2004). Is consciousness a gradual phenomenon? Evidence for an all-or-none bifurcation during the attentional blink. *Psychol. Sci.* 15, 720–728.

Shallice, T. (1972). Dual functions of consciousness. *Psychol. Rev.* 79, 383–393.

Sherrington, C. S. (1906). *The Integrative Action of the Nervous System*. New Haven, CT: Yale University Press.

Simpson, G. G. (1949). *The Meaning of Evolution*. New Haven, CT: Yale University Press.

Skinner, B. F. (1953). *Science and Human Behavior*. New York: Macmillan.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *J. Exp. Psychol.* 18, 643–662.

Taylor, J. L., and McCloskey, D. I. (1990). Triggering of preprogrammed movements as reactions to masked stimuli. *J. Neurophysiol.* 63, 439–446.

Taylor, J. L., and McCloskey, D. I. (1996). Selection of motor responses on the basis of unperceived stimuli. *Exp. Brain Res.* 110, 62–66.

Tononi, G., and Edelman, G. M. (1988). Consciousness and complexity. *Science* 282, 1846–1851.

Tye, M. (1999). Phenomenal consciousness: the explanatory gap as cognitive illusion. *Mind* 108, 705–725.

Ulhaas, P. J., Pipa, G., Lima, B., Melloni, L., Neuenschwander, S., Nikolic, D., and Singer, W. (2009). Neural synchrony in cortical networks: history, concept and current status. *Front. Integr. Neurosci.* 3:17. doi:10.3389/neuro.07.017.2009

Ulrich, R., Miller, J., and Schröter, H. (2007). Testing the race model inequality: an algorithm and computer programs. *Behav. Res. Methods Instrum. Comput.* 39, 291–302.

Varela, F., Lachaux, J. P., Rodriguez, E., and Martinerie, J. (2001). The brainweb: phase synchronization and large-scale integration. *Nat. Rev. Neurosci.* 2, 229–239.

Vroomen, J., and de Gelder, B. (2003). Visual motion influences the contingent auditory motion aftereffect. *Psychol. Sci.* 14, 357–361.

Wegner, D. M. (2002). *The Illusion of Conscious Will.* Cambridge, MA: MIT Press.

Wegner, D. M. (2003). The mind's best trick: how we experience conscious will. *Trends Cogn. Sci. (Regul. Ed.)* 7, 65–69.

Wegner, D. M., and Bargh, J. A. (1998). "Control and automaticity in social life," in *The Handbook of Social Psychology*, 4th Edn, Vol. 1, eds D. T. Gilbert, S. T. Fiske, and G. Lindsey (Boston, MA: McGraw-Hill), 446–496.

Wenke, D., Gaschler, R., and Nattkemper, D. (2007). Instruction-induced feature binding. *Psychol. Res.* 71, 91–106.

Zeki, S., and Bartels, A. (1999). Toward a theory of visual consciousness. *Conscious. Cogn.* 8, 225–259.

Zmigrod, S., and Hommel, B. (2011). The relationship between feature binding and consciousness: evidence from asynchronous multi-modal stimuli. *Conscious. Cogn.* 20, 586–593.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 14 June 2012; accepted: 09 September 2012;

Citation: Poehlman TA, Jantz TK and Morsella E (2012) Adaptive skeletal muscle action requires anticipation and "conscious broadcasting". Front. Psychology ***3****:369. doi: 10.3389/fpsyg.2012.00369*

This article was submitted to Frontiers in Cognition, a specialty of Frontiers in Psychology.

frontiers in
PSYCHOLOGY

HYPOTHESIS AND THEORY ARTICLE
published: 09 November 2012
doi: 10.3389/fpsyg.2012.00455

Intention concepts and brain-machine interfacing

Franziska Thinnes-Elker[1,2,3†], **Olga Iljina**[1,4†], **John Kyle Apostolides**[3,5], **Felicitas Kraemer**[6], **Andreas Schulze-Bonhage**[1], **Ad Aertsen**[2,3] and **Tonio Ball**[1,3]*

[1] Epilepsy Center, University Medical Center Freiburg, Freiburg, Germany
[2] Institute for Biology III, University of Freiburg, Freiburg, Germany
[3] Bernstein Center Freiburg, University of Freiburg, Freiburg, Germany
[4] Hermann Paul School of Linguistics, Freiburg, Germany
[5] Juniata College, Huntingdon, PA, USA
[6] Department of Philosophy and Ethics, Eindhoven University of Technology, Eindhoven, Netherlands

Edited by:
Rico Fischer, Technische Universität Dresden, Germany

Reviewed by:
Don Tucker, Electrical Geodesics, Inc. and the University of Oregon, USA
Malte Schilling, International Computer Science Institute Berkeley, USA

***Correspondence:**
Tonio Ball, Epilepsy Center, University Medical Center Freiburg, Engelbergerstr. 21, 79106 Freiburg, Germany.
e-mail: tonio.ball@uniklinik-freiburg.de

† Franziska Thinnes-Elker and Olga Iljina have contributed equally to this work.

Intentions, including their temporal properties and semantic content, are receiving increased attention, and neuroscientific studies in humans vary with respect to the topography of intention-related neural responses. This may reflect the fact that the kind of intentions investigated in one study may not be exactly the same kind investigated in the other. Fine-grained intention taxonomies developed in the philosophy of mind may be useful to identify the neural correlates of well-defined types of intentions, as well as to disentangle them from other related mental states, such as mere urges to perform an action. Intention-related neural signals may be exploited by brain-machine interfaces (BMIs) that are currently being developed to restore speech and motor control in paralyzed patients. Such BMI devices record the brain activity of the agent, interpret ("decode") the agent's intended action, and send the corresponding execution command to an artificial effector system, e.g., a computer cursor or a robotic arm. In the present paper, we evaluate the potential of intention concepts from philosophy of mind to improve the performance and safety of BMIs based on higher-order, intention-related control signals. To this end, we address the distinction between future-, present-directed, and motor intentions, as well as the organization of intentions in time, specifically to what extent it is sequential or hierarchical. This has consequences as to whether these different types of intentions can be expected to occur simultaneously or not. We further illustrate how it may be useful or even necessary to distinguish types of intentions exposited in philosophy, including yes- vs. no-intentions and oblique vs. direct intentions, to accurately decode the agent's intentions from neural signals in practical BMI applications.

Keywords: BMI, BCI, action intention, intentional, philosophy of mind

INTRODUCTION

Intentions lie at the heart of human goal-directed behavior and have been debated for centuries in the philosophy of mind. Such fundamental issues have been discussed as the role of rational thought in intention formation (Bentham, 1781; Kant, 1785; Wittgenstein, 1953; Davidson, 1963; Kiverstein, 2006; Mele, 2007), and the temporal dynamics in and across distinct stages or kinds of intending (Searle, 1983; Pacherie, 2006). Various definitions of intention have been given, and a number of classifications have been proposed.

Broadly speaking, intention can be conceived of as a mental state in some way linked to phenomena such as decision, agency, desire, and belief (e.g., Anscombe, 1963; Goldman, 1970; Bratman, 1987). It is widely, though not universally, assumed that intention is causal to intentional action (Davidson, 1963). Theories differ with respect to the question whether intentions count as distinctive mental states (the non-reductive approach) or not (the reductive approach), see Pacherie (2002) for a review and Setiya (2007) and Bratman (2009) for a recent discussion. The exact nature and definition of intention are thus a matter of debate. Here we proceed from the influential definition of intention proposed by Bratman (1987). It relies on a superordinate category of "pro-attitudes," which "play a motivational role" (1987, p. 15) in action. According to Bratman (1987), intentions and desires are distinctive mental states that fall into this category. A fundamental difference between the two is that intentions are "conduct-controlling" (1987, p. 16), whereas desires are "merely potential influencers of action" (1987, p. 16).

Owing to the advancements in neural-recording methodology over the last 50 years, various topographic, temporal, and semantic (content) manifestations of intentions in the human brain have been researched (Libet et al., 1983; Lau et al., 2004; Brass and Haggard, 2007, 2008; Haynes et al., 2007; Krieghoff et al., 2009; Bara et al., 2011) and are receiving further attention in cognitive neuroscience. The phenomenology and neurobiology of intentions are important to study for several reasons. A better understanding of causes and prerequisites for volitional behavior may aid objective evaluation of a person's actions in ethical and legal contexts (Haggard, 2008; Schleim, 2008). Furthermore, such knowledge may help to treat patients with intention-related disorders, such as anarchic hand and Tourette's syndromes (Haggard and Clark, 2003; Pacherie, 2007; Eddy et al., 2010; Edwards et al., 2011).

Conceptual input from the philosophy of intentions to other disciplines has previously proven useful. The belief-desire-intention model by Bratman (1987), for instance, was employed in computer science to develop the belief–desire–intention software model for programming intelligent agents (Rao and Georgeff, 1991). Similarly, philosophy may provide valuable input to the neuroscience of volitional action (Haggard, 2005; Mele, 2008; Pacherie and Haggard, 2010; Pacherie, 2011), and first attempts have recently been made to integrate philosophically-informed intention concepts into human neuroscience (Bara et al., 2011). Here, we propose that intention concepts from the philosophy of mind may be also usefully adopted by the emerging field of brain-machine interfacing (BMI) research and technology. In particular, we argue that intention concepts are important for BMIs utilizing higher-order intention-related brain activity, in contrast to BMIs that are based solely on inference of low-level movement parameters. The also widespread P300 BMI approaches as well as those based on learned self-regulation of brain signals remain outside the scope of this article. Recent reviews on these topics can be found in Fazel-Rezai et al. (2012) and Wolpaw et al. (2002), respectively.

The structure of this article is as follows. In Section "Insights into Intentions from Cognitive Neuroscience," we review current neuroscientific literature on intentions, and outline the core areas involved in intention-related processing in humans. In Section "Current Approaches to Brain-machine Interfacing," we address the basic principles that are currently employed in BMI-based restoration of motor and communication functions. In Section "Philosophical Taxonomies of Intentions and their Relevance to BMI," we summarize some influential philosophical notions and taxonomies of intentions, and illustrate their potential relevance for neuroscientific research in general and in particular for BMI. Finally, we draw conclusions and provide an outlook for future studies in Section "Conclusions and Outlook."

INSIGHTS INTO INTENTIONS FROM COGNITIVE NEUROSCIENCE

Interest in intention-related brain signals has grown in neuroscience over the last several decades. In their early electroencephalography (EEG) readiness-potential study, Libet et al. (1983) reported that cerebral activity before initiation of self-paced movements precedes the conscious intention to move over several 100 ms. The observed temporal differences led these authors to conclude that initiation of voluntary actions can begin unconsciously, and it is only some time later that we become aware of an intention to move. Although the reported findings and their interpretation were highly controversial (e.g., Keller and Heckhausen, 1990; Snyder et al., 1997; Haggard and Eimer, 1999), the article by Libet et al. (1983) contributed to the development of a vivid discussion about the nature of human free will, agency and voluntary movement, and was followed by a large amount of experimental studies and opinion articles concerning the neural correlates of intentional action. Consecutive research identified a widespread distribution of neural locations in the frontal, parietal, and even temporal lobes (**Figure 1**), arranged in extended cortical networks for intention-related processing (Haggard, 2008).

One cortical location that has been repeatedly activated in studies on intention-related processing is the posterior parietal cortex (PPC), in particular its inferior part (**Figure 1**). Initial evidence for the contribution of this region to intentional control comes from single-cell recordings in monkeys. Specifically, the parietal reach region (PRR) and the lateral intraparietal area (LIP) have been shown to exhibit effector-specific neuronal activity in delayed saccadic and reaching tasks (Andersen and Buneo, 2002; Quian Quiroga et al., 2006; Cui and Andersen, 2007; Andersen and Cui, 2009), suggesting that the PPC can convey neuronal information about what the animal intends to do (Snyder et al., 1997). In humans, involvement of parietal regions in intention-related processing was observed using electrical stimulation, which elicited a reported "urge to move" without consecutive execution (Assal et al., 2007; Desmurget et al., 2009), and in lesion studies showing that awareness of an intention to move is abnormal in

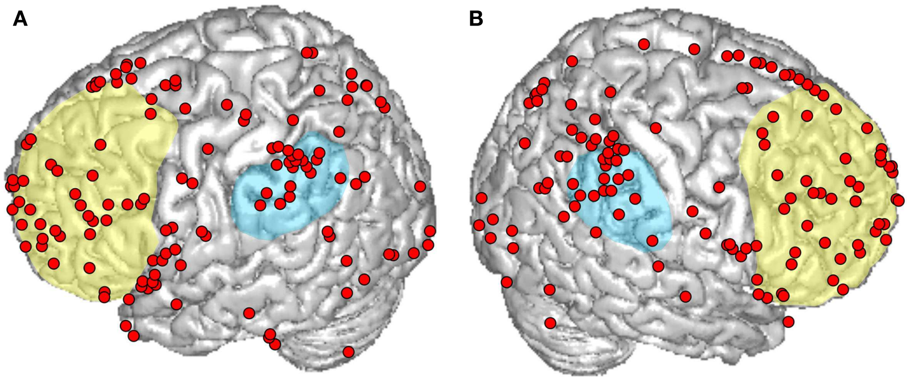

FIGURE 1 | An overview of cortical responses reported in recent functional magnetic resonance imaging (fMRI) and positron emission tomography (PET) studies that explicitly aimed at investigating intentions in healthy subjects. Peaks are plotted on a standard brain from SPM8 on **(A)** the left and **(B)** right hemisphere. The approximate locations of the prefrontal and the inferior parietal cortex are indicated in yellow and blue, respectively. Reported intention-related peaks in both hemispheres exhibit a widespread spatial distribution across the frontal, parietal, occipital, and temporal lobes.

patients with damage to the parietal cortex (Sirigu et al., 2004; Assal et al., 2007). Parietal contributions to intention encoding were also confirmed by a number of functional magnetic resonance imaging (fMRI) studies. In prospective memory tasks, both lateral and medial parietal regions showed increased blood-oxygen-level dependent (BOLD) responses that stretched from the precuneus into the anterior and posterior cingulate cortices, the intraparietal sulcus, and inferior parietal regions (Burgess et al., 2001; den Ouden et al., 2005; Eschen et al., 2007; Haynes et al., 2007; Soon et al., 2008; Gilbert, 2011; Benoit et al., 2012; Momennejad and Haynes, 2012). Investigations of non-delayed self-initiated movements reported similar neural responses in the inferior parietal lobe (Ball et al., 1999; Farrer et al., 2008; Krieghoff et al., 2009), in the intraparietal sulcus (Lau et al., 2004; Gallivan et al., 2011a,b), and in the anterior cingulate cortex (Cunnington et al., 2006; Mueller et al., 2007; Krieghoff et al., 2009), as opposed to externally triggered movements, which elicit no, or only attenuated activations in these regions (Jahanshahi et al., 1995; Jenkins et al., 2000; Mueller et al., 2007; Hoffstaedter et al., 2012).

In addition to parts of the parietal cortex, the prefrontal cortex (PFC, **Figure 1**) has been activated in many intention-related studies. Delayed intention paradigms revealed lateral and medial PFC responses, mostly in rostral prefrontal areas (Burgess et al., 2001; den Ouden et al., 2005; Simons et al., 2006; Poppenk et al., 2010; Gilbert, 2011; Benoit et al., 2012), whereas non-delayed intention experiments showed activity in the dorsal medial and lateral prefrontal regions (Lau et al., 2004; Cunnington et al., 2006; Rushworth, 2008; Gallivan et al., 2011a,b; Rosenberg-Katz et al., 2012). The frontopolar cortex (BA10) was suggested to represent a gateway mechanism for orienting attention toward external and internal stimuli, and to play a critical role in the encoding and storage of future intentions (den Ouden et al., 2005; Haynes et al., 2007; Soon et al., 2008; Uretzky and Gilboa, 2010). In accordance with the latter, clinical evidence shows that lesions in this area lead to the impaired ability to keep future intentions in mind for later execution (Burgess et al., 2001).

Intention-related information is also thought to be present in higher-order motor areas, including the supplementary motor area (SMA; Eccles, 1982; Fried et al., 1991, 2011; Jahanshahi et al., 1995; Ball et al., 1999; Jenkins et al., 2000; Lau et al., 2004; Soon et al., 2008; Hoffstaedter et al., 2012; Momennejad and Haynes, 2012), the pre-SMA (Lau et al., 2004, 2006; Cunnington et al., 2006; Mueller et al., 2007; Nachev et al., 2007), and in the dorsal and ventral premotor regions (Cunnington et al., 2006; Pesaran et al., 2006; Eschen et al., 2007; Gallivan et al., 2011a,b; Hoffstaedter et al., 2012). Since activity in the SMA and the pre-SMA typically occurs early and precedes movement execution (Fried et al., 2011), and considering that activation in the pre-SMA has been observed in relation to own intentions as opposed to own movements (Lau et al., 2004), these areas may contribute to intentional processes during preparation for action (Passingham et al., 2010).

Finally, the anterior insular cortex has been co-activated with several aforementioned areas in studies on intention encoding in the human brain (Jahanshahi et al., 1995; Mueller et al., 2007; Krieghoff et al., 2009; Hoffstaedter et al., 2012). Insular activation has been proposed to subserve evaluation of possible consequences of intentional actions (Brass and Haggard, 2010).

In addition to these core areas, intention-related activity has been reported in many other brain regions with a widespread distribution as shown in **Figure 1**, which presents an overview of cortical activation foci reported by recent neuroimaging studies that explicitly aimed at investigating intentions in healthy subjects. Using these criteria, we identified 22 studies (Burgess et al., 2001; Lau et al., 2004, 2006; den Ouden et al., 2005; Cunnington et al., 2006; Simons et al., 2006; Eschen et al., 2007; Haynes et al., 2007; Mueller et al., 2007; Farrer et al., 2008; Soon et al., 2008; Krieghoff et al., 2009; Poppenk et al., 2010; Gilbert, 2011; Hashimoto et al., 2011; Okuda et al., 2011; Benoit et al., 2012; Gilbert et al., 2012; Hoffstaedter et al., 2012; Momennejad and Haynes, 2012; Rosenberg-Katz et al., 2012) reporting a total amount of 303 cortical and subcortical intention-related peaks.

We performed an activation likelihood estimate analysis (ALE; as described in Mutschler et al., 2009) of these studies to statistically detect brain regions with responses that occur red reproducibly. This revealed only two clusters with significant ALE ($p < 0.05$, FDR-corrected). Both of them were located in the SMA (assigned to Brodmann area 6 with maxima at MNI coordinates −2; 16; 54 and −4; 14; 50, and with respective probabilities of anatomical assignment of 40 and 50% (Eickhoff et al., 2005). There may be several reasons why only these clusters were significant. First, the number of studies satisfying our strict selection criteria was limited. Future meta-analyses based on larger samples may reveal additional foci of reproducible neural responses. Second, as argued in Brass and Haggard (2008, p. 319), the spread of neural activity seen in neuroscientific literature on intentions may be because "intentional action has been treated as a unitary concept within neuroscience, even though experimental studies may focus on any of a number of different aspects of intentional action." Meta-analyses distinguishing different types and aspects of human intention may be necessary to reveal more reproducible neural responses. Applying the same idea to the field of BMI research, in the following section we discuss to what extent it may be useful or even necessary to integrate different types as well as temporal and semantic aspects of intentions to develop safe and efficient real-life BMI applications.

CURRENT APPROACHES TO BRAIN-MACHINE INTERFACING

Brain-machine interfaces allow humans to control technical devices through direct recordings of brain activity. To this end, the device – either intracranial (brain-implanted) or extracranial (fixed on the person's skull) – measures the brain activity of an agent, interprets ("decodes") the agent's intended action, and sends the corresponding execution command to an artificial effector system, such as a computer cursor, a prosthetic limb, or a wheelchair (**Figure 2A**). First clinical trials have demonstrated the success of the BMI principle for restoration of movement (Hochberg et al., 2006, 2012) and communication (Birbaumer et al., 1999; Guenther et al., 2009) in paralyzed individuals.

Brain-machine interfacing approaches may be categorized by the type of brain signal used (single-neuron activity, neural-population signals, etc.) and the invasiveness of the recording technique (Waldert et al., 2009). To assess the potential importance of intention concepts for practical BMI purposes, we shall characterize BMIs according to (i) the type of neural activity "input

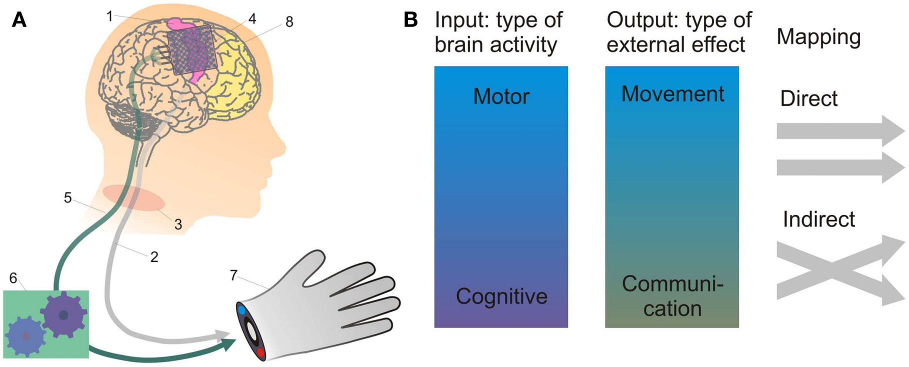

FIGURE 2 | The working principle and current approaches of BMI. In **(A)**, BMI-based control of a prosthetic hand by using recordings of brain activity is depicted. The projections from the motor cortex (1) constituting the cortico-spinal pathways (2) may be disrupted, e.g., by spinal-cord injury (3). Then, suitable electrodes (4) can be used to record persistent motor-cortical activity, which is transmitted by a technical connection (5; either wire-based or wireless) to a decoder (6) extracting control signals for an external actuator (7). If the primary motor cortex is destroyed, such as due to stroke, cognitive control signals may still be recorded from alternative areas such as the prefrontal cortex (8). As summarized in **(B)**, neural control signals may thus range from low-level motor signals, such as related to movement direction or velocity, to more high-level cognitive signals related to abstract action goals, subjective preferences, and intentions. Output signals may be used to restore movement (e.g., of an external actuator) or communication. The input-output mapping can be realized in a direct way, e.g., if right- and leftward movement-related neural activity controls the respective right- and leftward movements of an effector, the intention to grasp a cup is directly translated into the corresponding grasping action, or the intention to say the word "hello" is directly translated into speech. Indirect strategies would be, for example: using imagined leg vs. tongue movements to control right- vs. leftward movements of a robotic arm.

signal" used to decode information, (ii) the type of external output that is generated, and (iii) the kind of mapping between input and output (**Figure 2B**).

Regarding neural input, an important distinction can be made between BMIs using "low-level" motor signals, such as changes in neural activity related to movement direction or velocity, and BMIs utilizing "higher-level," cognitive signals. These may relate to subjective preferences or abstract action goals (Musallam et al., 2004). Between these extremes, there is a spectrum of more or less abstract/cognitive control signals that have been used, or are at least in principle usable, for BMI applications. Such signals can reflect that action plans are represented at different levels of abstraction in the brain (Bonini et al., 2011). Many current BMIs use low-level motor control signals recorded from primary or secondary motor areas (Hochberg et al., 2006, 2012; Moritz et al., 2008; Pistohl et al., 2012). BMIs based on this approach, however, still have much room for improvement in terms of decoding accuracy, especially in complex motor tasks. These and other challenges of present BMI technology are addressed further in a recent review by Schalk and Leuthardt (2011).

Cognitive signals may help to make BMI control more accurate. Based on the decoded abstract goals or intentions, intelligent autonomous external devices can perform lower-level computations, such as trajectories, that are necessary to achieve movement goals (Musallam et al., 2004). This approach may serve to lower bandwidth requirements for BMIs. Furthermore, if the brain structures that allow decoding movement-related signals (e.g., the primary motor cortex) are dysfunctional due to pathological processes, cognitive neural control signals, such as action goal- and intention-related activity from higher-order brain regions including premotor, posterior parietal, and PFC, may be used to substitute.

On the output side, the information decoded from either low- or high-level signals may be harnessed to generate movement (of a screen cursor, robotic arm, wheelchair, or even of a patient's own limbs via electrical stimulation of the muscles) or communication signals (as ringing an alarm bell, controlling a spelling device for writing, or synthesizing acoustic speech). Again, there are intermediate cases, such as if signals related to attempted right- and left-hand movement (a motor signal) were used to select a part of the alphabet or a letter in a spelling device (a communication output).

Different strategies may be used to map the input (brain) signal to the (externally-directed) output signal. We refer to those that aim to restore movement or speech functions with neural signals underlying the same function as "direct." For instance, a direct motor BMI would use brain signals related to left- vs. rightward movements to generate left- vs. rightward movements of an effector (Leuthardt et al., 2006; Milekovic et al., 2012). A direct speech BMI may use neural signals related to the respective phoneme (Blakely et al., 2008; Guenther et al., 2009; Pei et al., 2011), word (Kellis et al., 2010), semantic content (Wang et al., 2011), and context-dependent style (Derix et al., 2012) to generate matching speech output. Thus, neural activity related to the intended word "hello" would be decoded to spell "hello" in the BMI output (Kellis et al., 2010). In contrast, indirect approaches rely on neural input from tasks or modalities not directly related to output. For example, imagined leg vs. tongue movements may be used to control right- vs. leftward movements of a robotic arm. On this principle, Leuthardt et al. (2011) recently achieved BMI-based one-dimensional motor control using input signals related to production of overt and imagined phonemes. Indirect approaches have been widely used in non-invasive EEG-based BMI studies, since it is possible to select arbitrary tasks inducing highly distinctive global topographic EEG patterns, which can be robustly classified for BMI control.

The importance of intention-related brain signals and the potential role of intention concepts may vary depending on the BMI approach. A BMI based on low-level motor control signals may, at least to a certain extent, work without any such high-level information as intention-related signals. For approaches which do tap into intention-related processes in the brain, however, it may be useful or even necessary to take well-informed intention concepts into account, especially given a direct framework, i.e., if intentions are to be directly translated into the intended action.

Risks due to misinterpretation of neural control signals would be greatest for BMIs with effectors such as robotic arms or wheelchairs. In such applications, decoding of higher-order information with respect to the final goal of action as a whole may be a useful safety precaution, even if they primarily rely on low-level motor signals. In summary, intention concepts appear most relevant for *direct BMIs using cognitive neural control signals*, with both movement and speech output, but intention-related information may also constitute an auxiliary information channel for other types of BMIs.

PHILOSOPHICAL TAXONOMIES OF INTENTIONS AND THEIR RELEVANCE TO BMI

A properly designed intention-based BMI device should be able to clearly distinguish between different types of intentions. For instance, a patient using a BMI to steer a wheelchair may intend to turn right in a few seconds, right now, or next Wednesday, and the wheelchair must be sensitive to this temporal difference. At first glance, this distinction seems fairly trivial. Yet the question arises: How many different kinds of intentions can be identified by their temporal characteristics? And how are different types of intentions organized in time, that is, what are their individual dynamics, mutual transitions, and interactions? A number of intention theories (Searle, 1983; Brand, 1984; Bratman, 1987; Pacherie, 2006) have addressed the issue of timing and elaborated on various aspects of future- vs. present-oriented intentions.

Among other questions related to rational action, Bratman (1987) addressed differences between future- and present-directed intentions. According to his conceptual framework, future-directed intentions are formed prior to action and represent the product of deliberation whether or not to act in a certain way. An example is a future-directed intention to leave for Boston in April that has been formed in January (Bratman, 1987). In contrast, present-directed intentions inherit plans from future-directed intentions, and implement them in a current situation of action. Thus, if one has a future-directed intention to go to Boston in April, a present-directed intention may be to take a particular route that day and turn while driving to the airport, whereby the agent advances to complete his global plan. According to Bratman (1987), these two types of intentions are formed based on one's desires and beliefs as to whether the action in question is in some way beneficial and necessary to the conscious agent.

In a more recent philosophical paper, Pacherie (2006) adopted this terminology, referring to present- and future-directed intentions as P- and F-intentions, respectively. We will use these abbreviations from here on, also in cases where we do not refer to the specific theories by Pacherie (e.g., parts of **Figure 3**). In addition to these two types of intentions, Pacherie (2006) proposed a third category, the so-called motor-, or M-intentions, which inherit goals from present-directed intentions and initiate a motor program satisfying the spatial and temporal demands for action realization (Pacherie, 2006). One main reason for introducing this additional category was the consideration that, whereas both P- and F-intentions are subject to strong rationality constraints (Bratman, 1987), not all voluntary actions require deliberation. Examples are such automated, routine actions as a smoker reaching for a pack of cigarettes and realizing that she is doing it already in the process of reaching (Pacherie, 2006), or a person who unlocks his office door by mere habit of doing so every morning (Mele, 2007).

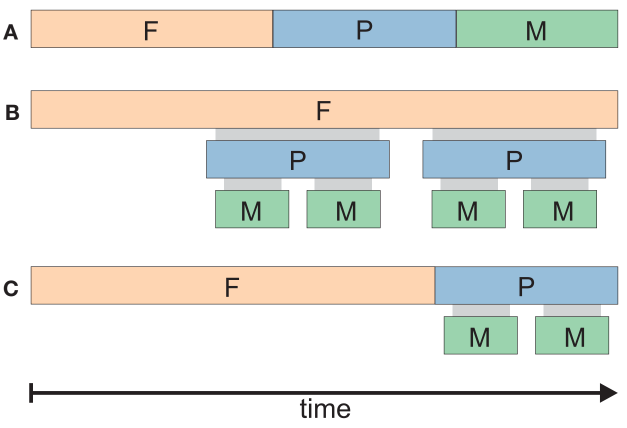

FIGURE 3 | (A) Schematic representation of sequential **(A)** vs. hierarchical **(B)** vs. **(C)** mixed models of intention organization in time. Three different scenarios are depicted for the case of a threefold intention concept, roughly corresponding to the F-, P-, and M- intentions proposed by Pacherie (2006). The durations of F-, P-, and M-intentions are depicted by colored bars. A mixture of hierarchical and sequential relations is shown in **(C)**. Which of these different scenarios is true in a given situation would have important consequences for attempts to decode intentions; for example, the detection of an F-intention would rule out the simultaneous presence of the corresponding P- and M-intentions in the purely sequential **(A)** but not in the hierarchical **(B)** model. Note that Pacherie's concept favors **(C)**, particularly in her recent work (Pacherie, 2008).

The F-, P-, and M-intentions have distinct functional roles. Based on Bratman's account of F-intentions (1987), Pacherie (2006, p. 3) assumes that F-intentions serve as "terminators of practical reasoning about ends, prompters of practical reasoning about means and plans, and intra- and interpersonal coordinators." The conscious P-intentions ensure "higher-level guidance and monitoring," whereas "lower-level guiding and monitoring functions should properly be assigned to M-intentions" (Pacherie, 2006, p. 5). But what is the exact *temporal* organization of intentions – is it hierarchical, or do intentions unfold sequentially?

A sequential model of intentions (illustrated in **Figure 3A**) would assume that the F-, P-, and M- intentions precede each other, and one type of intention stops once it has passed its goal onto the next type which directly follows. A hierarchical model (**Figure 3B**), though, would assume that all three types of intentions overlap in time and govern one another in synchrony.

Concerning F-intentions, Pacherie (2006, p. 4) states that they are "in principle detachable from the agent's current situation and [are] indeed commonly detached from it," and that "insofar as they are temporally separated from the action, [F-intentions] make no direct contribution to the experience of acting" (Pacherie, 2006, p. 14). In contrast, "P-intentions and M-intentions are both simultaneous with the action that they guide" (Pacherie, 2006, p. 14). While these views were still at least to some extent consistent both with models B and C (**Figure 3**), in her more recent work, Pacherie (2008) sides more clearly in favor of a mixed sequential-hierarchical model (**Figure 3C**), in which F-intentions occur before P- and M-, and the latter two types of intentions take place simultaneously.

The temporal model of intentions we inferred based on work by Pacherie (2006); **Figure 3C**), as well as the other two models (A and B in **Figure 3**), may be a useful reference to interpret observations of action- and intention-related neural activity at different temporal scales. Furthermore, these temporal models of intentions may entail important consequences for attempts to decode intentions. For example, the detection of an F-intention would rule out the simultaneous presence of the corresponding P- and M-intentions in the purely sequential but not in the hierarchical model. The decoding problems for BMI devices will differ accordingly.

If our experience of acting is directly governed by P- and M-, but not by F-intentions (Pacherie, 2006), it seems plausible that the neural correlates of F-intentions could considerably differ from that of both P- and M-intentions. Future-oriented intentions in neuroimaging studies have been mostly investigated in the context of prospective memory (Burgess et al., 2001; den Ouden et al., 2005; Simons et al., 2006; Eschen et al., 2007; Poppenk et al., 2010; Gilbert, 2011; Hashimoto et al., 2011; Okuda et al., 2011; Benoit et al., 2012; Gilbert et al., 2012; Momennejad and Haynes, 2012), and action intentions in a current situation of action have also been the focus of recent research (Cunnington et al., 2006; Gallivan et al., 2011a,b), revealing partially overlapping neural effects. A contrastive investigation into these two paradigmatic frameworks in the literature may be of interest in future studies including meta-analyses, and can be expected to reveal a topographically differential distribution of neural effects related to future- vs. present-directed intentions.

The idea that F-intentions are indeed formed prior to P-intentions, and do not directly contribute to the experience of action (Pacherie, 2006), may have important implications in the context of BMI. As the term implies, F-intentions deal with prospective plans (Bratman, 1987; Pacherie, 2006), so they may be used as possible coordinators for the fine-tuning of an intention decoder. Knowledge of the global plan of an intention is crucial for understanding action orientation as a whole (Bara et al., 2011). F-intentions could thus serve to improve the accuracy of a BMI, and ensure goal-oriented guidance of BMI-mediated action.

To differentiate the consciously experienced P-intentions from M-intentions, which are not subject to rationality constraints (Pacherie, 2006), is also essential for safe and efficient BMI-based restoration of motor functions. According to Pacherie (2006), p. 9), when M-intentions occur without P-intentions, they initiate "a competition among motor programs, with the program showing the strongest activation being triggered," such as in the above-mentioned example of reaching for a pack of cigarettes (Pacherie, 2006), possibly even when smoking is prohibited. BMI-based realization of M-intentions that do not inherit their goal from a P-intention may be, in some cases, dangerous, and in conflict with a higher-level "no-intention" (see below). On the other hand, unconscious intentions may play an important role in performance of automatic actions when no time for deliberation is available, e.g., when driving a car. Thus, there seems to be no general solution as to which cases of "isolated" M-intentions should be executed. Future research can be expected to shed more light on this issue.

Philosophical accounts may provide further theoretical ground for BMI research in their distinction between intentions to perform and intentions not to perform an action (Harman, 1986; Bratman, 1987; Setiya, 2011). For the sake of brevity, we suggest the terms *yes*- vs. *no-intentions.* Confusing these phenomena would severely compromise the safety of a BMI device. Intentional inhibition, which may be considered as a no-intention, has been investigated in human neuroscience (Brass and Haggard, 2007, 2008; Kühn and Brass, 2009; Kühn et al., 2009), identifying responses in the dorsal fronto-medial cortex distinct from areas implicated in what we refer to as yes-intentions (Brass and Haggard, 2007; Kühn et al., 2009). These findings may provide valuable information for emergent BMI technology, particularly to accurately interpret intended action vs. intended restraint.

Another distinction relevant to BMI is that between *direct* and *oblique* (i.e., indirect) intentions, as proposed by the English philosopher Jeremy Bentham in the late eighteenth century (1781, repr. 2000). Bentham (1781, p. 70) explains this distinction through a discrete relation between will, actions and consequences:

> "A consequence [...] may be said to be directly or lineally intentional, when the prospect of producing it constituted one of the links in the chain of causes by which the person was determined to do the act. It may be said to be obliquely or collaterally intentional, when although the consequence was in contemplation, and appeared likely to ensue in case of the acts being performed, yet the prospect of producing such consequence did not constitute a link in the aforesaid chain."

Bentham (1781, p. 71) exemplifies his account of direct and oblique intentions departing from a historical case of William II, king of England being deadly wounded by the nobleman Walter Tyrrel during a hunt. The circumstances of this incident remained unclear. According to Bentham, there are several possible ways to evaluate the intentionality of Sir Tyrrel's actions. One imaginable scenario is that the king is riding close to a stag, and Sir Tyrrel shoots his arrow with the aim to kill the stag; he is convinced that the shot is not dangerous to the king. The killing in this case occurs by accident and Bentham classifies it as unintentional. A second possibility is that Sir Tyrrel aims to kill the stag and shoots at it, although he is aware that the shot is as likely to kill the king as the stag. If Sir Tyrrel's shot kills the king in this scenario, his actions can be regarded as obliquely intentional. A third possibility is that Sir Tyrrel hates the king and shoots with no other aim than to kill him. In this latter case, Sir Tyrrel's actions classify as directly intentional. Thus, direct and oblique intentions differ as to whether the outcome of the action is actively sought-after (direct intentions) or a foreseeable "side effect" (oblique intentions).

Initially, this classification was developed to evaluate the degree of responsibility for harmful actions in the legal context. However, we believe that the distinction between direct and oblique intentions may be also of relevance to the emerging field of BMI. Imagine a person with a BMI-controlled prosthetic arm is having breakfast, and moves to reach a piece of bread, just behind his cup of coffee. The person aims to reach the bread (direct intention) and not to touch or topple the cup of coffee, although he understands that these consequences may occur (oblique intention). It is important that a BMI relying on inference of intentions does not confuse direct with oblique intentions, and gives priority to the execution of the former, to perform its user's effective wishes. To our knowledge, this direct vs. oblique distinction has not been investigated in cognitive neuroscience, and it is currently unclear which neural substrates support these different kinds of intentions.

Another important secernment is between what we call *mere urges* and *action intentions*. A review of the existing literature distinguishing intentions, urges, and desires, however, is beyond the scope of the present article (for literature on these distinctions, see Johnston, 2001; Mele, 2007). An urge may be phenomenologically described as a strong impulse toward an action. Urges are typically stimulus-evoked, such as an urge to scratch evoked by an itch or an urge to cry by a sad situation or thought. Urges may be delineated from desires in that desires have an evaluative element, i.e., the object of the desire is "desirable" and "good" in some way, while urges lack this (Scanlon, 1998, 2008). Within neuroscience, the neural underpinnings of urges have been, until now, most extensively investigated in the specific context of drug craving (Maas et al., 1998; Childress et al., 1999) and in electrical stimulation studies (Fried et al., 1991).

If the driving force of an urge becomes overwhelming, it may result in an "urged action" – even against one's intentions. However, it is a fundamentally important aspect of human behavior that urges can be controlled, and blocked if necessary. Here, we refer to an urge without any associated intention to perform an action as a *mere urge* – in contrast to an urge toward an action that is actually intended (following a similar idea as Pockett and Miller (2007), who distinguish a mere urge from an actual decision).

A BMI should likewise distinguish mere urges from action intentions. The relevance of this distinction becomes clear from the examples that follow. Imagine that a person with a BMI-controlled bionic arm becomes as angry at a rude conversation partner as to experience aggressive urges, such as to punch him for the offense. Punching the offender, however, does not correspond to the person's actual intentions. In this and similar cases, it is vitally important that the BMI device does not translate the mere urge into motor performance.

Another likely situation is that a person with a BMI-controlled arm is bitten by a mosquito and experiences an urge to scratch the bite. The person is aware that to scratch may further hurt the skin and make the itch even worse, so he decides to refrain from scratching. To prevent the execution of such unintended, and potentially dangerous movements, it will be necessary for the BMI to keep mere urges and intentions apart.

Whether the action is other- or agent-directed, an important question regarding misinterpretation of mere urges and intentions by BMI technology is: If someone is hurt in such a scenario, is the user responsible, or the manufacturer of the device? It seems plausible that a mere urge as defined above is not morally significant, and that a BMI application must be able to distinguish it from an action intention.

CONCLUSIONS AND OUTLOOK

Intention is often treated as a unitary concept in neuroscientific research (Brass and Haggard, 2008). Philosophy may help to discern types of intentions, and give a more differential account. Studies of intentions in neuroscience that have explicitly used philosophically-informed concepts are however rare at present. As a notable exception, Bara et al. (2011) investigated, using fMRI, several types of future intentions, namely, private, prospective social, and communicative intentions, incorporating classifications proposed by Searle (1983), Bratman (1987), Pacherie and Haggard (2010), as well as by Tomasello (2008). Bara et al. (2011) found that all of these intention types were associated with activation in the right temporo-parietal junction and the precuneus; the activity in the left temporo-parietal junction and the medial PFC, however, was specifically observed in relation to prospective social intentions.

In addition to these first steps toward incorporating intention concepts from philosophy into basic neuroscience (Bara et al., 2011), philosophical accounts may contribute to the development of BMIs based on cognitive control signals, as has been argued above. Examples in the previous section illustrate how a better understanding of temporal properties of intentions is important for safe and efficient BMI performance. Together, the reviewed literature strongly suggests that it would be particularly important to study the neural basis of intentions (i) at different temporal scales (Searle, 1983; Brand, 1984; Bratman, 1987; Pacherie, 2006) and (ii) taking into account the issue of hierarchical vs. sequential organization of intentions in time (Pacherie, 2006, 2008).

Further research to improve BMI may also benefit from understanding qualitative differences between various types of intentions that have been proposed in philosophy, including yes- vs. no-intentions (Harman, 1986; Bratman, 1987; Setiya, 2011), and direct vs. oblique intentions (Bentham, 1781). Only a few neuroscientific studies so far have investigated intentional inhibition of actions (Brass and Haggard, 2007, 2008; Kühn and Brass, 2009), which is apparently analogous to a no-intention, and of yet we are not aware of any study dedicated to the direct vs. oblique distinction.

The temporal and semantic components of intention seem to be strongly related, as the degree of content abstraction is generally higher in future- compared to present-directed intentions (Searle, 1983; Mele, 1989; Pacherie, 2006). Nevertheless, it is also imaginable that intentions directed at the same time in the future may still vary with respect to their level of abstraction. For instance, one may intend to go on a holiday next summer, or one may intend to go on a holiday to the same nice hotel in Ronda next summer. Conversely, intentions with different temporal targets may exhibit a higher degree of content similarity relative to other intentions with the same time to action execution. Content abstraction may hence be important to include into further empirical research on intention dynamics as an additional, at least partially independent factor. Insights from such investigations may be useful to determine the exact onset of an intended action in BMI-based movement restoration.

Alongside philosophical intention taxonomies, conceptual input from psychology and cognitive science may be of value. For instance, a conceptual framework incorporating the "what," "when," and "whether" components of intentional action has been proposed (Brass and Haggard, 2008) and neuroscientifically investigated (Brass and Haggard, 2007; Mueller et al., 2007; Krieghoff et al., 2009; Kühn and Brass, 2009; Kühn et al., 2009; Obhi et al., 2009; Hoffstaedter et al., 2012; Momennejad and Haynes, 2012). Further hallmark questions to be addressed in future interdisciplinary research are: (i) Can philosophical intention taxonomies be used as direct input for BMI studies, or do they first need to be operationalized to be applicable to neuronal data? (ii) What are the correspondences and differences between philosophical and psychological concepts of human intent? In addressing these questions, not only may neuroscience and BMI research benefit from cooperation with philosophy, but also vice versa: insights into the biological plausibility of different aspects and types of intentions proposed in the philosophy of mind may provide valuable empirical feedback, thereby closing the loop from neuroscience to philosophy.

ACKNOWLEDGMENTS

This work was supported by the Barbara-Wengeler Foundation, the German Federal Ministry of Education and Research (BMBF) grant 01GQ0420 to BCCN Freiburg, BMBF GoBio grant 0313891, and the European Platform for Life Sciences, Mind Sciences, and the Humanities. We would like to acknowledge the insightful discussions with Julian Kiverstein and thank Elisabeth Pacherie for helpful correspondence. The article processing fee was covered by the German Research Foundation (DFG) and the Albert Ludwigs University Freiburg in terms of the funding program "Open Access Publishing".

REFERENCES

Andersen, R. A., and Buneo, C. A. (2002). Intentional maps in posterior parietal cortex. *Annu. Rev. Neurosci.* 25, 189–220.

Andersen, R. A., and Cui, H. (2009). Intention, action planning, and decision making in parietal-frontal circuits. *Neuron* 63, 568–583.

Anscombe, G. E. M. (1963). *Intention*, 2nd Edn. Cambridge: Harvard University Press.

Assal, F., Schwartz, S., and Vuilleumier, P. (2007). Moving with or without will: functional neural correlates of alien hand syndrome. *Ann. Neurol.* 62, 301–306.

Ball, T., Schreiber, A., Feige, B., Wagner, M., Lücking, C. H., and Kristeva-Feige, R. (1999). The role of higher-order motor areas in voluntary movement as revealed by high-resolution EEG and fMRI. *Neuroimage* 10, 682–694.

Bara, B. G., Ciaramidaro, A., Walter, H., and Adenzato, M. (2011). Intentional minds: a philosophical analysis of intention tested through fMRI experiments involving people with schizophrenia, people with autism, and healthy individuals. *Front. Hum. Neurosci.* 5:7. doi:10.3389/fnhum.2011.00007

Benoit, R. G., Gilbert, S. J., Frith, C. D., and Burgess, P. W. (2012). Rostral prefrontal cortex and the focus of attention in prospective memory. *Cereb. Cortex* 22, 1876–1886.

Bentham, J. (1781). *An Introduction to the Principles of Morals and Legislation.* Kitchener, ON: Batoche Books. Reprint 2000.

Birbaumer, N., Ghanayim, N., Hinterberger, T., Iversen, I., Kotchoubey, B., Kübler, A., et al. (1999). A spelling device for the paralysed. *Nature* 398, 297–298.

Blakely, T., Miller, K. J., Rao, R. P. N., Holmes, M. D., and Ojemann, J. G. (2008). Localization and classification of phonemes using high spatial resolution electrocorticography (ECoG) grids. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* 2008, 4964–4967.

Bonini, L., Serventi, F. U., Simone, L., Rozzi, S., Ferrari, P. F., and Fogassi, L. (2011). Grasping neurons of monkey parietal and premotor cortices encode action goals at distinct levels of abstraction during complex action sequences. *J. Neurosci.* 31, 5876–5886.

Brand, M. (1984). *Intending and Acting: Toward a Naturalized Action Theory.* Cambridge: MIT Press.

Brass, M., and Haggard, P. (2007). To do or not to do: the neural signature of self-control. *J. Neurosci.* 27, 9141–9145.

Brass, M., and Haggard, P. (2008). The what, when, whether model of intentional action. *Neuroscientist* 14, 319–325.

Brass, M., and Haggard, P. (2010). The hidden side of intentional action: the role of the anterior insular cortex. *Brain Struct. Funct.* 214, 603–610.

Bratman, M. E. (1987). *Intention, Plans and Practical Reason.* Cambridge: Harvard University Press.

Bratman, M. E. (2009). Setiya on intention, rationality and reasons. *Analysis* 69, 510–521.

Burgess, P. W., Quayle, A., and Frith, C. D. (2001). Brain regions involved in prospective memory as determined by positron emission tomography. *Neuropsychologia* 39, 545–555.

Childress, A. R., Mozley, P. D., McElgin, W., Fitzgerald, J., Reivich, M., and O'Brien, C. P. (1999). Limbic activation during cue-induced cocaine craving. *Am. J. Psychiatry* 156, 11–18.

Cui, H., and Andersen, R. A. (2007). Posterior parietal cortex encodes autonomously selected motor plans. *Neuron* 56, 552–559.

Cunnington, R., Windischberger, C., Robinson, S., and Moser, E. (2006). The selection of intended actions and the observation of others' actions: a time-resolved fMRI study. *Neuroimage* 29, 1294–1302.

Davidson, D. (1963). Actions, reasons, and causes. *J. Philos.* 60, 685–700.

den Ouden, H. E. M., Frith, U., Frith, C., and Blakemore, S.-J. (2005). Thinking about intentions. *Neuroimage* 28, 787–796.

Derix, J., Iljina, O., Schulze-Bonhage, A., Aertsen, A., and Ball, T. (2012). "Doctor" or "darling"? Decoding the communication partner from ECoG of the anterior temporal lobe during non-experimental, real-life social interaction. *Front. Hum. Neurosci.* 6:251. doi: 10.3389/fnhum.2012.00251

Desmurget, M., Reilly, K. T., Richard, N., Szathmari, A., Mottolese, C., and Sirigu, A. (2009). Movement intention after parietal cortex stimulation in humans. *Science* 324, 811–813.

Eccles, J. C. (1982). Animal consciousness and human self-consciousness. *Experientia* 38, 1384–1391.

Eddy, C. M., Mitchell, I. J., Beck, S. R., Cavanna, A. E., and Rickards, H. E. (2010). Altered attribution of intention in Tourette's syndrome. *J. Neuropsychiatry Clin. Neurosci.* 22, 348–351.

Edwards, M. J., Moretto, G., Schwingenschuh, P., Katschnig, P., Bhatia, K. P., and Haggard, P. (2011). Abnormal sense of intention preceding voluntary movement in patients with psychogenic tremor. *Neuropsychologia* 49, 2791–2793.

Eickhoff, S. B., Stephan, K. E., Mohlberg, H., Grefkes, C., Fink, G. R., Amunts, K., et al. (2005). A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *Neuroimage* 25, 1325–1335.

Eschen, A., Freeman, J., Dietrich, T., Martin, M., Ellis, J., Martin, E., et al. (2007). Motor brain regions are involved in the encoding of delayed intentions: a fMRI study. *Int. J. Psychophysiol.* 64, 259–268.

Farrer, C., Frey, S. H., Horn, J. D. V., Tunik, E., Turk, D., Inati, S., et al. (2008). The angular gyrus computes action awareness representations. *Cereb. Cortex* 18, 254–261.

Fazel-Rezai, R., Allison, B. Z., Guger, C., Sellers, E. W., Kleih, S. C., and Kübler, A. (2012). P300 brain computer interface: current challenges and emerging trends. *Front. Neuroeng.* 5:14. doi:10.3389/fneng.2012.00014

Fried, I., Katz, A., McCarthy, G., Sass, K. J., Williamson, P., Spencer, S. S., et al. (1991). Functional organization of human supplementary motor cortex studied by electrical stimulation. *J. Neurosci.* 11, 3656–3666.

Fried, I., Mukamel, R., and Kreiman, G. (2011). Internally generated preactivation of single neurons in human medial frontal cortex predicts volition. *Neuron* 69, 548–562.

Gallivan, J. P., McLean, D. A., Smith, F. W., and Culham, J. C. (2011a). Decoding effector-dependent and effector-independent movement intentions from human parieto-frontal brain activity. *J. Neurosci.* 31, 17149–17168.

Gallivan, J. P., McLean, D. A., Valyear, K. F., Pettypiece, C. E., and Culham, J. C. (2011b). Decoding action intentions from preparatory brain activity in human parieto-frontal networks. *J. Neurosci.* 31, 9599–9610.

Gilbert, S. J. (2011). Decoding the content of delayed intentions. *J. Neurosci.* 31, 2888–2894.

Gilbert, S. J., Armbruster, D. J. N., and Panagiotidi, M. (2012). Similarity between brain activity at encoding and retrieval predicts successful realization of delayed intentions. *J. Cogn. Neurosci.* 24, 93–105.

Goldman, A. I. (1970). *A Theory of Human Action.* New York: Prentice-Hall.

Guenther, F. H., Brumberg, J. S., Wright, E. J., Nieto-Castanon, A., Tourville, J. A., Panko, M., et al. (2009). A wireless brain-machine interface for real-time speech synthesis. *PLoS ONE* 4, e8218. doi:10.1371/journal.pone.0008218

Haggard, P. (2005). Conscious intention and motor cognition. *Trends Cogn. Sci. (Regul. Ed.)* 9, 290–295.

Haggard, P. (2008). Human volition: towards a neuroscience of will. *Nat. Rev. Neurosci.* 9, 934–946.

Haggard, P., and Clark, S. (2003). Intentional action: conscious experience and neural prediction. *Conscious. Cogn.* 12, 695–707.

Haggard, P., and Eimer, M. (1999). On the relation between brain potentials and the awareness of voluntary movements. *Exp. Brain Res.* 126, 128–133.

Harman, G. (1986). *Change in View: Principles of Reasoning.* Cambridge: MIT Press.

Hashimoto, T., Umeda, S., and Kojima, S. (2011). Neural substrates of implicit cueing effect on prospective memory. *Neuroimage* 54, 645–652.

Haynes, J.-D., Sakai, K., Rees, G., Gilbert, S., Frith, C., and Passingham, R. E. (2007). Reading hidden intentions in the human brain. *Curr. Biol.* 17, 323–328.

Hochberg, L. R., Bacher, D., Jarosiewicz, B., Masse, N. Y., Simeral, J. D., Vogel, J., et al. (2012). Reach and grasp by people with tetraplegia using a neurally controlled robotic arm. *Nature* 485, 372–375.

Hochberg, L. R., Serruya, M. D., Friehs, G. M., Mukand, J. A., Saleh, M., Caplan, A. H., et al. (2006). Neuronal ensemble control of prosthetic devices by a human with tetraplegia. *Nature* 442, 164–171.

Hoffstaedter, F., Grefkes, C., Zilles, K., and Eickhoff, S. B. (2012). The "what" and "when" of self-initiated movements. *Cereb. Cortex* doi:10.1093/cercor/bhr391. [Epub ahead of print].

Jahanshahi, M., Jenkins, I. H., Brown, R. G., Marsden, C. D., Passingham, R. E., and Brooks, D. J. (1995). Self-initiated versus externally triggered movements. I. An investigation using measurement of regional cerebral blood flow with PET and movement-related potentials in normal and Parkinson's disease subjects. *Brain* 118(Pt 4), 913–933.

Jenkins, I. H., Jahanshahi, M., Jueptner, M., Passingham, R. E., and Brooks, D. J. (2000). Self-initiated versus externally triggered movements. *Brain* 123, 1216–1228.

Johnston, M. (2001). The Authority of affect*. *Philos. Phenomenol. Res.* 63, 181–214.

Kant, I. (1785). *Grundlegung zur Metaphysik der Sitten.* Riga: Johann Friedrich Hartknoch.

Keller, I., and Heckhausen, H. (1990). Readiness potentials preceding spontaneous motor acts: voluntary vs. involuntary control. *Electroencephalogr. Clin. Neurophysiol.* 76, 351–361.

Kellis, S., Miller, K., Thomson, K., Brown, R., House, P., and Greger, B. (2010). Classification of spoken words using surface local field potentials. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* 2010, 3827–3830.

Kiverstein, J. (2006). "An enactive theory of phenomenal intentionality," in *Content, Consciousness and Perception: Essays in Contemporary Philosophy of Mind*, eds. E. Di Nucci and C. McHugh (London: Cambridge Scholars Press), 190–211.

Krieghoff, V., Brass, M., Prinz, W., and Waszak, F. (2009). Dissociating what and when of intentional actions. *Front. Hum. Neurosci.* 3:3. doi:10.3389/neuro.09.003.2009

Kühn, S., and Brass, M. (2009). When doing nothing is an option: the neural correlates of deciding whether to act or not. *Neuroimage* 46, 1187–1193.

Kühn, S., Haggard, P., and Brass, M. (2009). Intentional inhibition: how the "veto-area" exerts control. *Hum. Brain Mapp.* 30, 2834–2843.

Lau, H., Rogers, R. D., and Passingham, R. E. (2006). Dissociating response selection and conflict in the medial frontal surface. *Neuroimage* 29, 446–451.

Lau, H. C., Rogers, R. D., Haggard, P., and Passingham, R. E. (2004). Attention to intention. *Science* 303, 1208–1210.

Leuthardt, E. C., Gaona, C., Sharma, M., Szrama, N., Roland, J., Freudenberg, Z., et al. (2011). Using the electrocorticographic speech network to control a brain–computer interface in humans. *J. Neural Eng.* 8, 036004.

Leuthardt, E. C., Schalk, G., Moran, D., and Ojemann, J. G. (2006). The emerging world of motor neuroprosthetics: a neurosurgical perspective. *Neurosurgery* 59, 1–14.

Libet, B., Gleason, C. A., Wright, E. W., and Pearl, D. K. (1983). Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential). The unconscious initiation of a freely voluntary act. *Brain* 106(Pt 3), 623–642.

Maas, L. C., Lukas, S. E., Kaufman, M. J., Weiss, R. D., Daniels, S. L., Rogers, V. W., et al. (1998). Functional magnetic resonance imaging of human brain activation during cue-induced cocaine craving. *Am. J. Psychiatry* 155, 124–126.

Mele, A. (2007). "Free will: action theory meets neuroscience," in *Intentionality, Deliberation and Autonomy: The Action-Theoretic Basis of Practical Philosophy*, ed. C. Lumer (Hampshire: Ashgate Publishing, Ltd.), 257–272.

Mele, A. R. (1989). Intention, belief, and intentional action. *Am. Philos. Q.* 26, 19–30.

Mele, A. R. (2008). Proximal intentions, intention-reports, and vetoing. *Philos. Psychol.* 21, 1–14.

Milekovic, T., Fischer, J., Pistohl, T., Ruescher, J., Schulze-Bonhage, A., Aertsen, A., et al. (2012). An online brain-machine interface using decoding of movement direction from the human electrocorticogram. *J. Neural Eng.* 9, 046003.

Momennejad, I., and Haynes, J.-D. (2012). Human anterior prefrontal cortex encodes the "what" and "when" of future intentions. *Neuroimage* 61, 139–148.

Moritz, C. T., Perlmutter, S. I., and Fetz, E. E. (2008). Direct control of paralysed muscles by cortical neurons. *Nature* 456, 639–642.

Mueller, V. A., Brass, M., Waszak, F., and Prinz, W. (2007). The role of the preSMA and the rostral cingulate zone in internally selected actions. *Neuroimage* 37, 1354–1361.

Musallam, S., Corneil, B. D., Greger, B., Scherberger, H., and Andersen, R. A. (2004). Cognitive control signals for neural prosthetics. *Science* 305, 258–262.

Mutschler, I., Wieckhorst, B., Kowalevski, S., Derix, J., Wentlandt, J., Schulze-Bonhage, A., et al. (2009). Functional organization of the human anterior insular cortex. *Neurosci. Lett.* 457, 66–70.

Nachev, P., Wydell, H., O'Neill, K., Husain, M., and Kennard, C. (2007). The role of the pre-supplementary motor area in the control of action. *Neuroimage* 36(Suppl. 2), T155–T163.

Obhi, S. S., Matkovich, S., and Chen, R. (2009). Changing the "when" and "what" of intended actions. *J. Neurophysiol.* 102, 2755–2762.

Okuda, J., Gilbert, S. J., Burgess, P. W., Frith, C. D., and Simons, J. S. (2011). Looking to the future: automatic regulation of attention between current performance and future plans. *Neuropsychologia* 49, 2258–2271.

Pacherie, E. (2002). "Intention," in *Encyclopedia of Cognitive Science*, ed. L. Nadel (London: MacMillan), 599–604.

Pacherie, E. (2006). "Toward a dynamic theory of intentions," in *Does Consciousness Cause Behavior? An Investigation of the Nature of Volition*, eds. S. Pockett, W. P. Banks, and S. Gallagher (Cambridge: MIT Press), 145–167.

Pacherie, E. (2007). The anarchic hand syndrome and utilization behavior: a window onto agentive self-awareness. *Funct. Neurol.* 22, 211–217.

Pacherie, E. (2008). The phenomenology of action: a conceptual framework. *Cognition* 107, 179–217.

Pacherie, E. (2011). Nonconceptual representations for action and the limits of intentional control. *Soc. Psychol.* 42, 67–73.

Pacherie, E., and Haggard, P. (2010). "What are intentions?," in *Conscious Will and Responsibility. A tribute to Benjamin Libet*, eds. L. Nadel and W. Sinnott-Armstrong (Oxford: Oxford University Press), 70–84.

Passingham, R. E., Bengtsson, S. L., and Lau, H. C. (2010). Medial frontal cortex: from self-generated action to reflection on one's own performance. *Trends Cogn. Sci. (Regul. Ed.)* 14, 16–21.

Pei, X., Barbour, D. L., Leuthardt, E. C., and Schalk, G. (2011). Decoding vowels and consonants in spoken and imagined words using electrocorticographic signals in humans. *J. Neural Eng.* 8, 046028.

Pesaran, B., Nelson, M. J., and Andersen, R. A. (2006). Dorsal premotor neurons encode the relative position of the hand, eye, and goal during reach planning. *Neuron* 51, 125–134.

Pistohl, T., Schulze-Bonhage, A., Aertsen, A., Mehring, C., and Ball, T. (2012). Decoding natural grasp types from human ECoG. *Neuroimage* 59, 248–260.

Pockett, S., and Miller, A. (2007). The rotating spot method of timing subjective events. *Conscious. Cogn.* 16, 241–254.

Poppenk, J., Moscovitch, M., McIntosh, A. R., Ozcelik, E., and Craik, F. I. M. (2010). Encoding the future: successful processing of intentions engages predictive brain networks. *Neuroimage* 49, 905–913.

Quian Quiroga, R., Snyder, L. H., Batista, A. P., Cui, H., and Andersen, R. A. (2006). Movement intention is better predicted than attention in the posterior parietal cortex. *J. Neurosci.* 26, 3615–3620.

Rao, A. S., and Georgeff, M. P. (1991). "Modeling rational agents within a BDI-architecture," in *Proceedings of the 2nd International Conference on Principles of Knowledge Representation and Reasoning*, San Mateo, 473–484.

Rosenberg-Katz, K., Jamshy, S., Singer, N., Podlipsky, I., Kipervasser, S., Andelman, F., et al. (2012). Enhanced functional synchronization of medial and lateral PFC underlies internally-guided action planning. *Front. Hum. Neurosci.* 6:79. doi:10.3389/fnhum.2012.00079

Rushworth, M. F. S. (2008). Intention, choice, and the medial frontal cortex. *Ann. N. Y. Acad. Sci.* 1124, 181–207.

Scanlon, T. (1998). *What We Owe to Each Other*. Cambridge: Harvard University Press.

Scanlon, T. (2008). *Moral Dimensions: Permissibility, Meaning, Blame.* Cambridge: Harvard University Press.

Schalk, G., and Leuthardt, E. C. (2011). Brain-computer interfaces using electrocorticographic signals. *IEEE Rev. Biomed. Eng.* 4, 140–154.

Schleim, S. (2008). The Brain on the stand? interactions between neuroscience, ethics and law. *J. Int. Biotechnol. Law* 5, 131–134.

Searle, J. R. (1983). *Intentionality, An Essay in the Philosophy of Mind.* New York: Cambridge University Press.

Setiya, K. (2007). *Reasons Without Rationalism.* Princeton, NJ: Princeton University Press.

Setiya, K. (2011). "Intention," in *The Stanford Encyclopedia of Philosophy*, ed. E. N. Zalta. Available at: http://plato.stanford.edu/archives/spr2011/entries/intention/

Simons, J. S., Schölvinck, M. L., Gilbert, S. J., Frith, C. D., and Burgess, P. W. (2006). Differential components of prospective memory? Evidence from fMRI. *Neuropsychologia* 44, 1388–1397.

Sirigu, A., Daprati, E., Ciancia, S., Giraux, P., Nighoghossian, N., Posada, A., et al. (2004). Altered awareness of voluntary action after damage to the parietal cortex. *Nat. Neurosci.* 7, 80–84.

Snyder, L. H., Batista, A. P., and Andersen, R. A. (1997). Coding of intention in the posterior parietal cortex. *Nature* 386, 167–170.

Soon, C. S., Brass, M., Heinze, H.-J., and Haynes, J.-D. (2008). Unconscious determinants of free decisions in the human brain. *Nat. Neurosci.* 11, 543–545.

Tomasello, M. (2008). *Origins of Human Communication.* Cambridge: MIT Press.

Uretzky, S., and Gilboa, A. (2010). Knowing your lines but missing your cue: rostral prefrontal lesions impair prospective memory cue detection, but not action-intention superiority. *J. Cogn. Neurosci.* 22, 2745–2757.

Waldert, S., Pistohl, T., Braun, C., Ball, T., Aertsen, A., and Mehring, C. (2009). A review on directional information in neural signals for brain-machine interfaces. *J. Physiol. Paris* 103, 244–254.

Wang, W., Degenhart, A. D., Sudre, G. P., Pomerleau, D. A., and Tyler-Kabara, E. C. (2011). Decoding semantic information from human electrocorticographic (ECoG) signals. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* 2011, 6294–6298.

Wittgenstein, L. (1953). *Philosophical Investigations.* Oxford: Blackwell.

Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., and Vaughan, T. M. (2002). Brain-computer interfaces for communication and control. *Clin. Neurophysiol.* 113, 767–791.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 09 August 2012; accepted: 09 October 2012; published online: 09 November 2012.

Citation: Thinnes-Elker F, Iljina O, Apostolides JK, Kraemer F, Schulze-Bonhage A, Aertsen A and Ball T (2012) Intention concepts and brain-machine interfacing. Front. Psychology ***3****:455. doi: 10.3389/fpsyg.2012.00455*

This article was submitted to Frontiers in Cognition, a specialty of Frontiers in Psychology.

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 06 September 2012
doi: 10.3389/fpsyg.2012.00312

Goal-side selection in soccer penalty kicking when viewing natural scenes

Matthias Weigelt [1] * and Daniel Memmert [2]

[1] Department of Sport and Health, University of Paderborn, Paderborn, Germany
[2] German Sports University Cologne, Institute of Cognitive and Team/Racket Sport Research, Cologne, Germany

Edited by:
Rico Fischer, Technische Universität Dresden, Germany

Reviewed by:
Peter Wühr, Technische Universität Dortmund, Germany
Gesine Dreisbach, University of Regensburg, Germany

***Correspondence:**
Matthias Weigelt, Department of Sport and Health, University of Paderborn, Warburger Straße 100, 33098 Paderborn, Germany.
e-mail: matthias.weigelt@uni-paderborn.de

The present study investigates the influence of goalkeeper displacement on goal-side selection in soccer penalty kicking. Facing a penalty situation, participants viewed photo-realistic images of a goalkeeper and a soccer goal. In the action selection task, they were asked to kick to the greater goal-side, and in the perception task, they indicated the position of the goalkeeper on the goal line. To this end, the goalkeeper was depicted in a regular goalkeeping posture, standing either in the exact middle of the goal or being displaced at different distances to the left or right of the goal's center. Results showed that the goalkeeper's position on the goal line systematically affected goal-side selection, even when participants were not aware of the displacement. These findings provide further support for the notion that the implicit processing of the stimulus layout in natural scenes can effect action selection in complex environments, such in soccer penalty shooting.

Keywords: anticipation, implicit action priming, action selection, sport performance, soccer penalty

INTRODUCTION

The ability to anticipate the consequences of ones' own actions and of the actions of other co-actors is an essential part of social interaction. Such anticipation skills are especially important for decision making in complex environments, such as in a sport setting like the soccer penalty kick, which has become one of the most prominent paradigms to investigate anticipation skills in sports (e.g., McGarry and Franks, 2000; Savelsbergh et al., 2005; Dicks et al., 2010a). In penalty kicking, two strategies have been observed for kickers when it comes to the selection of the left or right goal-side (e.g., Van der Kamp, 2006). The kicker can either anticipate the goal corner in advance and thus, select the goal-side independently of the goalkeeper, or base the decision on observing the goalkeeper's action (i.e., his/her jump direction) and react late during the execution of the penalty. An anticipation strategy in which the kicker selects the goal-side in advance has proven to be more successful than reacting to the goalkeeper's action during the run-up (Van der Kamp, 2006). From this finding, the question central to the present study arises: What kind of information do penalty kickers use to select their actions (i.e., the left or right goal-side) in advance? As we will show, these complex decisions are based on the visual processing of the action environment and basic spatial judgment.

Before taking a closer look at the visual information processes upon which the soccer kicker selects his/her goal corner during the penalty kick, we first consider the prominent theory of anticipatory behavioral control (ABC), advanced by Hoffmann (1993, 2009). According to this theory, peoples' anticipations are based on the acquisition of action-effect (A-E) representations. These A-E representations become stronger the more often a certain action leads to the desired effect. In soccer, the player learns the contingency between kicking a soccer ball in a certain way (e.g., instep kick) and the direction and/or trajectory of the ball's flight after the impact. Within the ABC-theory this is thought of "as being the primary learning process in the acquisition of behavioral competence" (Hoffmann, 2009, p. 22). Once the A-E representation is established, anticipating a certain effect (e.g., scoring on the left goal-side) will activate the appropriate action (e.g., instep kick with a small body rotation to the left while kicking with the inside of the right foot).

However, voluntary actions are not only linked to effect representations (Hommel et al., 2001), but also to the situational context in which a desired effect is consistently realized by a certain action. In fact, specific situational features become integrated into existing A-E representations. Hoffmann (2009, p. 22) considers this conditionalization of A-E representations as "being a secondary learning process." Importantly, conditionalized A-E representations will be *directly triggered*, when the situational features correspond to the represented condition (i.e., the situational context). This notion can be traced back to Lewin (1928), who used the German term "Aufforderungscharakter," to Ach (1913), who spoke of "voluntive Objektion," and/or to Gibson (1979), who proposed that objects in the environment provide "affordances" to act in a particular way. Together, these conceptions suggest that situational features trigger a certain habitual behavior, as long as people act in a specific context (e.g., soccer penalty).

Applied to the penalty situation in soccer, successful performance does not only rely on excellent kicking skills, which are based on well-established A-E representations, but also on sufficient information uptake during the visual processing of the environment, which is based on the continuous integration of situational features into existing A-E representations. During the penalty kick, processing a specific situational feature, such as the position of the goalkeeper on the goal line, may activate the corresponding A-E representation and thus, trigger a certain action in the kicker. For example, if the goalkeeper stands more on the right side of the goal, leaving a larger area on the left side uncovered, the kicker will kick to the left goal-side. Hence, when examining people's decision making in complex environments, one has to also consider the situational context in which the actions are carried out. This means that in the soccer penalty situation, anticipating the outcome of the kick (e.g., scoring on the left or right goal-side) does not only depend on good anticipation skills, but also on the sufficient processing of situational features (e.g., the exact goalkeeper's position on the goal line), which may trigger the corresponding A-E representations (e.g., a certain kicking behavior). The more general aim of the present study is therefore to investigate if goal-side selection (drawing on A-E representations) in soccer penalty kicking is influenced by the goalkeepers' position on the goal line (drawing on the processing of situational features).

The present study has been largely motivated by a recent study of Masters et al. (2007). In a video analysis of 200 penalty kicks in high-level soccer competitions (e.g., World Cups, European Championships, Africa Cup of Nations etc.), these authors observed that goalkeepers do not stand in the exact middle of the goal line in 96% of kicks, while, at the same time, professional kickers reliably select the greater goal-side. This observation led them to the question whether the penalty takers consciously perceived the goalkeeper's displacement or whether the selection of the greater goal-side was the result of implicit priming. To answer this question, Masters et al. (2007) designed three ingenious laboratory experiments with the aim to replicate and isolate the effects under controlled conditions. In Experiment 1, participants viewed a filled block, which was presented at different displacements to the left or right of a rectangle's center (scaled to 3% of a regular soccer goal). Participant's task was to indicate the larger area to the side of the rectangle. In Experiment 2, the filled block was replaced by an image of Oliver Kahn and the set-up was scaled to 44% of a regular soccer goal. The task for the participants was now to kick a soccer ball to the greater side of the goal (i.e., the side with the greater area). The same set-up was used in Experiment 3, but this time, participants were asked to only kick to the goal when they perceived the goalkeeper to be standing in the middle of the goal. The results demonstrated that participants were able to reliably indicate the greater area of the rectangle (Experiment 1) and to direct their kicks to the greater side of the goal (Experiment 2), even when executing the kick meant to indicate that they perceived the goalkeeper to be standing in goal center (Experiment 3).

Masters et al. (2007) related their findings to the empirical law of sensation from psychophysics, also known as the Weber–Fechner law (cf. Krueger, 1989). Essentially, this law captures the relationship between the (objective) physical world and the (subjective) psychological world of perception, and describes the just-noticeable difference between two physical stimuli varied along a single dimension (e.g., visual, auditory, tactile etc.). Following the empirical law of sensation, the difference of two physical stimuli can only be perceived when it overcomes a differential threshold. Applied to the penalty situation, whether or not the kicker will recognize the shift of the goalkeeper depends on the size relation between goal and goalkeeper, as well as on the viewing distance of the penalty kicker. Participants in Masters et al.'s study directed their kicks more often to the greater goal-side already for goalkeeper displacements of only 0.5%. This effect of implicit priming on goal-side selection was independent of the size of the stimulus display (scaled to 3 vs. 44% of the real penalty situation) and corresponded to the just-noticeable difference reliably found in line-bisection studies (e.g., Jewell and McCourt, 2000). The most surprising aspect of the findings was, however, that participants only became aware of the displacement when the goalkeeper position was shifted by 3% on the goal line. Hence, goal-side selection was driven by perceptual discriminations that were not consciously perceived by the participants.

On the basis of the original study by Masters et al. (2007), the present experiment investigates the effect of implicit priming on goal-side selection when presenting a natural scene (i.e., a goalkeeper in real soccer goal) to the kicker. In the original study, degraded stimulus images were used that consisted of a rectangle, which represented the goal, and a filled block (Experiment 1) or an image of Oliver Kahn (Experiments 2 and 3), which represented the goalkeeper. Interestingly, when looking at the image of Oliver Kahn, it appears that he was either shown with his arms behind the back or without arm. In any event, these stimuli did not represent the natural environment of the penalty situation, because in a regular game, the penalty kicker does not simply kick to a rectangle and at the same time, the goalkeeper does not passively await the kick with her/his arms behind her/his back. It is therefore the question whether a similar pattern of goal-side selection can be found when using photo-realistic images of the penalty situation.

This extension of the stimulus material to photo-realistic stimuli is in line with current approaches to investigate the mechanisms of the visual system in natural scenes (for a review see Felsen and Dan, 2005). Most of our knowledge about the visual system is gathered from experiments using simple stimuli, either displaying bright spots on dark backgrounds, or dark spots on light background. This methodological approach of simplistic stimulus presentation has also been used in line-bisection tasks (e.g., Lindell et al., 2007), and in the soccer penalty study by Masters et al. (2007) described above. Line-bisection tasks have been mostly used in basic research on visuospatial neglect (Lindell et al., 2007) and pseudoneglect (Jewell and McCourt, 2000). Recently, line-bisection performance has been directly related to more complex natural environments (Nicholls et al., 2010). These natural environments, in which people carry out our actions, however, are made up of rich colors and a distinct spatial structure. Importantly, the visual system has adapted to process the characteristic visual properties of natural scenes (Simoncelli and Olshausen, 2001; Kayser et al., 2004). In fact, there is evidence that the stimulus types often used in laboratory experiments are not representative of natural viewing behavior (Dorr et al., 2010). For example, when people look at natural scenes, their saccadic latencies are significantly shorter, allowing for faster reactions to potentially critical stimuli (White et al., 2008). Also, color information facilitates the processing of natural scenes (Delorme et al., 1999). Thus, from a methodological point of view, it is important to examine whether the results obtained from simple stimulus presentations [i.e., filled block (Experiment 1) and degraded black and white image of Oliver Kahn (Experiments 2 and 3) depicted on a white rectangle] by Masters et al. (2007) extend to task-contexts in which photo-realistic images of natural stimuli (i.e., real-world penalty scenario) are used.

To this end, photo-realistic stimulus images of a regular soccer goal and a goalkeeper standing on the goal line in a neutral goalkeeping posture were used in the present study. Specifically, the goalkeeper was displayed in a parallel stance with his knees slightly bend, his arms in a "ready-to-catch" position on the side, and the gaze straight ahead, focusing on the penalty taker. The goalkeeper was positioned on the goal line at different displacements and participant's task was to kick to the side with the greater goal area in the first part of the experiment. A stimulus image with the goalkeeper standing in the exact middle on the goal line was also included into the experiment (a condition not present in the original study by Masters et al., 2007). This was done to examine a potential bias of kicking to either the left or right side of the goal.

One general and one specific prediction were made for the present experiment. The general prediction referred to the goalkeeper's displacement and stated that participants would direct more kicks to the greater goal-side, even under conditions in which they were not aware of the displacement. If true, this would show that a similar pattern of goal-side selection can be found when using photo-realistic images of the penalty situation. The specific prediction related to the inclusion of the condition in which the goalkeeper was presented in the exact middle of the goal. Here, it is predicted that the right-footed participants of the present study select the right goal-side more often than the left goal-side. This prediction is based (purely) on inferences from the biomechanics of kicking. Accordingly, right-footed players will approach the ball from the left side, resulting in a run-up direction to the right. Continuing to kick the ball to the right goal-side is then easier than changing the kicking direction to the left goal-side. Surprisingly, and to the best of our knowledge, nothing is known about whether such a bias of kicking direction can be observed in competitive soccer.

MATERIALS AND METHODS

PARTICIPANTS

A total of 23 participants (nine females; mean age = 21.6 years; ranging from 18–27 years) with normal or corrected-to-normal vision took part in this experiment. All participants were sport science students at Bielefeld University and naïve to the purpose of the present experiment. However, none of the participants was an active soccer player, or had extensive practice in this sport. All reported to be right handed and right-footed. Before being tested, each participant gave his or her written informed consent. They were not paid for their participation. The study was approved by the local ethics committee and was carried out in accordance with the Helsinki Declaration of 1975.

APPARATUS AND STIMULI

A specific set-up was developed to take the penalty situation in soccer into the laboratory. To this end, the overall dimension of the set-up was downscaled to ~44%, while keeping the relative distances constant to the real penalty situation. Accordingly, the goal was 3.19 m wide (real goal = 7.32 m) and 1.07 m high (real goal = 2.44 m). The penalty spot was placed at a distance of 4.80 m to the goal line (real penalty spot = 11.00 m). These dimensions were similar to the one reported by Masters et al., 2007, Experiments 2 and 3). All pictures were displayed onto a large, white wall with a projector, which was installed at a height of 2.60 m and a distance of 6.10 m to the projection wall. The position of the projector was carefully chosen, so that the view of the stimulus image was not obstructed by the participant during the task. At the same time, the stimuli were also visible on the experimenter's laptop (shielded from view by the participant). In the action selection task, all shots were taken with a standard indoor soccer ball, made of hard foam.

Stimuli were taken with a digital camera on an outdoor soccer pitch. They displayed a goalkeeper wearing a goalkeeper's outfit and standing in a neutral goalkeeping posture in a standard size soccer goal. One picture was taken with the goalkeeper in the goal and one from the empty goal. Stimuli were then further edited on the PC with Corel Paint Shop Pro. To this end, the goalkeeper was carefully cut out and copied into the picture with the empty goal in one of nine positions, either in the goal's center or 1.5, 3, 6, and 12% to the left and right of center (see **Figure 1**). This resulted in a total of nine stimulus images. The displacements of the goalkeeper to the left or right of the center relate to 11, 22, 44, and 88 cm in the realistic penalty situation.

EXPERIMENTAL TASKS

The experiment consisted of two parts in which participants performed an *action selection task* and a *perception task*. In both tasks, participants viewed a photo-realistic, static image of a goalkeeper, who stood either in the middle of the goal, or was marginally shifted to the left or right of the goal center. The goalkeeper was depicted in a neutral goal keeping posture (see **Figure 1**). The *action selection task* required participants to direct a penalty shot to the "open corner" of the goal (i.e., motor response). Similar to Masters et al., 2007, Experiment 2), they were instructed to select the side with the greater area uncovered. If they felt unsure about which side represented the greater area, they were instructed to follow their first impression and to take the shot without further contemplation. The *perception task* required participants to indicate (i.e., verbal answer) whether the goalkeeper stood in the middle, or was shifted to the left or right, respectively. Again, if they felt unsure about the actual position of the goalkeeper, they were instructed to follow their first impression.

DESIGN AND PROCEDURE

All participants were tested first in the *action selection task*. They started with a practice block, in which each of the nine stimulus images was presented one at a time. The following test block consisted of 18 trials. Each trial started with the ball on the penalty spot and a blank screen for 1–2 s, whereupon, the experimenter presented the stimulus image on the screen. Then, the participant took a short approach of approximately 0.5–1.0 m and kicked the ball to the side, which she/he thought to present the greater area of the goal. Here it is important to note, that because participants were all right-footed, they approached the ball from the left side. After the shot was taken, the screen turned blank again and the ball was fetched (and returned to the penalty spot) by the participant. The next trial followed by presenting a new stimulus image. All stimulus images were presented in a pseudo-random order, which was predefined before the experiment and kept constant for each participant.

In the second part of the experiment, participants went on to be tested in the *perception task*. The procedure was similar to the first part, with one exception: The perception task did not require them to kick the ball to the goal, but instead to give a verbal judgment (i.e., explicit decision) of whether the goalkeeper stood in the middle of the goal, or to the left and right of the goal center.

FIGURE 1 | Depicted are five of nine stimulus displays used in the present study. The goal keeper was shown in a neutral goalkeeping posture, either in the goal's center or in one of four displacements to the left (not displayed here) or right of center.

Thus, the instruction regarding the perception task was different to the instruction in the action selection task. The perception task was included in order to learn more about conscious and unconscious information processing and to derive a more explicit measure of participant's perception of the goalkeeper's position on the goal line. This aspect of the present study is different from the procedures of the original study by Masters et al. (2007), where participant's perception of the goalkeeper's position was inferred from individual confidence ratings (representing an indirect measure of participant's perception). To this end, participants viewed each stimulus image while standing on the penalty spot. All stimulus images were presented in a new pseudo-random order.

Participants always started with the *action selection task* and then proceeded with the *perception task*. This order was not counterbalanced, because we wanted to avoid that participants would spend too much thought on the goalkeeper's displacement in the first part of the experiment. In fact, the displacement of the goalkeeper was not mentioned to the participants before they actually started the perception task. Participants were allowed to take a short break between the two parts and the whole experiment lasted about 20 min.

DATA COLLECTION

Participant's responses relative to the side of the goal (i.e., motor response) and the goalkeeper's position (i.e., verbal response) were noted by the experimenter on two separate experimental score sheets, for the action selection task and the perception task, respectively. These score sheets contained the experimental schedule for the presentation of the stimulus images. Accordingly, the experimenter wrote down if the participant directed her/his kick to the left or right side of the goal in the first part of the experiment and where the participant perceived the goalkeeper standing in the second part. Thereby, kicking accuracy was of no further interest, so that all trials were also counted in which the ball would have actually missed the goal to the left or right.

DATA ANALYSES

For the *action selection task*, data were analyzed for the different displacement conditions 1.5, 3, 6, and 12% of goal center. Since kicking side is a dichotomous variable, kicks to the smaller side of the goal were assigned a value of 0 and kicks to the greater side of 1. For each displacement condition, the sums of all trials were then divided by the number of trials times 100 to receive percentage values for kicking to the greater side for each participant. Planned comparisons (i.e., one-sample t-Tests) against chance level (50%) were conducted, beginning with the smallest displacement condition of 1.5% and continuing until a significant difference was reached. The data for the stimulus image in which the goalkeeper stood in the middle were analyzed separately to examine whether the (right-footed) participants had an implicit bias to direct their kicks to one or the other side of the goal.

For the *perception task*, participant's judgments were analyzed to whether they indicated the goalkeeper to be in the middle of the goal, or not. Thus, the variable goalkeeper position was treated as a dichotomous variable, although participants could further indicate a displacement to the left or right side. Verbal judgments of left or right side of goal center received a value of 0, whereas a value of 1 was given when participants indicated the goalkeeper to be standing in the middle of the goal. The data for the stimulus image in which the goalkeeper was not displaced on the goal line (that is, zero-displacement) were analyzed separately to examine participants' perceptual variability, e.g., perceiving the goalkeeper shifted to the left or right, when (in fact) he was not displaced, but stood in the exact middle of the goal. For the remaining conditions, the data was then further analyzed with planned comparisons (i.e., paired-samples t-Tests) between the zero-displacement condition and all displacement conditions, beginning with the smallest displacement condition of 1.5% and continuing until a significant difference was reached.

RESULTS

ACTION SELECTION TASK

When the goalkeeper was not displaced and presented in the exact middle of the goal, more kicks were directed to the right goal-side (27 out of 46 = 58.7% of the kicks) than to the left goal-side (19 out of 46 = 41.3% of the kicks). Accordingly, participants showed a kicking bias to the right goal-side. This kicking bias was in the magnitude of 17.4%.

Participant's goal-side selections under conditions in which the goalkeeper was displaced along the goal line are shown in **Figure 2**. The solid circles represent the percentages of selecting the greater side of the goal for the different displacement conditions. The mean percentages for the different displacements were 69.6, 78.3, 85.9, 97.8%, from smallest to largest respectively. The planned comparison of the smallest displacement condition revealed a significant difference, $t(22) = 3.600$; $p < 0.01$. Hence, participants selected the larger goal-side already for the smallest displacement of the goalkeeper (i.e., 1.5% of goal center).

PERCEPTION TASK

When the goalkeeper was not displaced and shown in the exact middle of the goal, participants indicated the middle position in the majority of trials (38 out of 46 trials = 82.6%), whereas on some occasions they reported the goalkeeper to be displaced (8

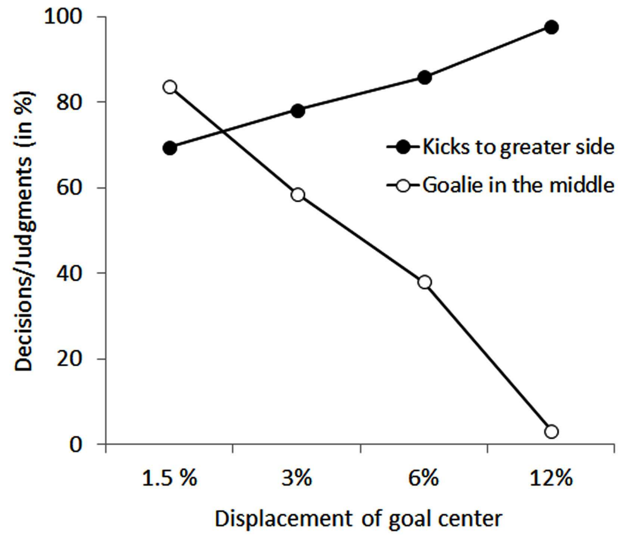

FIGURE 2 | Shows the pattern of results for kicks directed to the greater side of the goal as a function of goalkeeper's displacement and participant's expertise in the action selection task (AST, solid circles), as well as for judging the goalkeeper to be standing in the middle of the goal in the perception task (PT, open circles). Error bars indicate between-participant standard error.

out of 46 = 17.4% of trials), irrespective of the fact that he was not moved.

Figure 2 shows the percentages of trials in which participants perceived the goalkeeper to be standing in the middle of the goal, although he was displaced along the goal line (open circles). As can be seen, the further the goalkeeper was standing off center, the less often participant's perceived him in the middle. Accordingly, the percentages decreased from 83.7, 58.7, 38.0, to 3.3%. Planned comparisons between participants' perceptual judgments in the zero-displacement condition and all displacement conditions revealed the following (none-)significant differences: For the 1.5% condition, participants perception was not different to the zero-displacement condition and thus, they were not aware of the displacement, $t(22) = 0.182$, $p = 0.86$. For the displacement condition of 3%, participant's perceptual judgment of the goalkeeper position was significantly different from the zero-displacement condition. Hence, they became aware of the displacement when the goalkeeper was shifted by 3%, $t(22) = 2.975$, $p < 0.01$.

DISCUSSION

Anticipatory behavior does not only rely on well-established A-E representations, but also on the processing of the situational context in which these actions are carried out. When there is a high correspondence between the conditionalized A-E representations and the situational conditions, specific situational features will trigger the associated behavioral response by activating the conditionalized A-E representation. Hence, processing the action environment is essential for action selection. The general aim of the present study was to investigate whether an action instruction eases the selection of the open goal corner (when compared to a mere perception condition) as predicted by A-E theories (e.g., Hoffmann, 1993, 2009). Based on a recent study by Masters et al. (2007), who first reported implicit priming effects on goal-side selection, participants viewed photo-realistic images of a goalkeeper and a soccer goal. In the action selection task, they were asked to kick to the greater goal-side, whereas in the perception task, they indicated the position of the goalkeeper, who was either presented in the exact middle of the goal or was displaced at different distances to the left or right on the goal line. Two predictions were made: The general prediction related to implicit priming effects on participants' goal-side selection and stated that participants would direct more kicks to the greater goal-side, even if they were not aware of the goalkeeper's displacement. The specific prediction referred to a potential goal-side selection bias and predicted that the right-footed participants would kick more often to the right goal-side. The results of the present study are in line with these predictions and are discussed in the following.

With regard to the more general prediction of implicit priming effects on goal-side selection, the results of the action selection task confirmed that participants directed their kicks to the side of the goal with the greater area. This was already the case for the smallest displacement of 1.5% to the left or right, which corresponds to a distance of 11 cm in a real soccer goal. When asked to provide a verbal judgment of the goalkeeper's position under this condition in the perception task, participants stated that they perceived the goalkeeper to be standing in the middle of the goal in the great majority of trials. Importantly, participants were not able to discriminate this small displacement of 1.5% from the zero-displacement condition (i.e., goalkeeper displayed in goal center). Therefore, it can be argued that they were not aware of the displacement. With larger displacements of the goalkeeper away from the goal's center, participants more likely perceived him on the left or right side. They were becoming aware when the displacement was 3% and larger. This pattern of results replicates the implicit priming effects on penalty-kicking direction reported by Masters et al. (2007). It extends these findings, however, to a more realistic setting in which participants viewed photo-realistic images of a real soccer goal and a goalkeeper standing in a typical goalkeeping posture.

Another interesting observation was made for those stimuli, which displayed the goalkeeper to be standing in the exact middle of the goal. Here, participants indicated the goalkeeper to be standing in goal center in 82.6% of the trials. Participant's performance was well above chance, but not perfect under this condition. At the same time, participants perceived the goalkeeper to be standing in the middle, when he was (in fact) displaced by a small degree (i.e., in 83.7% of trials under a 1.5% displacement). What may be the reason for this variability in the accuracy of participants' perceptual judgments? It is possible that viewing natural scenes may have induced noise in the visual system and led to more (individual) variability in the processing of the spatial layout. Such individual variability during the visual processing of natural scenes has been reported elsewhere (Dorr et al., 2010) and may have affected participants' judgment. Unfortunately, this pattern of results cannot be compared to the results of Masters et al. (2007), because these authors never presented the goalkeeper in the exact middle of the goal, even though participants were asked to kick only if they perceived the goalkeeper to be standing in goal center (Experiment 3). Thus, participants were instructed to respond to an experimental condition that was actually not included in the experiment. It can only be speculated that on some occasions, participants would not have responded, even if the authors had displayed the goalkeeper in the middle. This hypothetical result would be similar to the finding of perceptual variability in the present experiment.

With regard to the specific prediction referring to a potential goal-side selection bias, results showed that participants directed their kicks more often to the right side of the goal when the goalkeeper was presented in the exact middle. This goal-side selection bias may be explained by the biomechanics of kicking. Arguably, continuing to kick the ball in the direction of the run-up from left to right is easier in terms of skill execution (i.e., kicking the ball) than changing the kicking direction to the left goal-side. There is an alternative explanation for this goal-side selection bias, however. This alternative explanation is based on the observation that in about 17% of the trials in which the goalkeeper was presented in the middle of the goal, participants erroneously reported him to stand off center. If this perceptual "error" were systematic and included only mislocations to the left, then this may also explain the higher number of kicks to the right goal-side. However, when examining the participant's judgments in the perception task for the smallest displacement (i.e., 1.5%) to the left side (goalie in the middle = 82.6%) and right side (goalie in the middle = 84.8%) separately and then comparing these number to the zero-displacement condition (goalie in the middle = 82.6%), a perceptual bias cannot be detected. Therefore, it seems rather unlikely, that the right-side kicking bias observed in the action selection task, being in the magnitude of 17.4%, can be explained by a perceptual bias. In any case, such a bias on goal-side selection has not been examined systematically for the penalty situation in competitive soccer. What has been reported, however, is that right-footed kickers score more often on the left goal-side (from their perspective), while left-footed kickers are more successful on the right side (Coloma, 2007). But this was not based on a more general kicking bias to one goal-side. More research is certainly needed to further determine the influence of the goal-side selection bias on kicking performance.

What kind of implications can be drawn for sports practice from the present experiment? For the penalty situation in soccer, specific performance-related instructions can be provided for goalkeepers to improve their success rate. Goalkeepers could use this knowledge about the implicit priming effect on goal-side selection strategically, by placing themselves a little more to their "weaker side" on the goal line. This will increase the likelihood of the penalty taker to kick to the opposite side of the goal, which corresponds with the "stronger side" of the goalkeeper. A goalkeeper displacement of 1.5% resulted in roughly 69.6% of kicks to the greater goal-side. At the same time, this small displacement, which corresponds to 11 cm in the real-size soccer goal, was not perceived by the participants, neither in the present study nor in the study of Masters et al. (2007). Instead of such a strategic displacement on the goal line, goalkeepers can also use explicit gestures. For example, an active goalkeeper that moves and waves her/his arms around can effectively distract the penalty taker (Wood and Wilson, 2010). Also, explicit signaling with specific pointing gestures to the left or right goal-side can render the upcoming kicking direction more predictable (Weigelt et al., in press).

In summary, decision making in soccer penalty kicking (i.e., goal-side selection) can be systematically influenced by the goalkeeper through the utilization of implicit (i.e., goalkeeper displacement) information strategies. The present experiment therefore adds further empirical evidence to the growing body of research on perception-action-coupling, using the soccer penalty situation as an experimental paradigm (e.g., Masters et al., 2007; Dicks et al., 2010a,b). It extends previous research by using photo-realistic stimulus material. The present findings provide valuable implications for specific performance-related instructions to benefit the performance of goalkeepers and penalty takers. To take the present findings from the laboratory to the field, future studies should examine the effectiveness of such performance-related instructions with representative task designs (cf. Dicks et al., 2009).

REFERENCES

Ach, N. (1913). "Über den Willen [On the will]," in *Untersuchungen zur Psychologie und Philosophie [Studies on Psychology and Philosophy]*, ed. N. Ach (Leipzig: Quelle & Meyer).

Coloma, G. (2007). Penalty kicks in soccer. *J. Sports Economics* 5, 530–545.

Delorme, A., Richard, G., and Fabre-Thorpe, M. (1999). Rapid processing of complex natural scenes: a role for the magnocellular visual pathways? *Neurocomputing* 26–27, 663–670.

Dicks, M., Button, C., and Davids, K. (2010a). Availability of advance visual information constrains association-football goalkeeping performance during penalty kicks. *Perception* 39, 1111–1124.

Dicks, M., Davids, K., and Button, C. (2010b). Individual differences in the visual control of intercepting a penalty kick in association football. *Hum. Mov. Sci.* 29, 401–411.

Dicks, M., Davids, K., and Button, C. (2009). Representative task designs for the study of perception and action in sport. *Int. J. Sport Psychol.* 40, 506–524.

Dorr, M., Martinetz, T., Gegenfurtner, K. R., and Barth, E. (2010). Variability of eye movements when viewing dynamic natural scenes. *J. Vis.* 10, 28.

Felsen, G., and Dan, Y. (2005). A natural approach to studying vision. *Nat. Neurosci.* 8, 1643–1636.

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Mahwah, NJ: Lawrence Erlbaum Associates.

Hoffmann, J. (1993). *Vorhersage und Erkenntnis: Die Funktion von Antizipationen in der menschlichen Verhaltenssteuerung und Wahrnehmung* [Anticipation and Cognition: The Function of Anticipations in Human Behavioral Control and Perception]. Göttingen: Hogrefe.

Hoffmann, J. (2009). "ABC: A psychological theory of anticipatory behavioral control," in *Anticipatory Behavior in Adaptive Learning Systems. From Psychological Theories to Artificial Cognitive Systems*, eds G. Pezzulo, M. V. Butz, O. Sigaud, and G. Baldassarre (Heidelberg: Springer), 10–30.

Hommel, B., Muesseler, J., Aschersleben, G., and Prinz, W. (2001). The theory of event coding (TEC): a framework for perception and action planning. *Behav. Brain Sci.* 24, 849–937.

Jewell, G., and McCourt, M. E. (2000). Pseudoneglect: a review and meta-analysis of performance factors in line bisection tasks. *Neuropsychologia* 38, 93–110.

Kayser, C., Körding, K. P., and König, P. (2004). Processing of complex stimuli and natural scenes in the visual cortex. *Curr. Opin. Neurobiol.* 14, 468–473.

Krueger, L. E. (1989). Reconciling Fechner and Stevens: toward a unified psychophysical law. *Behav. Brain Sci.* 12, 251–320.

Lewin, K. (1928). *Wille, Vorsatz und Bedürfnis* [Will, Intention, and Volition]. *Psychologische Forschung [Psychol. Res.]* 7, 330–385.

Lindell, A. B., Jalas, M. J., Tenovuo, O., Brunila, T., Voeten, M. J. M., and Hämäläinen, H. (2007). Clinical assessment of hemispatial neglect: evaluation of different measures and dimensions. *Clin. Neuropsychol.* 21, 479–497.

Masters, R. S. W., van der Kamp, J., and Jackson, R. C. (2007). Imperceptible off-center goalkeepers influence penalty-kick direction in soccer. *Psychol. Sci.* 18, 222–223.

McGarry, T., and Franks, I. M. (2000). On winning the penalty shoot-out in soccer. *J. Sports Sci.* 18, 401–409.

Nicholls, M. E. R., Loetscher, T., and Rademacher, M. (2010). Miss to the right: the effect of attentional asymmetries on goal-kicking. *PLoS ONE* 5, e12363. doi:10.1371/journal.pone.0012363.

Savelsbergh, G. J., van der Kamp, J., Williams, A. M., and Ward, P. (2005). Anticipation and visual search behaviour in expert soccer goalkeepers. *Ergonomics* 48, 1686–1697.

Simoncelli, E. P., and Olshausen, B. A. (2001). Natural image statistics and neural representation. *Ann. Rev. Neurosci.* 24, 1193–1216.

Van der Kamp, J. (2006). A field simulation study of the effectiveness of penalty kick strategies in soccer: late alterations of kick direction increase errors and reduce accuracy. *J. Sports Sci.* 24, 467–477.

Weigelt, M., Memmert, D., and Schack, T. (in press). Kick it like ballack: the effects of goalkeeping gestures on goal-side selection in experienced soccer players and soccer novices. *J. Cogn. Psychol.*

White, B. J., Stritzke, M., and Gegenfurtner, K. R. (2008). Saccadic facilitation in natural backgrounds. *Curr. Biol.* 18, 124–128.

Wood, G., and Wilson, M. R. (2010). A moving goalkeeper distracts penalty takers and impairs shooting accuracy. *J. Sports Sci.* 28, 937–946.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 14 June 2012; accepted: 07 August 2012; published online: 06 September 2012.

*Citation: Weigelt M and Memmert D (2012) Goal-side selection in soccer penalty kicking when viewing natural scenes. Front. Psychology **3**:312. doi: 10.3389/fpsyg.2012.00312*

This article was submitted to Frontiers in Cognition, a specialty of Frontiers in Psychology.

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 27 September 2012
doi: 10.3389/fpsyg.2012.00370

Actions seen through babies' eyes: a dissociation between looking time and predictive gaze

Moritz M. Daum[1,2]***, Manja Attig**[1]**, Ronald Gunawan**[1]**, Wolfgang Prinz**[1] **and Gustaf Gredebäck**[3]

[1] Research Group Infant Cognition and Action, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany
[2] Developmental Psychology Unit, Department of Psychology, University of Zurich, Zurich, Switzerland
[3] Department of Psychology, Uppsala University, Uppsala, Sweden

Edited by:
Dorit Wenke, Humboldt University at Berlin, Germany

Reviewed by:
Vincent Reid, University of Durham, UK
Markus Paulus, Ludwig Maximilian University, Germany

****Correspondence:***
Moritz M. Daum, Developmental Psychology Unit, Department of Psychology, University of Zurich, Binzmühlestrasse 14, Box 21, CH-8050 Zürich, Switzerland.
e-mail: daum@psychologie.uzh.ch

In this study, we explored the relation of two different measures used to investigate infants' expectations about goal-directed actions. In previous studies, expectations about action outcomes have been either measured after the action has been terminated, that is post-hoc (e.g., via looking time) or during the action is being performed, that is online (e.g., via predictive gaze). Here, we directly compared both types of measures. Experiment 1 demonstrated a dissociation between looking time and predictive gaze for 9-month-olds. Looking time reflected identity-related expectations whereas predictive gaze did not. If at all, predictive gaze reflected location-related expectations. Experiment 2, including a wider age range, showed that the two measures remain dissociated over the first 3 years of life. It is only after the third birthday that the dissociation turns into an association, with both measures then reflecting identity-related expectations. We discuss these findings in terms of an early dissociation between two mechanisms for action expectation. We speculate that while post-hoc measures primarily tap ventral mechanisms for processing identity-related information (at least at a younger age), online measures primarily tap dorsal mechanisms for processing location-related information.

Keywords: infancy, action perception, eye movements, looking time, predictive gaze, dorsal-ventral

INTRODUCTION

A hallmark of social-cognitive development is the ability to understand others' actions flexibly and quickly. Infants have been shown to do so at an early age. Furthermore, at this early age, infants interpret the various components that constitute actions, such as intentions and goals as well as movements and means involved in achieving goals (Wagner and Carey, 2005).

One way to assess infants' action perception is to measure their expectations about a forthcoming action, together with their responses when these expectations are met or violated. Two approaches have been predominantly used to do this. In the *post-hoc* approach, expectations are measured via *looking time*, for example in response to an observed action that is completed (e.g., Woodward, 1998). In contrast, the *online* approach is to measure expectations in anticipation of forthcoming action, for example, through *predictive gaze* during the observation of an ongoing action (e.g., Falck-Ytter et al., 2006). In the present study, we combined these two approaches in order to investigate how measures assessing infants' expectations post-hoc and online of are related to each other in development.

THE POST-HOC APPROACH

The measurement of looking times as an indicator of infants' cognition has been one of the most powerful tools in infancy research through the past 30 years. In a typical paradigm, infants are first habituated to a standard event. During the phase of habituation, infants build a representation of this specific event that allows them to form expectations about future events of a similar structure. Once a habituation criterion has been reached, two test events are presented that are variations of the habituation event. In one of the test events, the previously built expectations are met, in the other event, these expectations are violated. Longer looking times to one of the two test events indicate that the infants differentiated between the two events, that is, they could make use of the representation build previously and apply it to a novel situation.

In the context of infants' perception of goal-directed actions, post-hoc measures were used to assess whether infants' expectations about the outcome of an action are violated (resulting in longer looking time) or not when presented with test trials in which specific aspects of an observed action are altered compared to previously presented familiarization trials. To exemplify, Woodward (1998) habituated 6-month-olds to a hand reaching for one of two objects. In test trials, object locations were swapped and the hand either reached for the old object in a new location or the new object in the old location. Infants looked longer when the hand had reached for the new, relative to the old object, suggesting that they encoded the goal of the reaching action during familiarization and reacted with extended looking time when the agent changed its goal during test. Further studies using looking time have demonstrated that 6- to 12-month-olds encode goals of incomplete actions (Daum et al., 2008), the rationality of observed actions (Gergely et al., 1995), recognize the goal-directedness of successful, and failed reaching actions (Brandone and Wellman, 2009), and recognize goals of action sequences (Sommerville and Woodward, 2005), to list only a few. Looking time measurements allow a direct comparison between different sources of information (e.g., goal location vs. identity); however, infants' responses are measured with low spatial and temporal resolution making it difficult to relate looking time data to underlying processes, a fact that has been discussed by a large set of research in the past already (e.g., Aslin, 2007).

THE ONLINE APPROACH

The use of online prediction to investigate infants' cognitive processes and sensorimotor integration has a similarly long tradition in infancy research. Studies using predictive reaching have shown that at the same age that infants start to reach for stationary objects, they start to reach for slowly moving objects (von Hofsten and Lindhagen, 1979). Furthermore, infants' reaching movements have been shown to be predictive: Arm and hand movements are initiated before the target is within reaching distance, and are directed toward a future interception position (von Hofsten, 1980, 1983; Clifton et al., 1993). Infants' reaching and grasping abilities have been shown to be predictive in various other aspects like adjusting the orientation (Lockman et al., 1984; von Hofsten and Fazel-Zandy, 1984; von Hofsten and Johansson, 2009) or the aperture size of the hand relative to a target (von Hofsten and Rönnqvist, 1988), and by predicting the weight of an observed object (Mounoud and Bower, 1974).

Measuring infants' expectations online via predictive gaze is a relatively novel approach in infancy research (Gredebäck et al., 2010) although extensively used in adults (Flanagan and Johansson, 2003). This measure records an observer's eye movements and measures the ability to predict ongoing events (e.g., looking at the final state of an event before accomplishment).

A growing number of eye tracking studies has reported infants' abilities to predict the reappearance of objects that were shortly occluded (for a methodological review see Gredebäck and von Hofsten, 2007). This research has shown that infants as young as 4-month-olds already predict the reappearance of shortly occluded objects (Johnson et al., 2003; Rosander and von Hofsten, 2004). At the age of 6 months, infants' predictions are no longer constrained to linear motion paths but they now quickly adjust their expectations to new non-linear motion paths (Kochukhova and Gredebäck, 2007).

A second application of measuring predictive gaze has been reported from studies testing infants' categorization skills (McMurray and Aslin, 2004; Kovacs and Mehler, 2009; Addyman and Mareschal, 2010; Albareda-Castellot et al., 2011). McMurray and Aslin (2004), for example, developed an occlusion based anticipatory eye movement (AEM) paradigm where infants were presented with a training session in which one of two objects disappeared behind a T-shaped occluder and reappeared in one of two locations, depending on the identity of the moving object. Their results showed that infants learned to categorize different stimuli along a variety of stimulus dimensions such as color, orientation, or shape.

Measuring predictive gaze is specifically interesting in the context of investigating infants' perception of others' actions, as an action *per se* includes anticipation (von Hofsten, 2004). There are a number of studies using eye tracking to measure infants' expectations online via predictive gaze. These studies have, for example, demonstrated that 6-month-olds predict that food will be brought to the mouth (Kochukhova and Gredebäck, 2010) and that 12- to 14-month-olds predict the goal of manual object displacements (Falck-Ytter et al., 2006; Melzer et al., 2012) and reaching actions (Gredebäck et al., 2009; Kanakogi and Itakura, 2011; Cannon and Woodward, 2012). The measurement of eye movements in general and of predictive gaze shifts in specific allows a detailed mapping of the spatial and temporal dynamics of infants' action perception. In the same line as online measures offer advantages as compared to post-hoc measures, such as the track behavior on a fine-grained time scale, it has its limitations, for example, by constrained processing time and information.

COMPARING THE TWO APPROACHES

Post-hoc and online measures have not always revealed similarities in onset and development of action expectations, Cannon and Woodward (2012), for example, report predictive gaze shifts toward the correct target at the age of 11 months, while Woodward (1998) reports differences in looking times already being present at 5–6 months (or even earlier, as reported by Luo, 2011). One reason for this difference might be that the bases on which these expectations are built differ with respect to available information and time constraints. When measured post-hoc, expectations about an action are compared to the outcome of an action *after* it has been completed. The information about the action is *complete*. In contrast, when expectations are measured online, the measurement takes place *prior to* the completion of an observed action. The information about the action available is thus *incomplete*.

Given these differences, little is known about how these two measures relate to each other; whether they, for example, tap similar or different underlying cognitive systems. Only few studies have simultaneously used two different measures to assess infants' action expectations (Gredebäck and Melinder, 2010; Paulus et al., 2011b). Gredebäck and Melinder (2010) demonstrated that 6- and 12-month-olds' responses were more experience-dependent and developed later when measured online (via predictive gaze) than when measured post-hoc (via pupil dilations). Paulus et al. (2011b) showed that infants' predictions did not reflect their looking times in an adapted version of the rational action paradigm as reported by Gergely et al. (1995).

This is first evidence that action perception abilities might be based on different underlying mechanisms and that more attention is required to map out what processes are tapped when investigating infants' action perception.

Our aim here is to further explore the relation of the two approaches and the respective different measures that are used to investigate infants' action perception. Looking time studies have shown that infants expect actions to be organized around goal identities rather than goal locations (Woodward, 1998). A great majority of predictive gaze studies within the domain of action perception, however, used single goals at fixed locations (Falck-Ytter et al., 2006; Cannon et al., 2012) or an assembly of similar goal objects at the same location (Gredebäck et al., 2009), leaving the question open whether predictive gaze is based on goal identity or location (see Paulus et al., 2011b; Cannon and Woodward, 2012, for exceptions). We adapted the looking time paradigm introduced by Woodward (1998) that includes two different goals at two distinct locations and combined it with a predictive gaze paradigm (similar to McMurray and Aslin, 2004; Kochukhova and Gredebäck, 2007).

EXPERIMENT 1

In Experiment 1, infants' action expectations were measured via looking times and predictive gaze shifts. The paradigm that was used primarily followed the logic of Woodward (1998); infants were familiarized with an agent moving toward one of two objects. In a subsequent test phase, the positions of the two objects were swapped and the agent either moved toward the old object on a new movement path (old goal/new path event) or to the new object on the old movement path (new goal/old path event). In order to be able to measure looking times and predictive gaze shifts at the same time, the original paradigm was modified as follows.

First, to trigger predictive gaze shifts, we followed the rationale of the occlusion based AEM paradigm (McMurray and Aslin, 2004) by adding an circular occluder in the center of the screen (similar to Kochukhova and Gredebäck, 2007). The agent moved toward the occluder, disappeared below the occluder and reappeared at the side of one of two targets, see **Figure 1**.

Second, in order to have a well-defined agent that completely disappears behind the occluder, we replaced the human hand by an animated agent, a small red fish, who moved fish-like (i.e., by wiggling its tail). Animated agents have been successfully used in studies investigating infants' action expectations. Beginning with 6 months, infants are sensitive for the goal-directed behavior and the rationality of a wide range of human as well as animated agents (Csibra et al., 2003; Kamewari et al., 2005; Wagner and Carey, 2005; Csibra, 2008; Schlottmann and Ray, 2010). Using a paradigm similar to Woodward (1998), infants at the age of 6 months (Luo and Baillargeon, 2005) and even as young as 3 months (Luo, 2011) attribute goals to animated non-human agents. The agent that was used in the present study was designed to entail a variety of cues that have been shown to be help infants to perceive actions as goal-directed (e.g., self-propelledness, Biro and Leslie, 2007).

Third, we used a partially infant-controlled familiarization procedure and presented a fixed number of eight familiarization trials to each infant. This familiarization procedure has been successfully used in previous studies investigating infants' goal attribution abilities using modified versions of the original Woodward paradigm (Hofer et al., 2007, 2008).

We tested 9-month-old infants as infants at this age show a robust goal attribution effect and anticipate action goals (Hunnius and Bekkering, 2010). Our hypothesis about the looking time results was clear. Based on the previous findings mentioned above, we expected infants to look longer at new goal/old path events compared to old goal/new path events. The hypotheses concerning predictive gaze were less obvious. Based on previous results, two outcomes concerning the infants' predictions are conceivable. First, if infants attribute goals to an agent based on the identity of the goal as expected by their looking time (and as reported by Woodward, 1998), infants' predictions should likewise be related to the identity of the goal (Cannon and Woodward, 2012). Second, and in contrast, 6-month-olds rapidly learn location-related associations in occlusion based AEM paradigms (McMurray and Aslin, 2004; Kochukhova and Gredebäck, 2007). Based on these findings one might expect that infants' predictions would be related to the location of the goal.

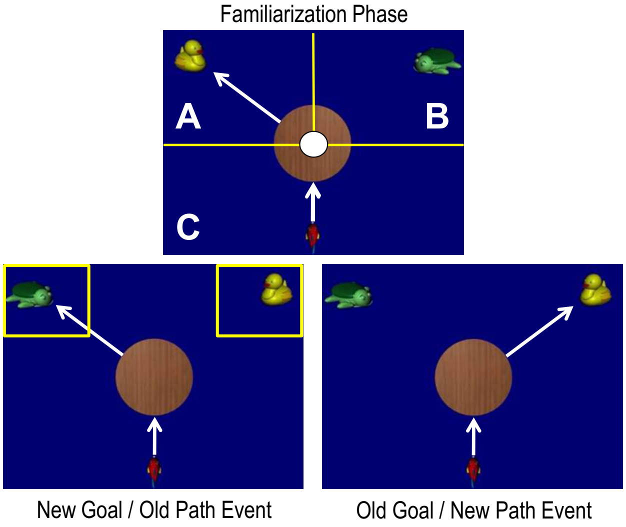

FIGURE 1 | An example of the stimulus presentation during familiarization trials (upper panel) and during the test trials (lower panels) in Experiment 1. The upper panel additionally indicates there appearance AOIs (not including the white area covered by the inner circle) for the calculation of predictive gaze. Reappearance AOI A indicates the area of a goal-related prediction. Reappearance AOI B indicates the area of a non-goal-related prediction. Additionally, start area C is the area where the gaze originated from previous to prediction. In the lower left panel, the target AOIs that were used to measure the looking time to the respective targets are depicted additionally.

METHOD

Participants

Participants were 9-month-olds ($n = 24$; 11 girls; $M = 9$ months; 5 days; 8.20–9.15). Nineteen additional infants were excluded due to fussiness (resulting in too few trials, $n = 16$) or experimenter errors ($n = 3$). Infants had to administer a sufficient number of trials for two dependent variables, looking times and predictive gaze shifts. The fact that only the very first test trial could be analyzed with respect to predictive gaze shifts (see also below) was one major cause for the high exclusion rate. Furthermore, the large number of infants excluded from analysis (44%) is not unusual for eye tracking studies (McMurray and Aslin, 2004) and does not reflect the average exclusion rate as this was much smaller in Experiment 2, see below. Infants were recruited from a database of parents who had agreed to participate in infant studies.

Apparatus

The laboratory was unfurnished except for the test equipment. The infants were seated in a car safety seat (Maxi Cosi Cabrio), which was placed in front of the eye tracker. The stimuli were presented, and gaze was measured using a Tobii 1750 near infrared eye tracker (Tobii AB, Stockholm, Sweden) with an infant add-on (precision: 1°, accuracy: 0.5°, sampling rate: 50 Hz). A nine-point infant calibration was used. During calibration, a blue and white sphere expanded and contracted (extended diameter = 3.3°) in synchrony with a sound. Viewing distance was approximately 60 cm, display size was 25° × 21°. For the measurement of the looking times, a camera was positioned above the monitor and recorded a close-up view of the infants, which was displayed on a control monitor. Looking times were measured online by two trained observers (to assess reliability).

Stimulus material

Stimulus material was generated using the software CINEMA 4D R10 and BodyPaint 3D (Maxon Computer GmbH, www.maxon.net). It consisted of an agent (a red colored fish with yellow tail, 2.9° × 1.0°), an occluder (wooden colored; radius = 6.7°), and two targets (a yellow duck, 3.4° × 2.9°) and a green turtle, 4.3° × 2.9°), all presented on a blue background, see **Figure 1**. The whole experiment consisted of eight familiarization trials, one intermediate trial where the positions of the targets were swapped (swap trials) and six test trials. The familiarization and the test trials consisted of the following sequence. First, the agent first jumped up and down three times accompanied by a sound to orient the infant to the screen (initial phase: 4000 ms). The agent then moved swimming-like with a wiggling tail toward the occluder and disappeared behind the occluder (pre-occlusion movement from movement onset until the agent completely disappeared behind the occluder: 2480 ms), the agent continued to move under the occluder (occlusion time: 920 ms), reappeared from behind the occluder and moved toward one of the targets upon reappearing (post-occlusion movement from the first frame of reappearance until the arrival at the goal: 3400 ms). Once at the goal object, the agent poked it three times while the goal object remained static (poking time: 2520 ms). The agent then remained motionless until the trial was terminated. Looking time measurement started when the agent touched the goal object until the infant had looked away for 2 or 60 s had elapsed, at which time the trial ended.

Prior to the test phase, infants were shown that the goal positions were swapped with no agent present. Subsequently, two different test events were presented three times each, in alternating order. In the *old goal/new path* event, the agent moved on a new path toward the old goal (i.e., constant goal identity, changed goal location). In the *new goal/old path* event, the agent moved on the old path toward a new goal (i.e., changed goal identity, constant goal location). Goal object, movement path, goal locations, and test event presented first were counterbalanced between subjects.

DATA ANALYSIS

Post-hoc measure – looking time

Analogous to Woodward (1998) and as described above, looking time toward the whole display was coded during all familiarization and test trials online from a control monitor by two trained observers who were unaware of the condition (inter-rater agreement was 83%).

Online measure – predictive gaze shifts

Two gaze measurements were calculated based on the previous study investigating predictive gaze shifts in an occlusion paradigm (Kochukhova and Gredebäck, 2007). For this, the area of the video presentation was divided into three further areas of interest (AOI; see **Figure 1**, upper panel). AOI A and B (reappearance AOIs) covered each 90° of the occluder edge. These areas extended both inside and outside the occluder, covering all but the final 2° near the occluder center and extending outwards to cover the entire amplitude of the agent's motion.

Gaze shifts were first categorized to be predictive or reactive. Predictive gaze shifts included all trials in which infants shifted their gaze across the occluder to target area A or B (see **Figure 1**) *before* the agent had been visible for 200 ms after occlusion. Reactive gaze shifts included all trials in which infants shifted their gaze across the occluder to target area A or B *after* the agent had been visible for 200 ms. This criterion was based on the average reactive saccadic latency to moving targets in adults (Engel et al., 1999) and infants (Gredebäck et al., 2006) and has been used in previous studies (e.g., Kochukhova and Gredebäck, 2007). Two primary scores, prediction rate and accuracy rate were calculated separately for each infant based on percentage scores.

Prediction rate. The prediction rate reports how often infants predicted the reappearance of the agent relative to the total number of attended trials. It is important to note that this measurement focuses on the timing of infants' gaze shift over the occluder and does not take into account at which location the infants predicted the agent to reappear.

Accuracy rate. Second, the accuracy rate reports where infants predicted the agent's reappearance; the number of predictions directed toward the target AOIs during familiarization was divided by the total number of gaze shifts (predictions and reactions) across the occluder. During familiarization, the *goal-related accuracy rate* (proportion of predictions toward the goal object) and the *non-goal-related accuracy rate* (proportion of predictions toward the other, non-goal object) were calculated. During test trials the *identity-related accuracy rate* (predictions being directed based on the identity of the goal during familiarization, i.e., toward the old goal on the new path) and the *location-related accuracy rate* (predictions being directed based on the location of the goal during familiarization, i.e., toward the new goal on the old path) were calculated.

Additional measures – specific looking times

Finally, the measurement of eye movement data allowed for a more detailed analysis of the position where infants were looking at what point in time during stimulus presentation. Accordingly, we were specifically interested in the proportion of looking time the infants spent looking at the areas of each of the two targets during the measurement of looking time. As for the looking times to the overall display, the measurement of these looking times was conducted after the agent had arrived at the respective goal during familiarization and test phase. In contrast to the measurement of the looking times to the overall display, this looking time measurement was calculated from the eye tracking data.

Additionally, to check whether the infants had seen the targets at their new locations, during the swap trial the looking time to the AOI around the two targets (target AOIs) was measured by counting the data points of infants gaze pattern that were located within this AOIs. The respective target AOIs are depicted in **Figure 1** (lower left panel) and covered 7.2° × 6.4° starting from the upper and left/right border of the stimulus display.

Inclusion criteria

With respect to looking time, all familiarization and test trials were analyzed. With respect to the calculation of prediction rate and accuracy rate, all familiarization trials and the first test trial were analyzed. Because the agent reappeared in all of the test trials, infants received feedback about the agent's behavior during the test trials immediately after the agent reappeared in the first test trial. For this reason, only the first test trial could be analyzed. In this first test trial, the infants had not yet received any feedback about the agent's behavior after the positions of the targets had been swapped.

To be included in the data analysis, infants had to provide valid data for at least four out of eight familiarization trials, the swap trial, and four out of six test trials including the first test trial. For the looking times, these inclusion criteria had to be met during the phase when the looking time was measured. For the gaze shifts, a trial was classified to be valid if infants had tracked the agent prior to the occlusion passage and if they fixated one of the two possible target AOIs before or after the agent reappeared.

RESULTS

In the results section we first report how many children and trials were included in the data analysis. Then, the data of the looking time as our post-hoc measure followed by the predictive gaze shifts as our online measure and then directly compare the two measures. This is followed by a more detailed analysis of the looking times toward different AOI.

Looking times were analyzed using parametric analyses of variance (ANOVA) and t-tests. The analyses of the prediction and accuracy rates were performed using non-parametric Wilcoxon Signed Ranks tests. The data level did not justify parametrical analyses because the categorization of the gaze shifts to be predictive or reactive resulted in nominal scaled data.

Inclusion rates

All infants provided *looking times* during the first two familiarization trials (*Fam12*) and the last two familiarization trials (*Fam78*). During test, one infant provided data only for one out of all six old goal/new path trials, all other infants provided looking time data for all six test trials.

With respect to the analysis of *gaze shifts*, the following numbers of trials were included in the final analysis (percentage scores reflect number of included trials relative to number of presented trials): *Total number of trials* equaled 206 (95.4%). *Number of Fam12 trials* equaled 47 (97.9%); *Number of Fam78 trials* equaled 45 (93.4%); *Number of Test1 trials* equaled 24 (100.0%). During Fam12, one infant provided data from only one trial; during Fam78, one infant provided no data, and one data from only one trial; during Test1, all infants provided data from the first test trial. For the parametric analyses, the data of the infant who provided no data during Fam78 (1 out of 24 participants) was replaced by the respective grand mean to keep data loss minimal.

Post-hoc measure – looking time

Looking time decreased from Fam12 ($M = 13.81$ s, SD = 5.72 s) to Fam78 ($M = 9.40$s, SD = 3.62s), $t(23) = 3.19$; $p < 0.01$. A 2 × 2 [Test Event (new goal/old path; old goal/new path) × Order (new goal/old path events first; old goal/new path event first)] ANOVA revealed a significant main effect of test event, the infants looked longer at new goal ($M = 9.57$ s, SD = 3.97 s) than old goal events ($M = 7.30$ s, SD = 3.15 s), $F(1, 22) = 6.77$, $p = 0.02$. There was neither a main effect of order nor an interaction of order and test, both Fs < 1. A non-parametric Sign test supported this result: 19 infants looked longer at new goal/old path events compared to five infants who looked longer at old goal/new path events, $p < 0.01$. As such, looking time results as a post-hoc measure replicate prior studies using the Woodward paradigm (Woodward, 1998; Luo and Baillargeon, 2005; Sommerville and Woodward, 2005) by demonstrating longer looking times when the agent moved along the same path as during familiarization to reach a new goal compared to when the agent moved on a new path to reach the same goal that was approached during familiarization.

Online measure – prediction rate

The infants performed predictive gaze shifts in roughly three-quarters of the trials, independent of the experimental phase. The prediction rate equaled 72.8% (SD = 22.2) over all familiarization trials and did neither change significantly from Fam12 ($M = 72.9$%, SD = 39.0) to Fam78 ($M = 71.7$%, SD = 35.6), $Z = 0.0$, $p = 1$, nor from Fam78 to Test1 ($M = 62.5$%, SD = 49.5), $Z = -0.36$, $p = 0.72$ (Wilcoxon test). These findings demonstrate that infants most often predicted that the fish will reappear from behind the occluder; this was true for both familiarization and test trials. In the following section we analyzed where the infants expected the agent to reappear.

Online measure – accuracy rate

Accuracy rates are presented in **Table 1**. During the familiarization, the goal-related accuracy rated was significantly higher than the non-goal-related accuracy rate during familiarization, $Z = -2.15$, $p = 0.03$. It did furthermore not change from Fam12 to Fam78, $Z = -0.24$, $p = 0.98$. The infants correctly predicted the agents' movement during familiarization.

As this analysis has shown that the infants had learned to correctly predict the reappearance of the agent during familiarization, the most relevant further analysis is to compare the (high) goal-related accuracy rate at the end of the familiarization phase (Fam78) to both the identity-related and the location-related accuracy rate in Test1. Comparing accuracy rates in the transition between familiarization (Fam78) and test trials (Test1) demonstrates that the goal-related accuracy rate was marginally higher than the identity-related accuracy rate in Test1, $Z = -1.58$, $p = 0.11$, and did not differ from the location-related accuracy rate in Test1, $Z = -0.56$, $p = 0.58$. These results indicate that infants performed less identity-related predictions in Test1 than goal-related predictions in Fam78, while the location-related predictions in Test1 did not differ from non-goal-related predictions in Fam78.

Comparison of looking time and eye movements

To directly compare looking time and predictive gaze, we compared the number of infants who looked longer at the new goal/old path events, thus did not expect a *change of identity* of the goal of the agent when measured post-hoc (19 out of 24, 79.2%) to the

Table 1 | Mean accuracy rates in % (and Standard Deviations) during the first and the last two familiarization trials (Fam 12, Fam78) and the test trials (first test trial in experiment 1 and first two test trials in experiment 2.

Experimental phase	Accuracy rate	Age						
		Experiment 1	Experiment 2					
		9 Months	All ages	9 Months	12 Months	24 Months	36 Months	Adults
Fam 12	Goal-related	47.9 (8.8)	40.3 (8.0)	47.8 (8.8)	39.6 (8.0)	29.2 (6.7)	54.2 (7.9)	30.4 (7.8)
	Non-goal-related	25.0 (6.7)	31.8(7.2)	30.4 (6.6)	37.5 (6.9)	27.1 (6.0)	31.3(7.9)	32.6(8.8)
Fam78	Goal-related	47.8 (8.2)	53.9 (8.4)	47.9 (9.3)	52.1 (8.2)	50.0(8.0)	62.2 (8.2)	54.2(8.5)
	Non-goal-related	23.9 (6.6)	18.9(6.3)	25.0 (7.4)	16.7(5.8)	12.5 (5.4)	19.6 (6.6)	20.8 (6.7)
Test	Identity-related	25.0 (9.0)	42.2 (8.5)	25.0 (7.4)	27.1 (8.0)	41.7(8.9)	58.3 (7.8)	58.7(8.3)
	Location-related	37.5(10.1)	34.3 (8.5)	47.9 (8.8)	41.7(9.2)	39.6 (9.0)	27.1 (7.4)	15.2(5.6)

number of infants showing *identity-related predictions* in Test1 (6 out of 24, 25%) using a Chi-square test, $\chi^2(1, N = 48) = 14.1$, $p < 0.001$. The number of infants who performed identity-related predictions during Test1 was much smaller than the number of infants who performed identity-related looking times. Interestingly, all six infants who performed identity-related predictions the first test trial showed respective identity-related looking times and looked longer in the new goal/old path test events.

Although this result has to be interpreted with great care, as the number of infants per cell is very small, it indicates that those infants, who show identity-related processing of the agent when measured online, do so as well when measured post-hoc. In contrast, the reverse is not true, infants who show identity-related processing when measured post-hoc, do not necessarily show identity-related processing when measured online.

Additional measures – proportion of looking times toward different AOIs

Finally, to look more closely at the infants looking during the different experimental phases, we calculated the proportion of time infants spent looking toward the two target AOIs (looking time toward respective AOI divided by the total looking time as respectively measured by the eye tracker). During familiarization, infants looked longer at the goal object (proportion of looking time: $M = 44.5\%$, $SD = 13.9$) than at the non-goal object ($M = 15.2\%$, $SD = 6.0$), $t(23) = 8.71$, $p < 0.001$. This looking behavior did not change from Fam12 (goal object: $M = 45.2\%$, $SD = 18.2$, non-goal object: $M = 19.6\%$, $SD = 9.2$) to Fam78 (goal object: $M = 41.9\%$, $SD = 16.7$, non-goal object: $M = 11.5\%$, $SD = 6.9$) indicated by a main effect of target (goal object vs. non-goal object), $F(1, 23) = 58.73$, $p < 0.001$ and no interaction with phase (Fam12 vs. Fam78), $F < 1$. A main effect of phase, $F(1, 23) = 11.59$, $p = 0.002$, indicates that the proportion of looking toward the two objects decreased from Fam12 to Fam78. This result indicates that during the familiarization phase, the infants primarily looked at the goal object where the agent was.

During the swap trial the infants looked equally long at the goal object ($M = 24.2\%$, $SD = 14.4$) and the non-goal object ($M = 31.7\%$, $SD = 16.8$), $t(23) = 1.31$, $p = 0.20$.

Finally, the looking proportions during the test events were analyzed using a 2 × 2 × 2 [Test Event (old goal/new path vs. new goal/old path) × Target (old goal vs. new goal) × Order (old goal/new path event presented first vs. new goal/old path event presented first)] repeated measures ANOVA with test event and target as within-subjects factors and order as between factor. This analysis only yielded a significant Test Event × Target interaction, $F(1, 22) = 51.59$, $p < 0.001$. During old goal test events, the infants looked longer at the old goal ($M = 38.8\%$, $SD = 18.9$) than at the other (new) goal ($M = 10.6\%$, $SD = 7.0$). During the new goal test events, the infants' looking behavior was reversed; they looked longer at the new goal ($M = 36.3\%$, $SD = 17.8$) than at the other (old) goal ($M = 12.9\%$, $SD = 9.1$). As during the familiarization trials, in the test trials, the infants primarily looked at the goal object where the agent was located. Furthermore, the infants looked at both objects during the swap trial and importantly, they looked equally long at the two targets during this trial. This indicates that the infants had observed that the positions of the two targets had been swapped.

DISCUSSION

The looking time results of Experiment 1 replicate previous findings (Woodward, 1998). When measured post-hoc, the infants looked longer when the agent moved on the old path toward a new goal compared to trials where the agent moved on a new path toward the old goal. Following the logic of Woodward (1998) infants built a representation about the agent's goal during familiarization. During test, when the positions of the targets had been swapped, the infants' expectation that the agent continues to move toward the old goal was met in the old goal/new path condition and was violated in the new goal/old path condition, resulting in extended looking times in the latter condition. In line with previous findings (Luo and Baillargeon, 2005; Csibra, 2008), this suggests that at the age of 9 months, infants are able to interpret a non-human agent's behavior as goal-directed when expectations are measured post-hoc.

Interestingly, the infants' predictive eye movements did not reflect their looking times. The results showed that the infants learned to correctly predict the reappearance of the agent during familiarization. However, during the first test trial infants showed a tendency to base their predictions on the location of the goal object as observed during familiarization.

A more detailed analysis of the looking times revealed that the infants predominantly looked toward the target that the agent was close to during both familiarization and test events. This finding has potential implications on the validity of the looking time task that will further be discussed in the Section "General Discussion."

An important issue to be raised at this point is the fact that the present design does not allow for a distinction between goal-anticipations and path-anticipations. There are two reasons why we did not differentiate between these two measures. First, due to the restriction to only one test trial given by the paradigm, dividing the infants' eye movements into goal- and path-directed gaze shifts resulted in a small number of test trials. Accordingly, the validity of such a measure would be limited. Second, more theoretically grounded, in a goal-directed action are goal and path mutually related. Anticipating the path an agent includes – at least in the present paradigm – the consideration of the goal the agent has. And vice versa, anticipating the goal includes the consideration of the path the agent takes. We did not differentiate between the two forms of anticipation.

To sum up, the main goal of the present Experiment1 was to compare an online with a post-hoc measure for infants' action perception and to test whether infants base their predictions of the goal of an observed action based on the identity of the goal as suggested by previous findings using post hoc looking time measures (first hypothesis; e.g., Woodward, 1998), or on the location of the goal (second hypothesis; e.g., McMurray and Aslin, 2004; Addyman and Mareschal, 2010; Albareda-Castellot et al., 2011; Paulus et al., 2011b) as observed during familiarization. There is no definite answer to this question. The present results point toward a dissociation between looking time and predictive gaze and allow therefore a rejection of the first hypothesis. It is, however, less clear, what the basis of the infants' prediction was as the results are ambiguous with respect to the two hypotheses. Furthermore, our conclusions are based on the performance of a few infants providing only one data point during the test trials. This might question the validity of the present data. For these reasons, we modified the paradigm used in Experiment 1 in order to replicated and extend the findings of Experiment 1 and to further explore the development of this potential dissociation.

EXPERIMENT 2

In Experiment 2, we modified the paradigm from Experiment 1 to make it (a) a more prediction-oriented eye tracking paradigm and (b) to further strengthen processes of goal attribution by adding, for example, action effects. Additionally, we included a wider age range to investigate the developmental trajectory of responses when expectations are measured online.

METHOD

Participants

We tested 9-month-olds ($n = 24$; 7 girls; $M = 9.3$; 8.17–9.13), 12-month-olds ($n = 24$; 9 girls; $M = 12.5$; 11.17–12.15), 24-month-olds ($n = 24$; 14 girls; $M = 24.2$; 23.15–24.14), 36-month-olds ($n = 24$; 8 girls; $M = 36.8$; 34.23–37.6), and adults ($n = 24$; 13 female; $M = 24$ years; 19–34 years). Additionally, fourteen 9-month-olds, six 12-month-olds, nine 24-month-olds, and two 36-month-olds were excluded from analysis due to fussiness or procedural errors.

Stimuli, apparatus, procedure, and data analysis

Stimulus material and procedure were adapted from Experiment 1 with the following modifications. We were concerned that the infants' predictions in Experiment 1 were biased by the long inter-trial periods. These were caused by the measurements of the infants' looking times resulting in periods up to 60 s depending on the infants' looking behavior. In order to present trials in a higher frequency we shortened the trials and did no longer measure infants' looking time.

Another reason why the infants did not predict the reappearance of the agent based on goal identity might have been that the stimulus presentation did not trigger goal attribution processes strong enough. For this reason, we strengthened these goal attribution processes by applying the following modifications: First, the targets were more distinct. In Experiment 1, both targets were animals; the infants might, thus, have processed both targets in terms of one category (animal) instead of two distinct targets (duck and turtle). We now followed more closely the targets as used by Woodward (1998) and replaced the turtle by an inanimate ball. Second, the agent's poking of the goal object now caused a salient effect (during the poking, the goal object moved up and down while making a laughing sound). Previous research has shown that adding and effect to an unfamiliar action helps 6-month-olds to interpret the respective action as object-directed (Hofer et al., 2007; Jovanovic et al., 2007).

Finally, in order to be able to analyze more test trials, no more feedback was provided during the test events; the agent never reappeared from behind the occluder. This allowed us to repeat the test trial, thereby gaining additional data that will reduce noise and provide a more solid assessment of individual infants' and children's prediction and accuracy rate.

Goal identity, movement path, and goal locations were counterbalanced between participants. Accuracy and prediction rates were calculated as in Experiment 1. In addition, movement times were the same as in Experiment 1, the duration of the swap trial was 15 s, the test trials were presented for 10 s after the agent disappeared behind the occluder. Participants of all age groups were only told to watch the movies closely, without further instructions.

RESULTS

In the results section we first report how many children and trials were included in the data analysis, then, prediction and accuracy rates are reported. Finally, the individual age groups are analyzed separately. As in Experiment 1, the data level did not justify parametrical testing, accordingly, the non-parametric analyses were performed using Kruskal–Wallis tests and Wilcoxon Signed Ranks tests.

Inclusion rates

The following numbers of trials were included in the final data analysis. The *total number of trials* (maximum: 240 trials; including 192 familiarization and 48 test trials) equaled 229 (95.4%) for the 9-month-olds, 223 (92.9%) for the 12-month-olds, 227 (94.6%) for the 24-month-olds, 228 (95.0%) for the 36-month-olds, and 232 (96.7%) for the adults. *Number of Fam12 trials* (maximum: 48 trials) equaled 46 (95.8%) for the 9-month-olds, the 12-month-olds, the 24-month-olds, and the adults, each, and equaled 47 (97.9%) for the 36-month-olds. One 9-month-old, and one adult provided no data during Fam12, two 12-month-olds, two 24-month-olds, and one 36-month-old provided only one trial, all other participants provided data in both trials. *Number of Fam78 trials* (maximum: 48 trials) equaled 42 (87.5%) for the 9-month-olds, 44 (91.7%) for the 12-month-olds, 46 (95.8%) for the 24-month-olds, 42 (87.5%) for the 36-month-olds, and 47 (97.9%) for the adults. One 36-month-old provided no data during, six 9-month-olds, four 12-month-olds, two 24-month-olds, four 36-month-olds, and one adult provided only one trial, all other participants provided data for both trials. *Number of Test12 trials* (maximum: 48 trials) equaled 48 (100.0%) for the 9-month-olds, 42 (87.5%) for the 12-month-olds, 43 (89.6%) for the 24-month-olds, 46 (95.8%) for the 36-month-olds, and 44 (91.7%) for the adults. Six 12-month-olds, five 24-month-olds two 36-month-old, and four adults provided only one trial, all other participants provided data in both trials.

Online measure – prediction rate

Prediction rate equaled 74.7% (SD = 25.8) over all familiarization trials and age groups. Kruskal–Wallis tests revealed no differences between the age groups for the prediction rates over all familiarization trials [$\chi^2(4, N = 120) = 7.07$, $p = 0.13$], as well as for Fam12 [$\chi^2(4, N = 120) = 9.03$, $p = 0.06$], Fam78 [$\chi^2(4, N = 120) = 4.16$, $p = 0.39$], and Test12 [$\chi^2(4, N = 120) = 2.58$, $p = 0.63$]. Over all age groups, prediction rate did not change from Fam78 trials ($M = 72.7\%$, SD = 39.5) to Test12 trials ($M = 76.5\%$, SD = 36.1), $Z = -1.02$, $p = 0.31$ (Wilcoxon test). The prediction rate in Experiment 2 was comparable for the different age groups and for the different experimental phases. Similar to Experiment 1, the participants predicted in almost three-quarter of the trials that the fish will reappear from behind the occluder, both during familiarization and test trials.

Online measure – accuracy rate

Familiarization phase. We first checked whether the accuracy rates during the familiarization phase changed over age using Kruskal–Wallis tests. This was not the case, neither for the goal-related accuracy rate, $\chi^2(4, N = 120) = 5.40$, $p = 0.25$, nor for the non-goal-related accuracy rate, $\chi^2(4, N = 120) = 2.65$, $p = 0.62$, see **Figure 2**. Accordingly, to test whether the two accuracy rates differed from each other during familiarization, data was collapsed across age groups. As in Experiment 1, during familiarization, the goal-related accuracy rate was higher than the non-goal-related accuracy rate, $Z = -5.66$, $p < 0.001$ (Wilcoxon test).

Test phase. The same analyses were performed for the accuracy rates during the test phase. As can be seen in **Figure 2**, the identity-related accuracy rate increased over age, $\chi^2(4, N = 120) = 15.37$, $p = 0.004$, and marginally, the location-related accuracy rate decreased over age, $\chi^2(4, N = 120) = 8.28$, $p = 0.08$ (both Kruskal–Wallis tests). Accordingly, the accuracy rates of the different age groups were analyzed separately and are reported in further detail in the following section.

Analysis of individual age groups

As in Experiment 1, participants of all age groups had learned to correctly predict the reappearance of the agent during familiarization. In the next step, we compared the (high) goal-related accuracy rate at the end of the familiarization phase (Fam78) to both the identity-related and the location-related accuracy rate in Test12. The results indicate that the 9-month-olds showed a marginally significant change when the goal-related accuracy rate was compared to the identity-related accuracy rate, $Z = -1.69$, $p = 0.09$, but no significant change when it was compared to location-related accuracy rate: $Z = -0.04$, $p = 0.97$. The same pattern was found in the 12-month-olds, who significantly changed their looking behavior from Fam78 to Test12, indicated by a significant decrease of the identity-related accuracy rate, $Z = -2.55$, $p = 0.01$, but no change to the location-related accuracy rate, $Z = -0.96$, $p = 0.37$. In the 24-month-olds, looking behavior did neither concerning the identity-related accuracy rate, $Z = 0.72$, $p = 0.42$, nor the location-related accuracy rate, $Z = -0.88$, $p = 0.38$. In the 36-month-olds, no change of looking behavior was observed to the identity-related accuracy rate, $Z = -0.59$, $p = 0.55$, but here, a significant change to the location-related accuracy rate was found, $Z = -2.85$, $p = 0.004$. The same pattern was found in the adults where no change to the identity-related accuracy rate was observed, $Z = -0.30$, $p = 0.77$, but again a significant change to location-related accuracy rate, $Z = -3.36$, $p = 0.001$.

These changes from familiarization to test phase indicate, that the youngest two age groups continued to predict the reappearance of the agent on the basis of the previously observed location of the goal, while the oldest two age groups continued to predict the reappearance of the agent on the basis of the previously observed identity of the goal.

Looking times during swap trials

Additionally, we calculated the proportion of looking time toward the AOIs of the two targets and analyzed them by means of a 2 × 5 [Target (goal, non-goal during familiarization) × Age (9 months, 12 months, 24 months, 36 months, adults)] ANOVA that showed that the infants looked equally long at both objects (goal: $M = 28.1$, SD = 16.1; non-goal: $M = 27.9$, SD = 17.0), $F < 1$, and that the looking time decreased over age, $F(4, 115) = 4.13$, $p = 0.004$. This age effect is based on the shorter looking times of the adults ($M = 20.4$, SD = 12.8) compared to all other age groups (9 months: $M = 30.3$, SD = 17.6; 12 months: $M = 28.5$, SD = 17.4; 24 months: $M = 31.7$, SD = 17.6; 36 months: $M = 29.2$, SD = 15.0), indicated by LSD-corrected post-hoc tests, all ps < 0.01. No other differences between the age groups were significant. The interaction of the two factors was not significant. To ensure that the participants did look at the objects one sample t-tests against zero were performed for the looking proportion toward the two objects separately for each age group that were all significant, all ps < 0.001. All participants thus looked at the object and they looked equally long at both objects during the swap trials.

DISCUSSION

Experiment 2 showed that the 36-month-olds and adults predicted the reappearance of the agent in the test trials based on the identity

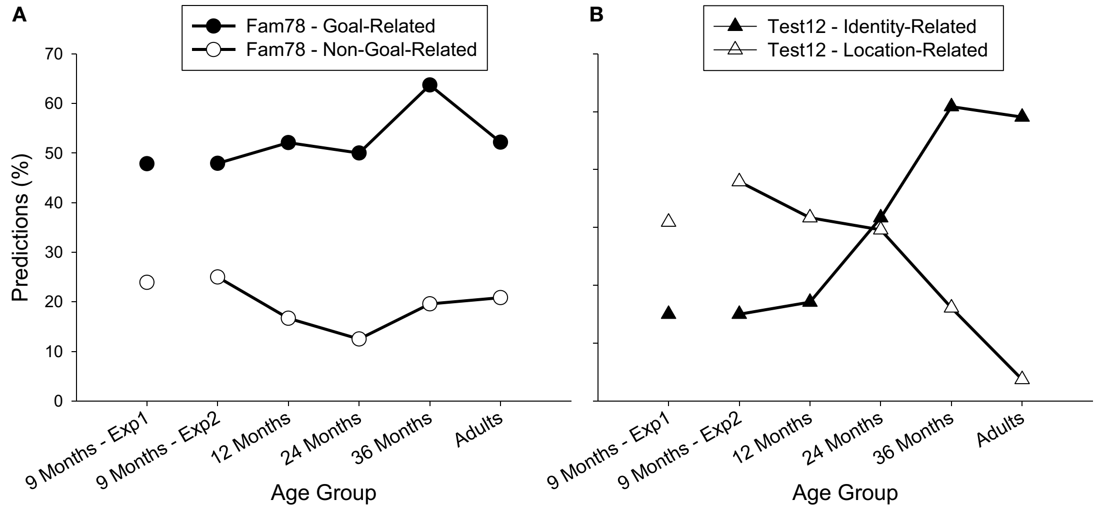

FIGURE 2 | Development of predictive gaze during the last two familiarization trials (A) and the test trials (B) across age groups in Experiment 1 and 2. (A): filled circles represent the goal-related accuracy rate; empty circles the non-goal-related accuracy rate. **(B)**: filled diamonds represent identity-related accuracy rate and empty diamonds the location-related accuracy rate.

of the goal of the observed action. In contrast, the 12-month-olds, and less clearly also the 9-month-olds based their predictions in the test trials on the location of the goal during the familiarization phase. The latter finding replicates the results of Experiment 1 indicating a dissociation between looking time as a post-hoc measure and predictive gaze as an online measure. Although infants do encode the identity of the action goal already at the age of 9 months when measured post-hoc, they base their predictions – though less clearly – on the location of the action goal. It is not before the age of 36 months, that children integrate goal identity in their predictions.

The ambiguous findings of the 9-month-olds might be explained by the fact that at this age, infants' capacity to predict action goals is in a developing phase. Recent studies have shown that infants start to predict action goals at 6 months of age (Hunnius and Bekkering, 2010; Kochukhova and Gredebäck, 2010) and that this capacity continues to develop over the following months of life (Kenward, 2010; Paulus et al., 2011b), however reliable predictive gaze shifts are often not found before the age of 12 months (Falck-Ytter et al., 2006; Melzer et al., 2012) or even older (Gredebäck et al., 2009). In the present study, there was no difference in the overall rate of predictions between the two younger age groups, already the 9-month-olds showed predictions in more then 70% of the trials, so the 9-month-olds were principally able to predict the agents' action. However, during the test phase, the 12-month-olds showed a clearer pattern of where their predictions were directed toward. The data finally shows that the 24-month-olds seem to be in a transition period, as their predictions were ambiguous. The looking times during the swap trials ensured that participants of all age groups have observed that the position of the targets has changed from the familiarization to the test phase. Potential causes and implications of this developmental trajectory from a dissociation between looking times and predictive gaze early in life to an association of the two measures later in life will be discussed below.

GENERAL DISCUSSION

In the present study, we compared post-hoc measures of children's expectations about an observed goal-directed action with online measures concerning their predictions about the same action by combining a looking time paradigm with a predictive gaze paradigm. The looking time results from Experiment 1 replicated Woodward's (1998) original findings. Nine-month-olds were shown to be sensitive to the identity of the goal of an observed action when measured post-hoc.

The results of the analysis of the infants' eye movements contrast the looking time results and showed at the age of 12 months (and less reliably at the age of 9 months) infants predicted the reappearance of the agent based on the location of the goal during an observed action and that it was not until the age of 3, that this dissociation disappeared and that children predicted the reappearance of the agent after occlusion based on goal identity. These findings indicate that post-hoc measures and online measures used to investigate children's action expectations are dissociated early in life. They further support the second hypothesis put forward in the introduction of Experiment 1 that early in life, infants continue to anticipate the reappearance of the agent in the test trials location-related.

This finding is, on the one hand, to some extent surprising as infants did encode the identity of the goal already at 9 months when their expectations were measured post-hoc, but they were not (yet) able to transfer this knowledge into their predictions. On the other hand, this findings is not that surprising as we know from previous findings that infants do take goal locations into account when predicting the future behavior of an agent (McMurray and Aslin, 2004; Addyman and Mareschal, 2010) and in the present paradigm, the infants were not only familiarized with the identity of the goal object but likewise with its location. The question remains, why this behavior changes with increasing age.

FROM DISSOCIATION TO INTEGRATION

One answer to this question is that the computational processes that are involved in processing observed actions are dissociated early in life and become associated later: Early in life, action expectations measured online seem to be organized around goal locations whereas action expectations measured post-hoc around goal identities. With increasing age, children then generally organize their action expectations primarily around goal identities. The nature of this dissociation can either be interpreted as temporal or procedural.

A temporal interpretation implies that the dissociation between the two measures reflects two successive states on the processing timeline of one common underlying mechanism. This mechanism would act location-conservatively in an early processing phase during the observation of an action, and identity-conservatively in a later processing phase upon completion of the action. Action expectations measured post-hoc and online thus rely on a different amount of information available. During development, the sensitivity to action goals shifts backward on the processing timeline. Early in life, infants can derive goals only through post-hoc comparison of their expectations with an observation, with sufficient information and processing time available. Only later in life can they already derive goals more quickly online, during the observation of an ongoing action. This temporal interpretation is further supported by the findings that 6-month-olds, who were not yet able to anticipate the goal of a feeding action when measured online did differentiate between rational and non-rational feeding actions when measured post-hoc using pupil dilation where the processing time was less constraint (Gredebäck and Melinder, 2010). However, recent findings from a manual search task are not consistent with this temporal interpretation. When 2-year-old children were provided with additional time to process an observed event and to plan a response to search for a hidden object, performance did not improve (Mash et al., 2006).

In contrast, according to a procedural interpretation, the dissociation between the two measures reflects a dissociation between two different mechanisms involved, one for processing goal location and another for goal identity. These two mechanisms could be separate early in life and only later become integrated under the lead of the identity-related mechanism. This assumption is reminiscent of the notion of the two visual pathways (Mishkin and Ungerleider, 1982; Goodale and Milner, 1992) as a possible underlying mechanism. The *ventral* (*what*) pathway is associated with the processing of goal identity. The *dorsal* (*where*/*how*) pathway provides online spatial control of movements required for action execution and mediates the processing of goal locations. Both pathways are connected to the frontal eye field (Schall, 2002) that is involved in visual processing and inhibitory control (Schall et al., 2002; Muggleton et al., 2010). Evidence from animals (Schroeder et al., 1998; Chen et al., 2007) and humans (Rao et al., 2003) further suggests a dorsal-over-ventral advantage showing faster processing of location-related compared to identity-related information. These findings are mirrored by the reported developmental pattern that in the first year of life, visual processing seems to be driven primarily by the dorsal pathway (Leslie et al., 1998; Mareschal et al., 1999). Mareschal and Johnson (2003) further showed that young infants have difficulty in integrating information coming from the two streams. The authors suggest that it is the affordance of a target that determines which of the two representations is maintained. When the targets were non-manipulable objects like faces or asterisks, the infants primarily responded to changes in identity like color and not to changes in location. In contrast, when the targets presented were manipulable toys, infants primarily responded to changes in location and not in identity. In the present study we presented manipulable toys as targets. Accordingly, the young infants might have primarily responded to the location of the object, processed by the dorsal stream, when their expectations were measured online and could not integrate this with the information about the object identity, processed by the ventral stream. Mareschal et al. (1999) further suggest that a dissociation or a developmental lag only occurs when it is necessary to integrate two sources of potentially conflicting information, about location and identity. As they say "This explanation predicts that tasks requiring access to only one imprecise source of information or tasks that are performed with a visible object will not result in a developmental lag. In contrast, any task that calls for the integration of cortically separable representations will fail unless performed with a visible object or with precise cortical representations." (p. 307). The advantage of the dorsal over the ventral stream found in the predictions of the younger children might therefore be based on the nature of the objects that were used as targets. The interpretation of Mareschal et al. (1999) are based on findings from the non-social domain and it remains a matter of further research to test whether the can be generalized to a social domain that includes animate cues as in the current paradigm and to test whether the dissociation found here can be modulated when non-manipulable objects are used as targets. It is, however, important to mention at this point, that in the present study, the objects used are the same for the two measures. Still, the infants do process the identity of the goal object when measured post-hoc but do only at a later age when measured online. The dissociation between the infants' looking times and their predictions can, thus, not be traced back solely to the manipulability of the targets.

Both the temporal and the procedural interpretation are not necessarily mutually exclusive. The above-mentioned differentiation between the mechanism acting location- vs. identity-conservatively in the temporal interpretation of course entails a procedural element, as well as the as the procedural interpretation includes a temporal element, such as the differences in processing speed. The emphasis of the respective interpretations, however, lies on the processing mechanism in the procedural interpretation and on differences in the processing time in the temporal interpretation.

Within this context, it is finally important to emphasize the functionality of this early dissociation. When predicting action goals, the perception-action system has limited time to make accurate estimations of future events. Focusing on location (rather than identity) might be a useful "heuristic" often providing accurate and fast estimations. This does not mean that infants are ignorant of goal identity. With sufficient information and time, infants adjust their behavior according to the configuration of goal locations and identities.

LOOKING TIMES TOWARD SPECIFIC AREAS OF INTEREST

Another important aspect of the present study that needs to be discussed is the finding of Experiment 1 that the 9-month-olds primarily looked at the goal object where the agent was both during familiarization and test phase. The interpretation of this finding can take two major routes. First, a radical interpretation would imply that data from studies using looking time as a post-hoc measure for infants' action expectations and their perception of goal-directed actions might be overly generous about infants' knowledge. As mentioned earlier, looking time tasks measure infants' expectations with fairly low spatial and temporal resolution (Aslin, 2007) and the results can easily be biased by low-level factors (see also Csibra, 2003). In the same line, it has been argued, that associative learning processes might subserve to a substantial part findings on early (social) cognition competences (Perner and Ruffman, 2005; Paulus et al., 2011a). For the following reasons, the present data might as well call into question the validity of the looking time tasks on action perception in general. During the familiarization phase, the infants spent more time looking at the familiarized (old) goal object than at the new goal object. This leads to an increase of the relative novelty of the new goal object in new goal/old location trials compared to old goal object in the old goal/new location trials and one cannot rule out that this relative novelty of surface features solely accounts for the increase in looking times in the respective test trials and that such a low-level explanation not only holds for the present data but for all data coming from studies using the same paradigm. However, there is evidence against such a low-level only explanation for the paradigm in general coming from previous looking time studies. These studies showed that at the same age, when infants do differentiate between the old goal/new path and new/goal/old path events when a familiar (grasping) action is performed by human agent, they do not show the same looking time pattern when a human agent performs an unfamiliar action (consisting in dropping the back of the hand on the object; Woodward, 1999; Hofer et al., 2005) or the human agent is replaced by an non-human agent performing the same action (e.g., mechanical claws, rods, occluders; Woodward, 1998). However, while the low-level factors of these studies can be assumed to be identical, the infants' looking times were not.

We favor an alternative interpretation of our data suggesting a developmental trajectory with the identity-related action expectation measured post-hoc at an early age being a precursor of the identity-related action expectation measured online at a later age. This interpretation is supported by the finding that all 9-month-olds who performed identity-related predictions in Experiment 1 showed respective identity-related looking times (i.e., longer looking to changes in goal identity compared to changes in goal location). In contrast, the reversed pattern could not be found; infants who showed identity-related looking times did not necessarily perform identity-related predictions. Similar developmental trajectories including dissociations between post-hoc measures and online measures early in life have been reported in tasks testing infants' knowledge about physical events. Expectations were measured online via manual search tasks where the children did not receive any feedback about the outcome of an observed event. Piaget, for example, has shown that infants do not manually search for hidden objects until they reach the age of 7.5–9 months. He concluded that it is not until this age that infants understand that hidden objects continue to exist (Piaget, 1952, 1954). In contrast, using post-hoc measures it has been shown that infants as young as 2.5 months do have some understanding about the continuity of hidden objects (Baillargeon et al., 1985; Wilcox et al., 1996). Similarly, when infants' knowledge about physical solidity was assessed, infants differentiated between expected and unexpected events already at the age of 4 months when their knowledge was measured post-hoc (Spelke et al., 1992). In contrast, when tested in a manual search task, toddlers at the age of 2.5 years still failed when they have to predict the position of an object behind a barrier (Berthier et al., 2000). Further studies comparing toddlers' knowledge about physical solidity in a within-subjects design showed that while toddlers failed to search at the correct location, they looked longer at an unexpected compared to an expected outcome of the same task (Hood et al., 2003; Mash et al., 2006). Keen (2003) concludes from these results that the perception of unexpected event outcomes seems to be a fundament upon which further knowledge about the world can be built. However, having knowledge (as assed via post-hoc measures) seems to be substantially different from being able to use that knowledge (as assessed via predictive gaze shifts or via manual search actions). Predictive gaze shifts, similar to manual search actions, not only require the evaluation of whether an observed event makes sense or not, the require an active – although not necessarily conscious – decision of where to shift gaze, a measure of the infants' expectation before the outcome of an event is perceivable. This requires the consideration of multiple potential outcomes and the selection of the most appropriate one. Infants are able to infer the outcome of an uncompleted event, when their looking time is measured (Daum et al., 2008, 2009), or when they have to imitate previously observed incomplete actions (Meltzoff, 1995; Hamlin and Woodward, 2005). Accordingly, the conclusion by Keen (2003) fits as well for the present findings that the post-hoc evaluation of a task as measured via looking times builds the basis and is a prerequisite on which the online processing can be built upon.

The slope of these developmental trajectories might very well vary between different domains. Here we presented an animated object and found that looking time and predictive gaze were dissociated over several years. In their recent study, Cannon and Woodward (2012) report earlier identity-related predictions when infants were presented with a grasping hand instead of an animated agent. The results of this study showed that infants at the age of 11 months were able to predict the goal of the grasping hand based on its identity during familiarization. One might interpret this finding as evidence for an earlier understanding of human actions compared to non-human actions, the fact that children at this age correctly predict the goal of a grasping action might, however, also be caused by the fact that the infants were less constraint in the timing of their predictions. The infants did not have to take into account the precise timing of the grasping action once it stopped but could shift their gaze to one of the two objects after the hand stopped. In contrast, in the present study, both spatial and temporal aspects were needed to be integrated very precisely in order to correctly predict the agent's behavior. This was not necessarily the case in the above-mentioned study. Furthermore, in the study by Cannon and Woodward (2012) the hand was never occluded during the presentation, while in the present study, in order to measure temporal and spatial aspects of prediction, the agent was occluded for a certain amount of time, what might have increased the difficulty of the task. Anyway, the prediction of action goals when measured online does still occur later, at the age of 11 months, than when measured post hoc (Woodward, 1998; Luo and Baillargeon, 2005; Luo, 2011). Further research will clarify how to what extent human and non-human actions are processed differently and on which basis predictions are made.

Finally, in the present study, we used a paradigm combining two visual measures of children's action perception. The dissociation found is potentially not restricted to these two measures, quite the contrary, as previous studies have shown, other measures such as looking time and manual grasping (Hood et al., 2003; Mash et al., 2003) or, dilation and predictive gaze (Gredebäck and Melinder, 2010) show comparable findings. This suggests that it is likely that for almost all measures of infant behavior, dissociations will be found as long as they are based on different temporal constraints, different amounts of information and, accordingly, tap potentially different underlying processing mechanisms.

FINAL CONCLUSION

In the present study, we explored the relation of two different measures used to investigate infants' expectations about goal-directed actions. We compared post-hoc measures of infants' expectations (via looking time) with online measures concerning their prediction (via predictive gaze). The looking times reflected identity-related expectations already at the age of 9 months. In contrast, predictive gaze pattern show that at a young age infants base their predictions primarily on the location of a goal object while it is only after the third birthday that predictive gaze reflects identity-related expectations as well.

ACKNOWLEDGMENTS

We would like to thank the parents and their children who participated in this study. We also wish to thank Caterina Boettcher for the acquisition of the infants, and our student research assistants for help in data collection. The study was supported by the Swedish Research Council (2011-1528).

REFERENCES

Addyman, C., and Mareschal, D. (2010). The perceptual origins of the abstract same/different concept in human infants. *Anim. Cogn.* 13, 817–833.

Albareda-Castellot, B., Pons, F., and Sebastián-Gallés, N. (2011). The acquisition of phonetic categories in bilingual infants: new data from an anticipatory eye movement paradigm. *Dev. Sci.* 14, 395–401.

Aslin, R. N. (2007). What's in a look? *Dev. Sci.* 10, 48–53.

Baillargeon, R., Spelke, E. S., and Wasserman, S. (1985). Object permanence in five-month-old infants. *Cognition* 20, 191–208.

Berthier, N. E., DeBlois, S., Poirier, C., Novak, M., and Clifton, R. (2000). Where's the ball? Two- and three-year-olds reason about unseen events. *Dev. Psychol.* 36, 394–401.

Biro, S., and Leslie, A. M. (2007). Infants' perception of goal-directed actions: development through cue-based bootstrapping. *Dev. Sci.* 10, 379–398.

Brandone, A. C., and Wellman, H. M. (2009). You can't always get what you want: infants understand failed goal-directed actions. *Psychol. Sci.* 20, 85–91.

Cannon, E. N., and Woodward, A. L. (2012). Infants generate goal-based action predictions. *Dev. Sci.* 15, 292–298.

Cannon, E. N., Woodward, A. L., Gredebäck, G., von Hofsten, C., and Turek, C. (2012). Action production influences 12-month-old infants' attention to others' actions. *Dev. Sci.* 15, 35–42.

Chen, C. M., Lakatos, P., Shah, A. S., Mehta, A. D., Givre, S. J., Javitt, D. C., and Schroeder, C. E. (2007). Functional anatomy and interaction of fast and slow visual pathways in macaque monkeys. *Cereb. Cortex* 17, 1561–1569.

Clifton, R. K., Muir, D. W., Ashmead, D. H., and Clarkson, M. G. (1993). Is visually guided reaching in early infancy a myth? *Child Dev.* 64, 1099–1110.

Csibra, G. (2003). Teleological and referential understanding of action in infancy. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 358, 447–458.

Csibra, G. (2008). Goal attribution to inanimate agents by 6.5-month-old infants. *Cognition* 107, 705–717.

Csibra, G., Biro, S., Koos, O., and Gergely, G. (2003). One-year-old infants use teleological representations of actions productively. *Cogn. Sci.* 27, 111–133.

Daum, M. M., Prinz, W., and Aschersleben, G. (2008). Encoding the goal of an object-directed but uncompleted reaching action in 6- and 9-month-old infants. *Dev. Sci.* 11, 607–619.

Daum, M. M., Vuori, M. T., Prinz, W., and Aschersleben, G. (2009). Inferring the size of a goal object from an actor's grasping movement in 6- and 9-month-old infants. *Dev. Sci.* 12, 854–862.

Engel, K. C., Anderson, J. H., and Soechting, J. F. (1999). Oculomotor tracking in two dimensions. *J. Neurophysiol.* 81, 1597–1602.

Falck-Ytter, T., Gredebäck, G., and von Hofsten, C. (2006). Infants predict other people's action goals. *Nat. Neurosci.* 9, 878–879.

Flanagan, J. R., and Johansson, R. S. (2003). Action plans used in action observation. *Nature* 424, 769–771.

Gergely, G., Nadasdy, Z., Csibra, G., and Biro, S. (1995). Taking the intentional stance at 12 months of age. *Cognition* 56, 165–193.

Goodale, M. A., and Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends Neurosci.* 15, 20–25.

Gredebäck, G., Johnson, S., and von Hofsten, C. (2010). Eye tracking in infancy research. *Dev. Neuropsychol.* 35, 1–19.

Gredebäck, G., and Melinder, A. (2010). Infants' understanding of everyday social interactions: a dual process account. *Cognition* 114, 197–206.

Gredebäck, G., Örnkloo, H., and von Hofsten, C. (2006). The development of reactive saccade latencies. *Exp. Brain Res.* 173, 159–164.

Gredebäck, G., Stasiewicz, D., Falck-Ytter, T., von Hofsten, C., and Rosander, K. (2009). Action type and goal type modulate goal-directed gaze shifts in 14-month-old infants. *Dev. Psychol.* 45, 1190–1194.

Gredebäck, G., and von Hofsten, C. (2007). "Taking an action perspective on infant's object representations," in *From Action to Cognition*, eds C. von Hofsten and K. Rosander (Amsterdam: Elsevier), 265–282.

Hamlin, J. K., and Woodward, A. L. (2005). "7-month-old infants differentially imitate purposeful and non-purposeful actions," in *Poster Presented at the 2005 Biennial Meeting of the Society for Research in Child Development (SRCD)*, Atlanta, GA.

Hofer, T., Hauf, P., and Aschersleben, G. (2005). Infants' perception of goal-directed actions performed by a mechanical device. *Infant Behav. Dev.* 28, 466–480.

Hofer, T., Hauf, P., and Aschersleben, G. (2007). Infant's perception of goal-directed actions on television. *Br. J. Dev. Psychol.* 25, 485–498.

Hofer, T., Hohenberger, A., Hauf, P., and Aschersleben, G. (2008). The impact of maternal interaction style on infant action understanding. *Infant Behav. Dev.* 31, 115–126.

Hood, B. M., Cole-Davies, V., and Dias, M. (2003). Looking and search measures of object knowledge in preschool children. *Dev. Psychol.* 39, 61–70.

Hunnius, S., and Bekkering, H. (2010). The early development of object knowledge: a study of infants' visual anticipations during action observation. *Dev. Psychol.* 46, 446–454.

Johnson, S. P., Amso, D., and Slemmer, J. A. (2003). Development of object concepts in infancy: evidence for early learning in an eye-tracking paradigm. *Proc. Natl. Acad. Sci. U.S.A.* 100, 10568–10573.

Jovanovic, B., Király, I., Elsner, B., Gergely, G., Prinz, W., and Aschersleben, G. (2007). The role of effects for infant's perception of action goals. *Psychologia* 50, 273–290.

Kamewari, K., Kato, M., Kanda, T., Ishiguro, H., and Hiraki, K. (2005). Six-and-a-half-month-old children positively attribute goals to human action and to humanoid-robot motion. *Cogn. Dev.* 20, 303–320.

Kanakogi, Y., and Itakura, S. (2011). Developmental correspondence between action prediction and motor ability in early infancy. *Nat. Commun.* 2, 341.

Keen, R. E. (2003). Representation of objects and events: why do infants look so smart and toddlers look so dumb? *Curr. Dir. Psychol. Sci.* 12, 79–83.

Kenward, B. (2010). 10-month-olds visually anticipate an outcome contingent on their own action. *Infancy* 15, 337–361.

Kochukhova, O., and Gredebäck, G. (2007). Learning about occlusion: initial assumptions and rapid adjustments. *Cognition* 105, 26–46.

Kochukhova, O., and Gredebäck, G. (2010). Preverbal infants anticipate that food will be brought to the mouth: an eye tracking study of manual feeding and flying spoons. *Child Dev.* 81, 1729–1738.

Kovacs, A. M., and Mehler, J. (2009). Cognitive gains in 7-month-old bilingual infants. *Proc. Natl. Acad. Sci. U.S.A.* 106, 6556–6560.

Leslie, A. M., Xu, F., Tremoulet, P. D., and Scholl, B. J. (1998). Indexing and the object concept: developing "what" and "where" systems. *Trends Cogn. Sci. (Regul. Ed.)* 2, 10–18.

Lockman, J. J., Ashmead, D. H., and Bushnell, E. W. (1984). The development of anticipatory hand orientation during infancy. *J. Exp. Child. Psychol.* 37, 176–186.

Luo, Y. (2011). Three-month-old infants attribute goals to a non-human agent. *Dev. Sci.* 14, 453–460.

Luo, Y., and Baillargeon, R. (2005). Can a self-propelled box have a goal? Psychological reasoning in 5-month-old infants. *Psychol. Sci.* 16, 601–608.

Mareschal, D., and Johnson, M. H. (2003). The "what" and "where" of object representations in infancy. *Cognition* 88, 259–276.

Mareschal, D., Plunkett, K., and Harris, P. (1999). A computational and neuropsychological account of object-oriented behaviours in infancy. *Dev. Sci.* 2, 306–317.

Mash, C., Keen, R. E., and Berthier, N. E. (2003). Visual access and attention in two-year-olds' event reasoning and object search. *Infancy* 4, 371–388.

Mash, C., Novak, E., Berthier, N. E., and Keen, R. E. (2006). What do two-year-olds understand about hidden-object events. *Dev. Psychol.* 42, 263–271.

McMurray, B., and Aslin, R. N. (2004). Anticipatory eye movements reveal infants' auditory and visual categories. *Infancy* 6, 203–229.

Meltzoff, A. N. (1995). Understanding the intentions of others: re-enactment of intended acts by 18-month-old children. *Dev. Psychol.* 31, 838–850.

Melzer, A., Prinz, W., and Daum, M. M. (2012). Production and observation of contralateral reaching: a close link by 12 months of age. *Infant Behav. Dev.* 35, 570–579.

Mishkin, M., and Ungerleider, L. G. (1982). Contribution of striate inputs to the visuospatial functions of parieto-preoccipital cortex in monkeys. *Behav. Brain Res.* 6, 57–77.

Mounoud, P., and Bower, T. G. R. (1974). Conservation of weight in infants. *Cognition* 3, 29–40.

Muggleton, N. G., Chen, C.-Y., Tzeng, O. J. L., Hung, D. L., and Juan, C.-H. (2010). Inhibitory control and the frontal eye fields. *J. Cogn. Neurosci.* 22, 2804–2812.

Paulus, M., Hunnius, S., Vissers, M., and Bekkering, H. (2011a). Bridging the gap between the other and me: the functional role of motor resonance and action effects in infants' imitation. *Dev. Sci.* 14, 901–910.

Paulus, M., Hunnius, S., van Wijngaarden, C., Vrins, S., van Rooij, I., and Bekkering, H. (2011b). The role of frequency information and teleological reasoning in infants' and adults' action prediction. *Dev. Psychol.* 47, 976–983.

Perner, J., and Ruffman, T. (2005). Infants' insight into the mind: how deep? *Science* 308, 214–216.

Piaget, J. (1952). *The Origins of Intelligence in Children.* New York: International Universities Press.

Piaget, J. (1954). *The Construction of Reality in the Child.* New York: Basic Books.

Rao, H. Y., Zhou, T. G., Zhuo, Y., Fan, S., and Chen, L. (2003). Spatiotemporal activation of the two visual pathways in form discrimination and spatial location: a brain mapping study. *Hum. Brain Mapp.* 18, 79–89.

Rosander, K., and von Hofsten, C. (2004). Infants' emerging ability to represent occluded object motion. *Cognition* 91, 1–22.

Schall, J. D. (2002). The neural selection and control of saccades by the frontal eye field. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 357, 1073–1082.

Schall, J. D., Stuphorn, V., and Brown, J. W. (2002). Monitoring and control of action by the frontal lobes. *Neuron* 36, 309–322.

Schlottmann, A., and Ray, E. (2010). Goal attribution to schematic animals: do 6-month-olds perceive biological motion as animate? *Dev. Sci.* 13, 1–10.

Schroeder, C. E., Mehta, A. D., and Givre, S. J. (1998). A spatiotemporal profile of visual system activation revealed by current source density analysis in the awake macaque. *Cereb. Cortex* 8, 575–592.

Sommerville, J. A., and Woodward, A. L. (2005). Infants' sensitivity to the causal features of means-end support sequences in action and perception. *Infancy* 8, 119–145.

Spelke, E. S., Breinlinger, K., Macomber, J., and Jacobson, K. (1992). Origins of knowledge. *Psychol. Rev.* 99, 605–632.

von Hofsten, C. (1980). Predictive reaching for moving objects by human infants. *J. Exp. Child. Psychol.* 30, 369–382.

von Hofsten, C. (1983). Catching skills in infancy. *J. Exp. Psychol. Hum. Percept. Perform.* 19, 75–85.

von Hofsten, C. (2004). An action perspective on motor development. *Trends Cogn. Sci. (Regul. Ed.)* 8, 266–272.

von Hofsten, C., and Fazel-Zandy, S. (1984). Development of visually guided hand orientation in reaching. *J. Exp. Child. Psychol.* 38, 208–219.

von Hofsten, C., and Johansson, K. (2009). Planning to reach for a rotating rod: developmental aspects. *Z. Entwicklungspsychol. Padagog. Psychol.* 41, 207–213.

von Hofsten, C., and Lindhagen, K. (1979). Observations on the development of reaching for moving objects. *J. Exp. Child. Psychol.* 28, 158–173.

von Hofsten, C., and Rönnqvist, L. (1988). Preparation for grasping an object: a developmental study. *J. Exp. Psychol. Hum. Percept. Perform.* 14, 610–621.

Wagner, L., and Carey, S. (2005). 12-month-old infants represent probable endings of motion events. *Infancy* 7, 73–83.

Wilcox, T., Nadel, L., and Rosser, R. (1996). Location memory in healthy preterm and fullterm infants. *Infant Behav. Dev.* 19, 309–323.

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition* 69, 1–34.

Woodward, A. L. (1999). Infants' ability to distinguish between purposeful and non-purposeful behaviors. *Infant Behav. Dev.* 22, 145–160.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 29 May 2012; accepted: 09 September 2012; published online: 27 September 2012.

*Citation: Daum MM, Attig M, Gunawan R, Prinz W and Gredebäck G (2012) Actions seen through babies' eyes: a dissociation between looking time and predictive gaze. Front. Psychology **3**:370. doi: 10.3389/fpsyg.2012.00370*

This article was submitted to Frontiers in Cognition, a specialty of Frontiers in Psychology.

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 09 October 2012
doi: 10.3389/fpsyg.2012.00391

Goal salience affects infants' goal-directed gaze shifts

***Ivanina Henrichs*[1]*, Claudia Elsner*[2]*, Birgit Elsner*[1] *and Gustaf Gredebäck*[2]**

[1] *Department of Psychology, University of Potsdam, Potsdam, Germany*
[2] *Department of Psychology, Uppsala University, Uppsala, Sweden*

Edited by:
Dorit Wenke, Humboldt University at Berlin, Germany

Reviewed by:
Victoria Southgate, University of London, UK
Anne Melzer, Max Planck Institute for Human Cognitive and Brain Sciences, Germany

****Correspondence:***
Ivanina Henrichs, Department of Psychology, University of Potsdam, Karl-Liebknecht-Street 24/25, 14476 Potsdam, Germany.
e-mail: ivanina.henrichs@uni-potsdam.de

Around their first year of life, infants are able to anticipate the goal of others' ongoing actions. For instance, 12-month-olds anticipate the goal of everyday feeding actions and manual actions such as reaching and grasping. However, little is known whether the salience of the goal influences infants' online assessment of others' actions. The aim of the current eye-tracking study was to elucidate infants' ability to anticipate reaching actions depending on the visual salience of the goal object. In Experiment 1, 12-month-old infants' goal-directed gaze shifts were recorded as they observed a hand reaching for and grasping either a large (high-salience condition) or a small (low-salience condition) goal object. Infants exhibited predictive gaze shifts significantly earlier when the observed hand reached for the large goal object compared to when it reached for the small goal object. In addition, findings revealed rapid learning over the course of trials in the high-salience condition and no learning in the low-salience condition. Experiment 2 demonstrated that the results could not be simply attributed to the different grip aperture of the hand used when reaching for small and large objects. Together, our data indicate that by the end of their first year of life, infants rely on information about the goal salience to make inferences about the action goal.

Keywords: anticipation, eye movement, salience, infant, action understanding

INTRODUCTION

The ability to anticipate other people's actions is crucial for the planning and control of one's own actions in accordance with the actions of others. Already at the age of 6–9 months, infants are able to predict others' goal-directed actions (Southgate et al., 2010; Kanakogi and Itakura, 2011). Moreover, around their first year of life, they anticipate a variety of different manual actions such as reaching (Cannon and Woodward, 2012), placing objects inside a container (Falck-Ytter et al., 2006), or everyday feeding actions (Gredebäck and Melinder, 2010).

A considerable amount of literature indicates a close relationship between infants' ability to anticipate observed actions and their motor ability of the same actions (Gredebäck and Kochukhova, 2010; Gredebäck and Melinder, 2010). To illustrate, Kanakogi and Itakura (2011) demonstrated that 6-month-olds' emerging motor ability to perform grasping actions corresponded to their ability to anticipate the goal of observed grasping actions. Similarly, Falck-Ytter et al. (2006) found that 12- but not 6-month-olds were able to anticipate the goal of a manual action, consisting of transporting balls to a container. Because 12- but not 6-month-olds have extensive experience with reaching and placing actions, the authors interpreted these data as evidence for the link between motor experience and action understanding. Additionally, Gredebäck and Melinder (2010) found that 12- but not 6-month-olds' anticipatory performance of observed feeding actions was correlated with their lifetime experience being fed.

Apart from motor experience, there are other influence factors as well. For instance, when observing other people act on objects, infants are commonly faced with scenes where multiple objects with different shape and size are available (Ambrosini et al., 2011). Indeed, information about the properties of the goal is crucial for the planning and control of one's own actions (Castiello, 2005). Research on human prehension indicates that object parameters such as size, shape, and weight have a great impact on the execution of grasping actions in adults (Smeets and Brenner, 1999; Castiello, 2005). Recently, it was found that the properties of the goal have also an impact on adults' goal anticipations (Eshuis et al., 2009). To illustrate, in Ambrosini et al.'s (2011) study, adults observed action events in which a hand was reaching for and grasping one of two differently sized goal objects. In one condition, the hand was pre-shaped so that adults could use the grip information in order to predict the goal of the ongoing action (a whole hand grip for the big object and a precision grip for the small object). In the no-shape condition, the hand moved with a closed fist configuration to the goal objects. Results showed that in the pre-shape condition, adults looked at the correct goal object ahead of time, with earlier gaze-arrival times at the large goal object compared to the small goal object. Interestingly, even in the no-shape condition, adults looked ahead of time toward the large object. This effect was ascribed to the salience of the large object. Analogously, Eshuis et al. (2009) presented adults with videos in which a human agent was moving a toy frog toward a bucket. In one condition the transporting action was followed by end-effects: when the toy entered the bucket water ripples were shown and frog croaking was played. In a control condition, there were no end-effects. Eshuis et al. (2009) found an earlier gaze-arrival time at the action goal in the end-effects condition compared to the no-effects condition, indicating an impact of goal salience on adults' goal anticipations.

To date, little is known about the degree to which the properties of the goal influence infants' online assessment of others' actions. For instance, Cannon et al. (2012) used similar set-up to that applied in Falck-Ytter et al.'s (2006) study, presenting 12-month-old infants with events in which a human agent was placing three balls into a bucket. Cannon et al. (2012) found later gaze-arrival times compared to those found for 12-month-olds in Falck-Ytter et al.'s (2006) study. More specifically, in Falck-Ytter et al.'s study, infants' gaze shifts passed the threshold of 0 ms, indicating that they were able to look at the goal ahead of time, whereas infants' gaze shifts in Cannon et al.'s study did not. Cannon et al. attributed this effect to a procedural difference between the two studies. Namely, while in the first study there were end-effects accompanying the arrival of the ball into the bucket (an artificial sound was played and a face pattern imposed on the bucket), there were no such end-effects in the latter study.

Although these studies indirectly support the notion that goal salience might have an impact on goal anticipations, this idea has not yet been directly addressed in infants. Hence, the following experiments seek to investigate the impact of goal salience on infants' ability to anticipate reaching actions. In two experiments, we demonstrate that the visual salience of the goal object has an impact on infants' goal anticipations. This effect cannot be simply attributed to the different grip aperture of the hand when reaching for small and large objects. Our results indicate that action prediction in infancy might be influenced by the properties of the goal such as the size of the goal object.

EXPERIMENT 1

In Experiment 1, we investigated the influence of the visual salience of the goal on 12-month-old infants' ability to anticipate reaching actions performed by a human agent. We presented videos in which a human hand reached for one goal object. In the high-salience condition, the hand reached for a large goal object, whereas in the low-salience condition, it reached for a small goal object. In order to investigate infants' goal anticipations, we measured their predictive gaze shifts (Gredebäck et al., 2010). If infants use information about the salience of the action goal, then they should show earlier gaze-arrival times in the high-salience condition compared to the low-salience condition (Ambrosini et al., 2011). If goal salience does not have an impact on infants' processing of reaching actions, then gaze performance between conditions should not differ. To our knowledge, this is the first infant study to directly investigate the influence of goal salience on infants' goal anticipation.

MATERIAL AND METHODS

Participants

The final sample consisted of 24 12-month-old infants, 12 in each condition (6 females in each condition). The mean age was 365 days ($SD = 7$) in the high-salience condition and 366 days ($SD = 7$) in the low-salience condition. An additional two infants were excluded because of fussiness or calibration failure. Parents were contacted by phone and signed a consent form prior to their participation. The study was approved by the Regional Ethic Committee according to the 1964 Declaration of Helsinki. Each family was given a gift certificate (approximately 10 Euro) for participation.

Stimuli and apparatus

Gaze was measured with a Tobii T120 near infrared eyetracker (Tobii, Stockholm, Sweden) with an infant add-on; monitor size 17″; accuracy 0.5°, sampling rate 60 Hz. A standard five-point calibration was used (Gredebäck et al., 2010). Infants were presented with videos (25.5 × 19.1 visual degrees) of a human hand reaching for one goal object placed on a table.

Each video began with a still frame giving a view of a wooden table top, filmed from above, with either one small or one large rectangular blue object positioned in the middle of the screen. After 500 ms, a human hand entered the scene from above and moved to the upper middle of the table (500–960 ms). It rested motionless on the table (960–1800 ms) and then reached for the goal object (1800–2960 ms). The hand grasped the object (2960–3760 ms), and rested on the object for the last 1240 ms (see **Figure 1**). Each video lasted for approximately 5000 ms. There were two videos, one in which the hand reached for the small goal object (see **Figure 1A**) and one in which the hand reached for the large goal object (see **Figure 1B**).

Procedure

During the experiment, infants sat on their caregivers' lap in a curtained experimental room and performed a calibration procedure first. Then infants were presented with videos of either the high-salience or low-salience condition, interleaved with brief animations designed to reorient their attention to the screen. There were 14 trials in each condition. Each family spent approximately 15 min in the lab.

Data reduction and analysis

Infants' predictive gaze shifts during the reaching action were analyzed. Two areas of interest (AOIs) were created to cover the hand (5.8 visual degrees horizontal extension) and the goal object (5.6 visual degrees horizontal extension) the hand was reaching for (see **Figure 1**). The size of the AOI covering the goal object was identical in both conditions. Infants first had to fixate the hand for at least 200 ms and then shift their gaze to the goal AOI. Mean gaze-arrival times were calculated by subtracting the time when infants first looked inside the goal AOI from the time when the observed hand first entered the same AOI. Thus, positive numbers refer to a gaze-arrival before the hand arrived at the goal AOI, value of zero indicates gaze-arrival at the same time as the hand and negative numbers a gaze-arrival after the hand arrived at the goal AOI. Gaze shifts were classified as functionally predictive if they occurred before the hand entered the goal AOI, that is, if the 95% confidence interval with lower boundary for each group mean was above 0. This threshold is conservative and ensures that infants actually look at the correct location ahead of time (Gredebäck et al., 2010). It has previously been used in the majority of action prediction studies in infancy (Falck-Ytter et al., 2006; Gredebäck and Melinder, 2010; Kanakogi and Itakura, 2011).

Data from each action were included if infants attended to the hand for at least 200 ms and fixated the goal AOI no later than 1000 ms after the hand had entered the goal AOI. All included

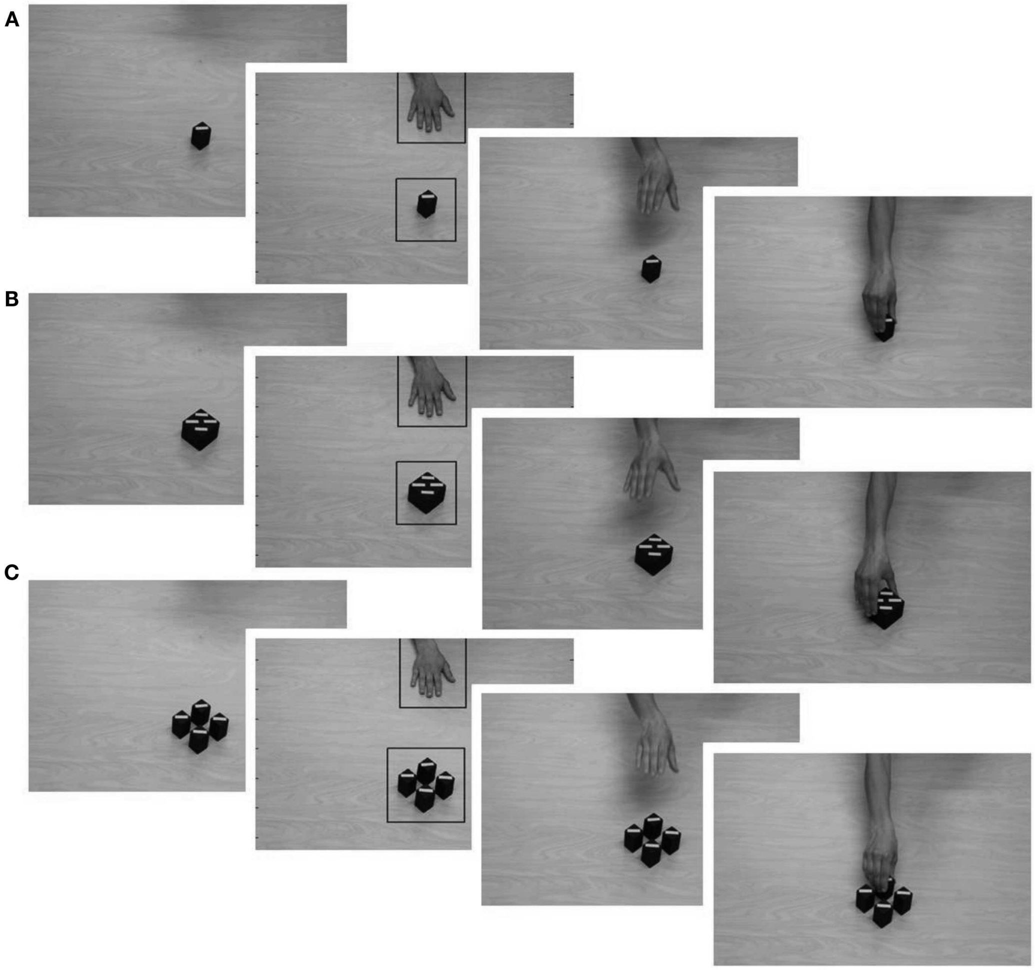

FIGURE 1 | Snapshots of the action sequence in each video, depicting the beginning of each movie, the hand resting on the table, and the reaching and grasping action in the low-salience (A) and high-salience (B) condition in Experiment 1 and the low-salience (A) and high-salience (C) condition in Experiment 2. Areas of interest (AOIs) for the hand and for each goal object are marked with black rectangles.

infants had minimum five out of 14 valid trials. Mean gaze-arrival times were aggregated over trials 1–9. The last five trials were excluded because of lack of attention. Mean gaze-arrival times were compared between conditions using independent *t* tests. Effect sizes were calculated using Cohen's *d*. Both linear and curvilinear regression analyses were run to analyze learning effects across trials 1–9 in each condition. The regression line with the highest explained variance (linear or curvilinear) was reported and displayed in **Figure 3**.

RESULTS

Overall gaze-arrival time

A comparison of the aggregated mean gaze-arrival times of trials 1–9 revealed a significant difference between conditions, $t(15.62) = 2.52$, $p = 0.023$, $d = 1.27$. Infants in the high-salience condition showed significantly earlier mean gaze-arrival times than infants in the low-salience condition (see **Figure 2**).

Infants' mean gaze-arrival times in the high-salience condition passed the threshold of 0 ms, 95% CI [182, 704], whereas infants' mean gaze-arrival times in the low-salience condition did not, 95% CI [−49, 192], indicating that only infants in the prior group were able to fixate the goal object ahead of time.

Learning effects

Within the first two trials, infants in the high-salience condition learned to predict the goal object of the reaching action and their performance improved throughout the experimental session. This learning effect is best described using the logarithmic function, $y = 197.59\ln(x) + 181.39$, expressing a rapid improvement of gaze performance over the course of trials (see **Figure 3**), $R^2_{\text{adj}} = 0.52$, $F(1, 8) = 9.84$, $p = 0.02$. By contrast, infants in the low-salience condition did not show improvement of performance throughout the experimental session (see **Figure 3**).

DISCUSSION

The purpose of Experiment 1 was to investigate the impact of the visual salience of the goal object on infants' goal anticipations of observed reaching actions. We found that 12-month-old infants

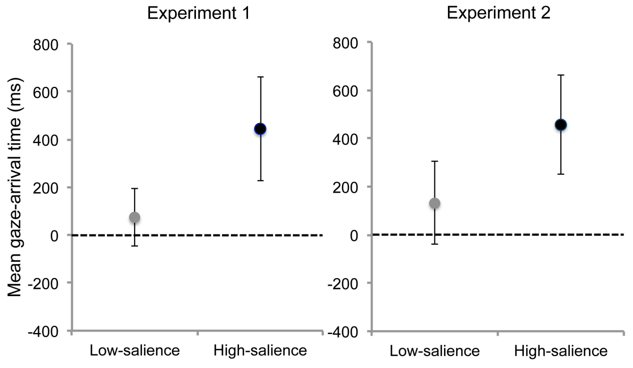

FIGURE 2 | Mean gaze-arrival times (in ms) relative to the hand's arrival time for the aggregated means of trials 1–9 for the high- and low-salience condition in Experiments 1 and 2. Error bars depict 95% confidence intervals. The horizontal line represents the threshold of 0 ms and differentiates predictive from reactive gaze shifts. Values above 0 correspond to earlier arrival of gaze relative to the arrival of the hand at the goal area.

exhibited gaze shifts significantly earlier when the observed hand reached for the large goal object as compared to when it reached for the small goal object. Additionally, only infants in the high-salience condition were able to look at the goal ahead of time, fixating the goal object before the hand arrived at the goal AOI. Therefore, our data indicate that by the end of their first year of life, infants rely on information about the properties of goal objects to make inferences about the action goal. To our knowledge, this is the first infant study to directly demonstrate this effect during the observation of reaching actions.

In Falck-Ytter et al.'s (2006) study, 12-month-olds looked at the goal ahead of time when observing a human agent transporting balls into a bucket, indicating predictive gaze shifts. By contrast, same-aged infants in Cannon et al.'s (2012) study were not able to fixate the goal object ahead of time when observing comparable transporting actions. Cannon et al. attributed these differences to the goal being more salient in Falck-Ytter et al.'s study. In our study, mean gaze-arrival times in the high-salience condition (443 ms) were comparable to those in previous infant studies on action prediction, indicating that the salience of the goal in this condition was similar to the goal salience in these studies (Falck-Ytter et al., 2006; Gredebäck and Melinder, 2010; Kanakogi and Itakura, 2011). By contrast, when infants observed actions during which the hand reached for a small goal object, they were not able to fixate the goal object ahead of time, resulting in later mean gaze-arrival times (72 ms), comparable to that found by Cannon et al.

In the current investigation, infants in the high-salience condition rapidly learned to track the reaching action in a predictive manner within only a few trials, whereas infants in the low-salience condition did not so. Interestingly, the learning curve in the high-salience condition was highly similar to that found by Kochukhova and Gredebäck (2007) for occluded non-social action events. It seems that the salience of the goal affects infants' learning to anticipate goal-directed actions throughout the experimental session.

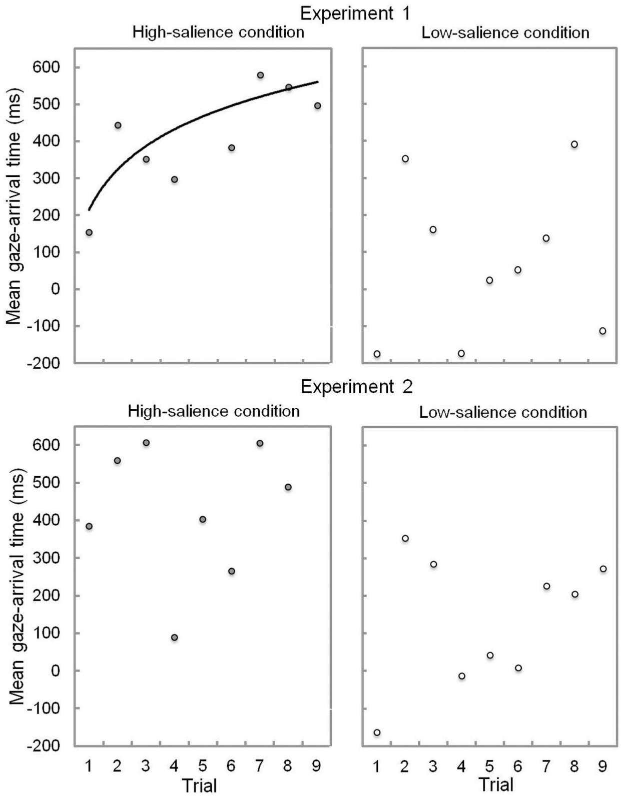

FIGURE 3 | Mean gaze-arrival times (in ms) over trials 1–9 in the high- and low-salience condition in Experiments 1 and 2. The solid curve depicts the regression line with most explained variance. Note: no significant regression line could be fitted for the low-salience condition in Experiment 1 as well as for the high- and low-salience conditions in Experiment 2.

Taken together, our data indicate that the properties of the goal object have an impact on infants' goal anticipations. Particularly, the visual salience of the goal object contributed to the overall predictive gaze shifts in the high-salience condition. This notion is supported by Ambrosini et al.'s (2011) study in which adults exhibited earlier gaze-arrival times when the hand was reaching for a large goal object as compared to when it was reaching for a small goal object. However, Ambrosini et al. attributed this effect to the pre-shaping of the hand rather than to the visual salience of the goal object. More specifically, the hand was pre-shaped to a whole hand grip to reach for the large object and to a precision grip to reach for the small object, the latter requiring more time to be processed than the prior (Ambrosini et al., 2011). The idea that observers use information about the hand shape when processing others actions, is also supported by other studies. For instance, Fischer et al. (2008) demonstrated that adults rapidly inferred the goal object of an observed grasping action from the shape of the actor's hand in a reaction-time study. Moreover, already at the age of 6 months, infants are able to infer the size of the goal object from the pre-shaping of the grasping hand (Daum et al., 2009). Although we used a power grip in both conditions, the configuration of the hand actually differed between conditions. More specifically, the aperture between all fingers and the thumb was larger in the high-salience condition than in the low-salience condition. Thus, it might be that the earlier gaze-arrival times in the high-salience condition were not only due to the size of the goal object but also to the wider grip of the hand. This possibility was addressed in Experiment 2.

EXPERIMENT 2

In Experiment 2, we tested the assumption that the grip aperture accounts for the difference in gaze performance between conditions in Experiment 1. We presented videos in which a human hand reached for one small goal object in both conditions. In the high-salience condition, there were four small objects available, which preserved the overall higher visual salience of the goal area in this condition. The low-salience condition was identical to that of Experiment 1. Thus, in Experiment 2, all infants saw the hand reaching for a small goal object resulting in a narrow grip in both conditions. If the pre-shaping of the hand had the greatest impact on infants' predictive gaze shifts in Experiment 1, we would expect no difference in the mean gaze-arrival times between conditions in Experiment 2. If however, the visual salience of the goal was crucial for the predictive gaze shifts, then infants in the high-salience condition should still show earlier gaze-arrival times than infants in the low-salience condition.

MATERIAL AND METHODS

Participants

The final sample consisted of 24 12-month-old infants, 12 in each condition (6 females in each condition). None of the infants had participated in Experiment 1. The mean age was 365 days ($SD = 7$) in the high-salience condition and 364 days ($SD = 9$) in the low-salience condition. An additional three infants were excluded because of fussiness or calibration failure.

Stimuli and apparatus

The stimuli and apparatus were identical to that of Experiment 1 with the following exception. In the high-salience condition, infants saw a movie in which the hand was grasping for one of four small rectangular objects positioned next to each other, forming a rectangular form. The total size of the four small objects was comparable to that of the large goal object used in Experiment 1. The hand reached for and grasped the nearest of the four small objects which was exactly on the same position as the small object in the low-salience condition (see **Figure 1C**). In the low-salience condition, infants were presented with an action event identical to that in Experiment 1 (see **Figure 1A**). Thus in both movies, the hand was shaped to a narrow power grip when reaching for the goal object.

Procedure, data reduction, and analysis

The procedure, data reduction, and analyses were identical to that of Experiment 1.

RESULTS

Overall gaze-arrival time

There was a significant difference between conditions, $t(22) = 2.40$, $p = 0.025$, $d = 1.02$. Infants in the high-salience condition showed significantly earlier mean gaze-arrival times than infants in the low-salience condition (see **Figure 2**). Mean gaze-arrival times did not differ between Experiment 1 and 2 neither for the high-salience condition, $t(22) = 0.08$, $p = 0.93$, $d = 0.03$, nor for the low-salience condition, $t(22) = 0.52$, $p = 0.61$, $d = 0.22$.

Infants' gaze-arrival times in the high-salience condition passed the criterion of 0 ms, 95% CI [252, 664], whereas infants' gaze-arrival times in the low-salience condition did not, 95% CI [−43, 301], suggesting that only infants in the first group were able to fixate the goal object prior to the arrival of the hand at the goal AOI.

Learning effects

There was no linear or curvilinear regression line fitting the data, indicating no learning effects in Experiment 2 (see **Figure 3**).

DISCUSSION

In Experiment 2, we addressed the possibility that the higher gaze performance in the high-salience condition compared to the low-salience condition in Experiment 1 was due to the wider grip of the hand rather than to the visual salience of the large object. Although the grip aperture was identical in both conditions, we found a significant difference between conditions, indicating that infants in the high-salience condition (four small objects available) exhibited gaze shifts much earlier than infants in the low-salience condition (one small object available). Moreover, the mean gaze-arrival times in the high-salience condition in Experiment 2 ($M = 458$ ms) were similar to those in Experiment 1 ($M = 443$ ms). Furthermore, just like in Experiment 1, only infants in the high-salience condition were able to look at the goal object ahead of time, before the hand arrived at the goal AOI. As soon as the goal is highly salient, infants anticipate the goal of a reaching hand in a functional way no matter if the hand is pre-shaped in a wide or narrow power grip. Thus, it seems that gaze performance is not only affected by subtle motor information (see Falck-Ytter, 2012), but also by the object-related properties such as the size of the goal object. To our knowledge, this is the first study to disentangle the contribution of these two factors on infants' goal-directed gaze shifts.

Interestingly, in Experiment 2, we failed to find any learning effects during the experimental session. Because most infant studies on action prediction either did not find or they did not report learning effects (Falck-Ytter et al., 2006; Gredebäck et al., 2009; Kanakogi and Itakura, 2011; Cannon et al., 2012), it is difficult to explain the presence of learning effects in Experiment 1 and the absence of those in Experiment 2. It might be that a larger sample size is required in order to find clearly visible learning effects. Alternatively, although the overall size of the goal area in the high-salience condition was kept similar between experiments, in Experiment 1, the hand was approaching the goal object pre-shaped in a wide power grip, whereas in Experiment 2 the reaching hand was pre-shaped in a narrow power grip. It might be that the learning effect in the high-salience condition in Experiment 1 was influenced by both the salience of the goal object and the grip aperture used during the reach. Future research should address the factors influencing infants' learning when observing others' manual actions.

GENERAL DISCUSSION

This study is the first to demonstrate that infants' goal-directed gaze shifts are modulated by the visual salience of the goal object. Twelve-month-olds in Experiment 1 exhibited predictive gaze shifts significantly earlier when the observed hand reached for a large as compared to a small goal object, which is consistent with Ambrosini et al.'s (2011) findings with adults. Interestingly, Ambrosini et al. attributed the difference in gaze performance in their study to the pre-shaping of the hand rather than to the visual salience of the goal object. Although we kept the grip aperture constant between conditions in Experiment 2, infants in the high-salience (large goal area) condition still fixated the goal earlier than infants in the low-salience (small goal area) condition. Hence, our data indicate that it is the visual salience of the goal object what accounted for differences in gaze performance between conditions.

One difference between the two studies was that in the present investigation the reaching hand was always shaped to a power grip, only slightly varying its aperture depending on the size of the goal object. By contrast, in Ambrosini et al.'s (2011) study the hand was pre-shaped to a power or precision grip depending on the to-be-grasped object. The authors argued that the precision grip needs more time to be processed compared to the power grip. It might be that infants' processing of a power grip is independent from the exact distance between the fingers and the thumb. Additionally, in the adult study, the large and the small objects were both present during the reaching action. Thus, another likely explanation is that information about the exact kinematics of the handgrip is crucial in situations, where multiple objects are present and the goal of the reaching action cannot be predicted in advance (Falck-Ytter, 2012). This idea is supported by an adult study in which participants were able to predict the goal of an ongoing action from the kinematics of the moving hand without prior knowledge of the agent's intention (Rotman et al., 2006). As soon as a single goal object is available, infants might only pay attention to global kinematic information from the moving arm, neglecting more subtle motor information such as the grip aperture.

It might be that infants' earlier gaze-arrival times in the high-salience condition were driven by a general selective process modulated by the size of the goal object. This notion is supported by adult studies indicating that large objects capture attention in visual search tasks (Proulx, 2010). Thus, it is possible that a larger object captures more attention leading to earlier gaze-arrival times irrespective of the action type observed. Indeed, in Ambrosini et al.'s (2011) study, adults exhibited predictive gaze shifts to the large object even when the hand moved to the goal objects with a closed fist configuration. This effect was ascribed to the higher visual salience of the large object. Analogously, in Eshuis et al.'s (2009) study, apart from the human agent condition, there was a self-propelled condition, in which the frogs moved to the bucket on their own. Just as in the human agent condition, the transporting action was either followed by end-effects or not. Eshuis et al. found that as soon as end-effects accompany the transporting action, adults exhibit earlier gaze shifts irrespective of the action type, indicating a great impact of goal salience on adults' goal anticipations.

Alternatively, recent research indicates that the mirror neuron system (MNS) is involved in the processing of others' goal-directed actions (Rizzolatti and Craighero, 2004; Gallese et al., 2009). In their seminal study, Flanagan and Johansson (2003) demonstrated that when observing others' manual actions, adults exhibit similar predictive eye movements to those found when they execute the action themselves. This phenomenon is described by a direct matching mechanism within the MNS, in which observed actions are mapped onto the observer's motor representation of the same action (Rizzolatti et al., 2001; Rizzolatti and Craighero, 2004). Evidence for the direct matching hypothesis was also found in infants (Rosander and von Hofsten, 2011). Moreover, a considerable amount of research indicates that infants' ability to anticipate observed actions is tightly linked to their own motor experience with the same actions (Falck-Ytter et al., 2006; Gredebäck and Kochukhova, 2010; Gredebäck and Melinder, 2010; Kanakogi and Itakura, 2011). For instance, in Falck-Ytter et al.'s (2006) study, 6- and 12-month-old infants observed action sequences, consisting of a transporting balls to a container. While in one condition the balls were transported by a human agent, in another condition they moved on their own. Results indicated that 12- but not 6-month olds were able to fixate the goal ahead of time, but only when the human agent performed the action. Because 12- but not 6-month-olds have extensive experience with transporting actions, Falck-Ytter et al. interpreted these data as evidence for the link between motor experience and action understanding. Furthermore, Kanakogi and Itakura (2011) found that 6- but not 4-month-old infants were able to anticipate grasping actions and that infants' gaze performance corresponded to their emerging motor ability to perform grasping actions. Additionally, in control conditions including non-functional and non-human actions, they tracked those actions in a reactive manner. Similarly, Kochukhova and Gredebäck (2010) demonstrated that 6-month-olds anticipate that food is brought to the mouth, while combing actions and self-propelled spoons were tracked in a reactive manner. Together these studies demonstrate that infants' ability to predict others' actions is modulated by their motor experience with the same actions.

However, all of the above-mentioned studies only varied the type of the action, keeping the goal salience constant between conditions. By contrast, in the present investigation, we presented the same reaching action in both groups, varying the size of the goal object between conditions. As a result, 12-month-old infants were only able to anticipate the goal of the reaching action when the goal was highly salient. By comparison, infants in the low-salience condition failed to track the reaching action in a predictive manner. This is a surprising result given the fact that by the end of their first year of life, infants have gained extensive experience with reaching actions and are therefore supposed to be able to anticipate the reaching actions of others (Rosander and von Hofsten, 2011; Cannon et al., 2012). Thus, our data extend previous findings, suggesting that infants' action prediction is not only modulated by motor experience but also by the properties of the goal. This is in line with what was found by Falck-Ytter et al. (2006) and Cannon et al. (2012) who used similar action sequences in their studies. Namely, in the presence of end-effects, 12-month-olds in Falck-Ytter et al.'s study were able to predict the goal of the transporting action in a functional way, whereas in the absence of such effects in Cannon et al's study, infants failed to functionally predict the action goal. Hence, infants might not only need motor experience with a particular action, but also salient goals and end-effects in order to reliably predict those actions.

To sum up, this is the first infant study to find a direct evidence for the impact of goal salience on infants' goal anticipations of observed reaching actions. More specifically, our data suggest that in a simple reaching action setting, a highly salient goal facilitates infants' gaze shifts from the reaching hand to the goal object, enabling them to look at the goal object ahead of time. By contrast, in the case of low-salience, infants fail to track the reaching action in a predictive manner. It might be that a highly salient goal draws infants' attention irrespective of the action type observed, indicating a general selective process. However, given the evidence from previous research, it is more likely that goal salience interacts with infants' motor experience with the observed action. Future research should disentangle the role of these factors, varying both the action type and the salience of the goal. Only when we take into consideration the complex structure of predictive gaze shifts, we can understand how infants learn about the actions of others.

ACKNOWLEDGMENTS

This research was supported by the German Research Foundation (DFG, EL 252/5-2), by the Swedish Research Council (2011-1528), and by the Marie-Curie ITN RobotDoc (2010–2013). We thank parents and infants for their participation in the study and reviewers for valuable comments on an earlier version of this manuscript.

REFERENCES

Ambrosini, E., Costantini, M., and Sinigaglia, C. (2011). Grasping with the eyes. *J. Neurophysiol.* 106, 1437–1442.

Cannon, E. N., and Woodward, A. L. (2012). Infants generate goal-based action predictions. *Dev. Sci.* 15, 292–298.

Cannon, E. N., Woodward, A. L., Gredebäck, G., von Hofsten, C., and Turek, C. (2012). Action production influences 12-month-old infants' attention to others' actions. *Dev. Sci.* 15, 35–42.

Castiello, U. (2005). The neuroscience of grasping. *Nat. Rev. Neurosci.* 6, 726–736.

Daum, M. M., Vuori, M. T., Prinz, W., and Aschersleben, G. (2009). Inferring the size of a goal object from an actor's grasping movement in 6- and 9-month-old infants. *Dev. Sci.* 12, 854–862.

Eshuis, R., Coventry, K. R., and Vulchanova, M. (2009). Predictive eye movements are driven by goals, not by the mirror neuron system. *Psychol. Sci.* 20, 438–440.

Falck-Ytter, T. (2012). Predicting other people's action goals with low-level motor information. *J. Neurophysiol.* 107, 2923–2925.

Falck-Ytter, T., Gredebäck, G., and von Hofsten, C. (2006). Infants predict other people's action goals. *Nat. Neurosci.* 9, 878–879.

Fischer, M. H., Prinz, J., and Lotz, K. (2008). Grasp cueing shows obligatory attention to action goals. *Q. J. Exp. Psychol.* 61, 860–868.

Flanagan, J., and Johansson, R. S. (2003). Action plans used in action observation. *Nature* 424, 769–771.

Gallese, V., Rochat, M. Cossu, G., and Sinigaglia, C. (2009). Motor cognition and its role in the phylogeny and ontogeny of action understanding. *Dev. Psychol.* 45, 103–113.

Gredebäck, G., Johnson, S., and von Hofsten, C. (2010). Eye tracking in infancy research. *Dev. Neuropsychol.* 35, 1–19.

Gredebäck, G., and Kochukhova, O. (2010). Goal anticipation during action observation is influenced by synonymous action capabilities, a puzzling developmental study. *Exp. Brain Res.* 202, 493–497.

Gredebäck, G., and Melinder, A. (2010). Infants' understanding of everyday social interactions: a dual process account. *Cognition* 114, 197–206.

Gredebäck, G., Stasiewicz, D., Falck-Ytter, T., Rosander, K., and von Hofsten, C. (2009). Action type and goal type modulate goal-directed gaze shifts in 14-month-old infants. *Dev. Psychol.* 45, 1190–1194.

Kanakogi, Y., and Itakura, S. (2011). Developmental correspondence between action prediction and motor ability in early infancy. *Nat. Commun.* 2, 341.

Kochukhova, O., and Gredebäck, G. (2007). Learning about occlusion: initial assumptions and rapid adjustments. *Cognition* 105, 26–46.

Kochukhova, O., and Gredebäck, G. (2010). Preverbal infants anticipate that food will be brought to the mouth: an eye tracking study of manual feeding and flying spoons. *Child Dev.* 81, 1729–1738.

Proulx, M. J. (2010). Size matters: large objects capture attention in visual search. *PLoS ONE* 5, e15293. doi:10.1371/journal.pone.0015293

Rizzolatti, G., and Craighero, L. (2004). The mirror-neuron system. *Annu. Rev. Neurosci.* 27, 169–192.

Rizzolatti, G., Fogassi, L., and Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nat. Rev. Neurosci.* 2, 661–670.

Rosander, K., and von Hofsten, C. (2011). Predictive gaze shifts elicited during observed and performed actions in 10-month-old infants and adults. *Neuropsychologia* 49, 2911–2917.

Rotman, G., Troje, N. F., Johansson, R. S., and Flanagan, J. (2006). Eye movements when observing predictable and unpredictable actions. *J. Neurophysiol.* 96, 1358–1369.

Smeets, J. B. J., and Brenner, E. (1999). A new view on grasping. *Motor Control.* 3, 237–271.

Southgate, V., Johnson, M. H., El Karoui, I., and Csibra, G. (2010). Motor system activation reveals infants' online prediction of others' goals. *Psychol. Sci.* 21, 355–359.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 12 June 2012; accepted: 19 September 2012; published online: 09 October 2012.

*Citation: Henrichs I, Elsner C, Elsner B and Gredebäck G (2012) Goal salience affects infants' goal-directed gaze shifts. Front. Psychology **3**:391. doi: 10.3389/fpsyg.2012.00391*

This article was submitted to Frontiers in Cognition, a specialty of Frontiers in Psychology.

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 10 August 2012
doi: 10.3389/fpsyg.2012.00287

Sequential modulation of cue use in the task switching paradigm

Mike Wendt*[1]*, Aquiles Luna-Rodriguez*[1], *Renate Reisenauer*[2], *Thomas Jacobsen*[1] and *Gesine Dreisbach*[2]

[1] Helmut-Schmidt-University/University of the Federal Armed Forces Hamburg, Hamburg, Germany
[2] University of Regensburg, Regensburg, Germany

Edited by:
Dorit Wenke, Humboldt University at Berlin, Germany

Reviewed by:
Marco Steinhauser, University of Konstanz, Germany
Baptist Liefooghe, Ghent University, Belgium

****Correspondence:***
Mike Wendt, Experimental Psychology Unit, Helmut-Schmidt-University/University of the Federal Armed Forces Hamburg, Holstenhofweg 85, 22043 Hamburg, Germany.
e-mail: mike.wendt@hsu-hh.de;
Gesine Dreisbach, Department of Experimental Psychology, University of Regensburg, Universitätsstraße 31, 93053 Regensburg, Germany.
e-mail: gesine.dreisbach@psychologie.uni-regensburg.de

In task switching studies, pre-cuing of the upcoming task improves performance, indicating preparatory activation of the upcoming task-set, and/or inhibition of the previous task-set. To further investigate cue-based task preparation, the authors presented both valid and invalid task cues in a task switching experiment involving three tasks. Consistent with previous findings, a validity effect in terms of higher reaction times on invalidly compared to validly cued tasks was obtained. However, this validity effect was reduced following invalidly cued trials, suggesting dynamic adjustment in terms of decreased cue-based preparation after being misled. Performance was particularly impaired when the current task was the one that was invalidly cued on the preceding trial. This finding may reflect either particular reluctance to prepare or persisting inhibition of the erroneously prepared task-set from the pre-trial.

Keywords: task switching, cue validity, sequential modulation, preparation

INTRODUCTION

In task switching paradigms, participants frequently switch between two or more tasks. Typically the tasks comprise overlapping or identical sets of stimuli, therefore correct task execution critically depends on adoption of the correct task-set. In such situations, performance markedly improves with the option to prepare for the upcoming task. For instance, reaction times (RTs) and error rates decrease when the interval between a cue indicating the identity of the upcoming task and the imperative stimulus increases, more so on task switch than on task repetition (e.g., Meiran, 1996; for a review Kiesel et al., 2010).

Contrasting with the laboratory situation in which pre-knowledge about upcoming task demands can be provided with perfect validity, in real-life settings preparation for an impending activity is almost always associated with some degree of uncertainty. Several studies have addressed the question how the cognitive system deals with such conditions by using advance cues (or task sequence probability) to signal the occurrence of a specific task with differential probabilities. A general finding of such manipulations is that task performance increases with increasing task probability (Ruthruff et al., 2001; Hübner et al., 2004a; Dreisbach and Haider, 2006). Here, we will use informative cues that announce one specific task which – in some (25%) cases – will be followed by a different than the announced task. In the following, we will use the term valid cuing if the cue is followed by the announced task and the term invalid cuing, if the cue is unexpectedly followed by a different task.

The repeatedly observed decreased task performance under conditions of lower likelihood of task occurrence suggests that task preparation is gradually adjusted to its assumed utility. To date, however, still little is known about the mechanism that governs this processing adjustment. One possibility is that the cue usage depends on its recently experienced utility. Such sequential adjustment of cue-based task preparation bears some resemblance with another well documented form of sequential control adjustments. More precisely, reduced response interference from a distractor stimulus feature following trials associated with response conflict has been taken to reflect conflict-induced control adjustment (Botvinick et al., 2001; see also Gratton et al., 1992).

Applying such trial-to-trial adjustment to cue-based task preparation, one might expect that participants engage less in preparation for a cued task after recent invalid cuing. In the current study, we explored this possibility by analyzing task performance as a function of both cue validity on the current and the directly preceding trial.

Invalid task cuing may have additional consequences to possible adjustment of task preparation. One plausible notion is that competition from the erroneously prepared task-set due to invalid cuing may increase and thus trigger extra control measures in terms of reactive inhibition of the "wrong" task-set. In this connection, Hübner et al. (2004a) observed larger task switch costs if the preceding trial involved invalid rather than valid pre-cuing. Based on the assumption of enhanced competition from an erroneously prepared task, the authors attributed this impairment to reactive task-set inhibition (i.e., particular costs of switching to an inhibited task). Because the experiments in that study comprised only two tasks, a task switch following an invalidly cued trial always implied switching to the previously invalidly cued task. Due to this confound it was not possible to decide whether the post-invalid increase of switch cost was indeed due to reactive inhibition of the erroneously prepared task or reflects a general switching impairment after being misled – possibly brought about by reduced preparation. In the current study, we therefore used a task switching paradigm with *three* tasks, which allowed us to compare switching to the task which was erroneously prepared on the previous trial with switching to a different task.

To summarize, the current study investigated aftereffects of erroneous preparation of a task due to invalid cuing. Our main question was whether invalid cuing results in reduced preparatory engagement on the following trial. Furthermore, we wanted to know whether particular costs emerge when switching to a task which was invalidly cued on the preceding trial. In Experiment 1, we administered only task switch trials. In Experiment 2, we extended our investigation to task repetition trials.

EXPERIMENT 1

To investigate current and subsequent consequences of preparation for a not-to-be-executed task, we applied a frequently used task switching paradigm involving three tasks afforded by the same set of stimuli. We presented advance cues which signaled with 75% likelihood the occurrence of a specific task. Because the target stimuli were completely ambiguous regarding the current task, additional information regarding the identity of the relevant task had to be provided in the case of invalid cuing. This was done by presenting a second, coherently valid, task cue, simultaneously with the target stimulus. To ensure that participants did not ignore the advance cues, no simultaneous cues were presented on validly cued trials.

METHOD

Participants

Fourteen female and six male students of the University of Regensburg participated on a voluntary basis. They ranged in age from 19 to 33 years.

Apparatus and stimuli

Participants viewed the screen from a distance of about 60 cm. All target stimuli were presented in white color on a dark gray background and occurred inside a rectangular frame, which was centered on the screen. The digits 1–9, except 5 served as target stimuli. The target stimulus was always presented in the center of the screen and extended 0.7–0.9 cm horizontally and 1.1 cm vertically. Depending on the currently relevant task, participants were instructed to classify the character as odd or even, smaller or larger than five or as extreme or medial (i.e., 1, 2, 8, 9 vs. 3, 4, 6, 7). Responses were given on a standard QWERTZ keyboard. Participants were instructed to press the left key ("y"-key) for *smaller*, and the right key ("m"-key) for *larger*. The S-R assignment in the *odd/even* and the *medial/extreme* tasks was counterbalanced across participants.

PROCEDURE

On each trial, the task and the target stimulus was chosen randomly with the only constraint that no task was repeated on a subsequent trial. The target stimulus remained on the screen until a response key was pressed. Throughout each block of trials the rectangular frame was shown. It was filled with one of three colors to indicate the upcoming task. On a random 25% of trials this color cue did not match the upcoming task but the other task that was not presented on the preceding trial. Yellow indicated the odd/even task, cyan indicated the smaller/larger task, and purple indicated the extreme/medial task. These task cues were shown 500 ms after a response key was pressed, and remained on the screen for 100 ms, followed by a blank screen (except for the rectangular frame) for 400 ms, after which the target stimulus was presented. In case of an invalid trial, the target stimulus was presented with a simultaneous task cue, "overruling" the advance cue. In case of a validly cued task, no additional task cue was presented with the target stimulus.

Participants were instructed to identify the target by pressing the assigned response key as quickly as possible while avoiding errors. In case of an incorrect response, error feedback occurred for 800 ms slightly below the center of the screen. After three practice blocks of 20 trials each (the first block comprising only odd/even decisions, the second block comprising only smaller/larger decisions, and the third block comprising only extreme/medial decisions), participants were administered 10 blocks of 99 trials each. They were allowed to rest between blocks.

RESULTS

The first three trials of each experimental block were considered "warm-up" trials and did not enter the statistical analyses. In addition, data from trials with RTs deviating more than 2 standard deviations from the mean RT of each experimental condition per participant were considered outliers and were also excluded from the analyses.

In an Analysis of Variance (ANOVA) with repeated measures on the factors Validity on Current Trial (valid, invalid), and Validity on Preceding Trial (valid, invalid) on the mean RTs, both main effects reached significance. There was an invalidity cost of 328 ms, $F(1, 19) = 188.35$, $p < 0.001$, and post-invalid slowing of 112 ms, $F(1, 19) = 87.52$, $p < 0.001$. As can also be seen in **Figure 1**, post-invalid slowing was confined to validly cued trials, resulting in a two-way interaction, $F(1, 19) = 28.21$, $p < 0.001$. The invalidity cost amounted to 422 and 235 ms after validly and invalidly cued trials, respectively. The corresponding ANOVA on error rates revealed a significant effect of Validity on the Preceding Trial, $F(1, 19) = 5.03$, $p = 0.037$, and marginally significant effect of Validity on the Current Trial, $F(1, 19) = 4.28$, $p = 0.053$. Both factors entered into an interaction, $F(1, 19) = 5.51$, $p = 0.030$, reflecting that following an invalid trial, invalid trials were associated with less errors than validly cued trials (3.7 vs. 6.0%) whereas there was no difference after validly cued trials (4.1 vs. 3.9%).

To further explore the effects of previous invalid cuing, we ran a second analysis, confined to data from task switch trials following invalidly cued trials, only. More precisely, to find out, whether the observed post-invalid adjustments were further modulated by the specific task that was invalidly cued on the pre-trial, we

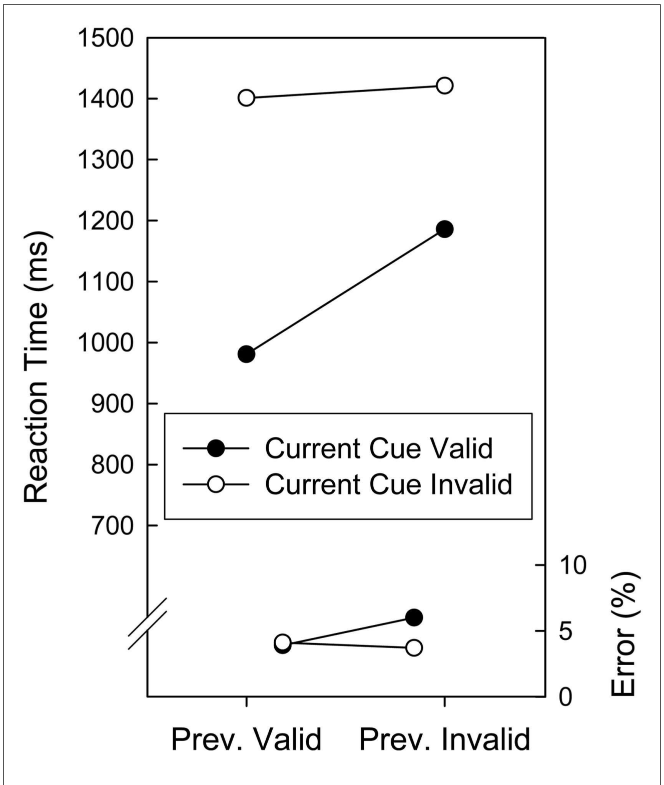

FIGURE 1 | Mean reaction times and error proportions of Experiment 1 as a function of cuing validity of the current and preceding trials.

ran an ANOVA with the factors Validity on Current Trial (valid, invalid) and Task (Invalidly) Cued on Preceding Trial (different, same as current task). The mean RT and error data are depicted in **Figure 2**, left panel. As can be seen, switching to the task that was invalidly cued on the preceding trial impaired performance in both RTs and errors [$F(1, 19) = 4.04$, $p = 0.059$, and $F(1, 19) = 5.07$, $p = 0.036$, respectively]. This impairment did not differ for validly and invalidly cued trials [$F(1, 19) < 1$, and $F(1, 19) = 1.27$, $p = 0.273$, for RTs and errors].

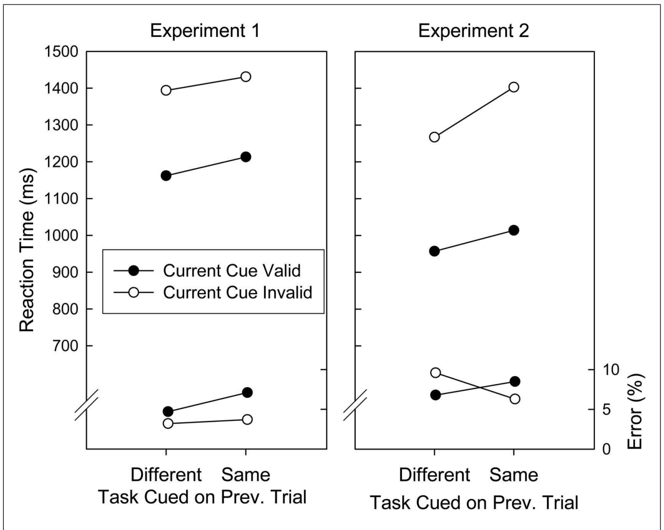

FIGURE 2 | Mean reaction times and error proportions of Experiment 1 and 2 for task switch trials following an invalidly cued trial as a function of cuing validity and whether the switch was made to the previously (invalidly) cued task.

DISCUSSION

Experiment 1 replicated costs of invalid task cuing previously found with other procedures (Ruthruff et al., 2001; Dreisbach et al., 2002; Hübner et al., 2004a). Interestingly, in the current study, the invalidity cost was confined to RTs, suggesting that participants adopted a response strategy which ensured constant performance accuracy across valid and invalid cuing conditions. As expected on the assumption of trial-to-trial adjustment of task preparation on the basis of its previous utility, the invalidity cost was reduced when the preceding trial was invalid, suggesting a general reluctance to engage in cue-based task preparation after having been misled.

Performance on trials associated with invalid cuing was particularly impaired when it implied switching to the task which was (invalidly) cued on the preceding trial. This extra cost is consistent with the assumption that an invalidly prepared task-set becomes subject to reactive inhibition. However, an alternative interpretation must be considered which relates to a confound with the specific cue sequence. More precisely, validly cued trials associated with switching to the previously cued task (e.g., aBaA with uppercase letters denoting tasks and lowercase letters denoting cues, respectively), and invalidly cued trials associated with switching to the remaining task (e.g., cBcA) are both necessarily associated with a cue repetition. That is, a lack of preparation following invalid switches should impair performance for valid task switches (aBaA), because here preparation would activate the correct task. Conversely, on invalid switches (following invalid switches, cBcA), in which preparation would activate an incorrect task-set, performance would be improved. Our results are thus also consistent with the assumption that participants are particularly reluctant to engage in cue-based preparation for the task which was invalidly cued on the preceding trial. It might also be conjectured that the lack of preparation is bound to usage of the previously misleading cue. However, the fact that Hübner et al. (2004a, Experiment 4) found increased switch costs after invalid cuing even when different cues were used to indicate the same task, argues against this possibility. Although our data do not allow deciding between reactive inhibition and task-specific lack of preparation, the fact that the invalidity cost manifested only in RTs whereas the extra cost of switching to the previously cued task occurred in both RTs and errors suggests that different processes may underlie the two effects.

EXPERIMENT 2

Experiment 2 closely resembled Experiment 1, the main modification being that, on each trial, the task was chosen randomly, resulting in an expected proportion of one third task repetition trials. The purpose of Experiment 2 was twofold. First, we wanted to replicate the sequential modulation of cuing validity obtained in Experiment 1. Second, extending the procedure to task repetitions allowed us to assess both the impact of current and previous invalid cuing on task repetition performance as well as to compare task switch performance when the task was invalidly cued as a (different) task switch vs. when it was cued as a task repetition. Noteworthy in this regard, in previous studies of cuing validity, in which only two tasks were used (or in which cues were used which could only be followed by two tasks, one of them constituting a task repetition), invalid cuing of a task switch implied an actual task repetition whereas invalid cuing of a task repetition implied an actual task switch. Inasmuch as preparation for task repetitions and switches involves different processes, comparing invalidity costs on task repetition and switch trials is confounded by this factor.

METHOD

Participants

Five female and 25 male students of the Helmut-Schmidt-University/University of the Federal Armed Forces Hamburg participated in exchange for partial course requirements. They ranged in age from 20 to 27 years.

Apparatus and stimuli

Apparatus and stimuli were the same as in Experiment 1 with the exception that responses were given by pressing one of two response keys which were mounted on an external rectangular keyboard (10 cm × 18 cm). The response keys extended 1.0 cm × 1.0 cm and were separated by 8.0 cm (parallel to the keyboards long axis). Participants pressed the response keys with the index or middle fingers of their left and right hand.

PROCEDURE

The procedure was identical to the procedure of Experiment 1 with the following exceptions. First, the task was chosen randomly on each trial, resulting in an expected proportion of one third task repetition trials. After an incorrect response, the identical trial was repeated. Such repetitions were discarded from the analyses and not counted as trials.

RESULTS

The same exclusion criteria as in Experiment 1 were applied. Additionally, to ensure identical preparation conditions, we excluded data from trials which were invalidly cued as a task repetition. We also excluded direct stimulus repetitions because these have been shown to facilitate responding selectively on trials in which the task or cue repeats (Hübner et al., 2004b).

In an ANOVA with repeated measures on the factors Validity on Current Trial (valid, invalid), Validity on Preceding Trial (valid, invalid), and Task Sequence (repetition, switch) on the mean RTs, all main effects were significant. There was an invalidity cost of 365 ms, $F(1, 29) = 222.12$, $p < 0.001$, post-invalid slowing of 70 ms, $F(1, 29) = 21.09$, $p < 0.001$, and a task switch cost of 168 ms, $F(1, 29) = 29.33$, $p < 0.001$. The invalidity cost was larger after a valid trial than after an invalid trial (482 vs. 247 ms), $F(1, 29) = 159.59$, $p < 0.001$. Furthermore, the invalidity cost was larger for task switches than for task repetitions (389 vs. 340 ms), $F(1, 29) = 11.11$, $p < 0.003$, and this was further modulated by a three-way interaction involving all factors, $F(1, 29) = 15.48$, $p < 0.001$. As can be seen in **Figure 3**, the reduction of the invalidity cost after an invalid predecessor trial was more pronounced on task repetition than on task switch trials. To examine this result pattern in more detail, planned comparisons were run, contrasting the invalidity cost on task repetition and task switch trials, separately. Both comparisons reached significance, $F(1, 29) = 6.37$, $p < 0.02$, and $F(1, 29) = 19.54$, $p < 0.001$, respectively, demonstrating that after a validly cued trial the invalidity cost affected task repetitions more strongly than task switches, whereas the opposite pattern occurred after an invalidly cued trial. No significant results were found in the corresponding error analysis (all ps > 0.24).

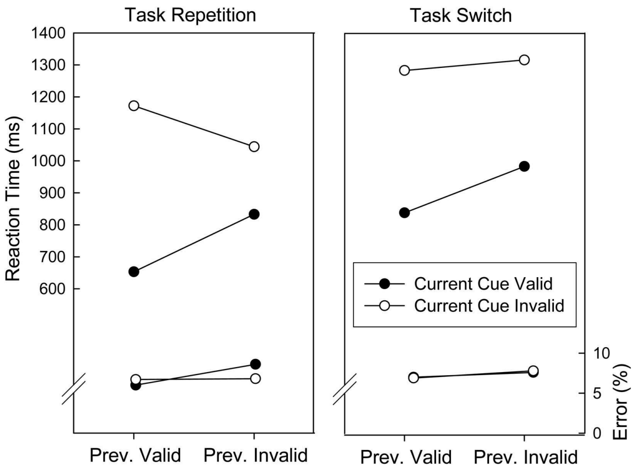

FIGURE 3 | Mean reaction times and error proportions of Experiment 2 as a function of cuing validity of the current and preceding trials and task sequence.

Analogously to Experiment 1, we ran an ANOVA, confined to data from task switch trials following an invalidly cued trial, with the factors Validity on Current Trial (valid, invalid) and Task (Invalidly) Cued on Preceding Trial (different, same as current task). The mean RT and error data are depicted in **Figure 2**, right panel. Regarding RTs, switching to a task that was invalidly cued on the preceding trial again impaired performance by 97 ms, $F(1, 19) = 14.60$, $p < 0.001$. This impairment did not differ for validly and invalidly cued trials, $F(1, 19) = 2.33$, $p = 0.138$. There were no significant effects in the error analysis (all ps > 0.20)[1].

DISCUSSION

Experiment 2 replicated the sequential modulation of cuing validity found in Experiment 1 and extended it to task repetition trials. Again, the invalidity cost was confined to RTs, whereas errors were kept at a constant level. Furthermore, Experiment 2 replicated the extra cost when switching to a previously invalidly cued task.

Intriguingly, the reduction of the invalidity cost after invalid trials was further modulated by task sequence. More precisely, task repetitions were associated with a larger invalidity cost after a valid trial and a smaller invalidity cost after an invalid trial than task

---

[1] To compare consequences of invalid preparation for a task repetition and for a task switch, we conducted an ANOVA involving only invalidly cued task switches with the factors Validity on Preceding Trial and Cuing (Task Repetition, Task Switch). Regarding RTs, trials which were invalidly cued as a task repetition were generally facilitated (by 74 ms after a valid predecessor trial and by 48 ms after an invalid predecessor trial), $F(1, 29) = 19.54$, $p < 0.001$. The two factors did not interact, $F < 1$. There were no significant effects in the corresponding error analysis (all Fs < 1).

switches. This modulation cannot be attributed to a difference in the preparation processes because at the time of preparation participants did not know whether a task repetition or switch would follow. By consequence, it has to be assumed that task repetitions are more strongly affected by the variation in cue-based preparation after valid and invalid trials than task switches. At first sight, it seems plausible to assume that a recently instantiated task-set may be more easily re-activated after being misled, thus predicting a lower invalidity cost for task repetitions than for task switches after an invalidly cued trial. However, findings of relative performance impairment when switching back to a task compared to when switching to a task not executed on previous trials (i.e., ABA vs. CBA task sequences) have been accounted for in terms of inhibition of the to-be-abandoned task-set (Mayr and Keele, 2000; Mayr, 2007; for a review see Koch et al., 2010). Assuming that backward inhibition is implemented during preparation, invalidity costs should be larger on task repetition trials (i.e., because inhibition has to be overcome) than on task switch trials. Given the broad empirical evidence that backward inhibition depends on appropriate preparation for the new task (Dreisbach et al., 2002, Experiment 5; Mayr and Keele, 2000; Hübner et al., 2003; Dreisbach and Haider, 2006; Kuhns et al., 2007, Experiment 3; see also Li and Dupuis, 2008), it seems conceivable that backward inhibition suffers from post-invalid reduction of preparatory activity. That is, the reduction of preparation after an invalidly cued trial (which normally goes along with the inhibition of the just executed task) then reduces backward inhibition accordingly. Backward inhibition might thus account for both the larger invalidity cost on task repetition trials after a valid predecessor trial (more invalid preparation and thus stronger backward inhibition on invalid trials), and the smaller invalidity cost on task repetitions after an invalid predecessor trial (less preparation and thus less backward inhibition on invalid trials).

GENERAL DISCUSSION

When people switch between simple cognitive tasks, performance benefits from advance cuing of the identity of the upcoming task. In particular, task performance increases with the probability of occurrence, suggesting that preparation is gradually adjusted to its expected utility. The current study provides evidence for adjustment of task preparation on the basis of its utility on the directly preceding trial. On both task switch (Experiment 1 and 2) and task repetition trials (Experiment 2) costs of invalid task cuing were strongly reduced when the directly preceding trial also involved an invalid cue. This sequential modulation resembles trial-to-trial adjustment effects regarding the processing of task-irrelevant stimulus features (Gratton et al., 1992; Botvinick et al., 2001) and suggests that participants engage less in cue-based task preparation processes after being misled.

Extending previous studies of invalid task cuing, we used a task switching paradigm with three tasks which allowed us to deconfound invalid task cuing and task sequence by assessing performance on both task repetition trials and on task switch trials after cuing of a task switch. Contrasting with previous results of additive or under additive interactions of cuing validity and task sequence (Ruthruff et al., 2001; Dreisbach et al., 2002; Hübner et al., 2004a) we observed an overall larger invalidity cost on task switches than on task repetitions. This interaction is difficult to interpret, however, given the modulation by previous cuing validity, that is, the fact that the reduction of the invalidity cost was more pronounced on task repetition than on task switch trials. A possible explanation is to assume that performance on invalidly cued task repetition trials is particularly impaired by anticipatory backward inhibition and that preparation following an invalidly cued trial lacks this component. Future research is necessary to disentangle preparatory activation of the set for an upcoming task and inhibition of the set of the preceding task in more detail.

In addition to assessing the sequential modulation of the invalidity cost, our experimental set-up allowed us to look more specifically at the performance on task switch trials following an invalidly cued trial when switching to a previously cued task vs. when switching to the remaining task. In both experiments, we observed an extra cost in the former case. This finding can be explained by particular reluctance to prepare for a task (or to use a task cue) which is associated with recent invalid preparation. Alternatively, it is conceivable that invalid task preparation enhances task competition which then triggers control measures of task-set inhibition (Hübner et al., 2004a).The results of the current study thus provide a new example for online, trial-to-trial adjustment of cognitive processing, clearly demonstrating that the degree of task preparation depends on previous success, possibly reflecting both anticipatory and reactive inhibition to ensure efficient performance regarding a currently relevant task in the face of exogenously and endogenously evoked competition.

ACKNOWLEDGMENTS

We thank Carmen Hefer and Falco Walther for data collection. The article processing fee was sponsored by Deutsche Forschungsgemeinschaft DFG.

REFERENCES

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., and Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychol. Rev.* 108, 624–652.

Dreisbach, G., and Haider, H. (2006). Preparatory adjustment of cognitive control in the task switching paradigm. *Psychon. Bull. Rev.* 13, 334–338.

Dreisbach, G., Haider, H., and Kluwe, R. H. (2002). Preparatory processes in the task-switching paradigm: evidence from the use of probability cues. *J. Exp. Psychol. Learn. Mem. Cogn.* 28, 468–483.

Gratton, G., Coles, M. G., and Donchin, E. (1992). Optimizing the use of information: strategic control of activation of responses. *J. Exp. Psychol. Gen.* 121, 480–506.

Hübner, M., Dreisbach, G., Haider, H., and Kluwe, R. H. (2003). Backward inhibition as a means of sequential task-set control: evidence for reduction of task competition. *J. Exp. Psychol. Learn. Mem. Cogn.* 29, 289–297.

Hübner, M., Kluwe, R. H., Luna-Rodriguez, A., and Peters, A. (2004a). Task preparation and stimulus-evoked competition. *Acta Psychol. (Amst)* 115, 211–34.

Hübner, M., Kluwe, R. H., Luna-Rodriguez, A., and Peters, A. (2004b). Response selection difficulty and asymmetrical costs of switching between tasks and stimuli: no evidence for an exogenous component of task-set reconfiguration. *J. Exp. Psychol. Hum. Percept. Perform.* 30, 1043–1063.

Kiesel, A., Steinhauser, M., Wendt, M., Falkenstein, M., Jost, K., Philipp, A. M., and Koch, I. (2010). Control and interference in task switching – a review. *Psychol. Bull.* 136, 849–874.

Koch, I., Gade, M., Schuch, S., and Philipp, A. M. (2010). The role of inhibition in task switching: a review. *Psychon. Bull. Rev.* 17, 1–14.

Kuhns, D., Lien, M.-C., and Ruthruff, E. (2007). Proactive versus reactive task-set inhibition: evidence from flanker compatibility effects. *Psychon. Bull. Rev.* 14, 977–983.

Li, K. Z., and Dupuis, K. (2008). Attentional switching in the sequential flanker task: age, location, and time course effects. *Acta Psychol. (Amst)* 127, 416–27.

Mayr, U. (2007). "Inhibition of task sets," in *Inhibition in Cognition*, eds G. David and C. M. MacLeod (Washington, DC: American Psychological Association),27–44.

Mayr, U., and Keele, S. W. (2000). Changing internal constraints on Action: the role of backward inhibition. *J. Exp. Psychol. Gen.* 129, 4–26.

Meiran, N. (1996). Reconfiguration of processing mode prior to task performance. *J. Exp. Psychol. Learn. Mem. Cogn.* 22, 1423–1442.

Ruthruff, E., Remington, R. W., and Johnston, J. C. (2001). Switching between simple cognitive tasks: the interaction of top-down and bottom-up factors. *J. Exp. Psychol. Hum. Percept. Perform.* 27, 1404–1419.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 11 June 2012; accepted: 23 July 2012; published online: 10 August 2012.

Citation: Wendt M, Luna-Rodriguez A, Reisenauer R, Jacobsen T and Dreisbach G (2012) Sequential modulation of cue use in the task switching paradigm. Front. Psychology ***3****:287. doi: 10.3389/fpsyg.2012.00287*

This article was submitted to Frontiers in Cognition, a specialty of Frontiers in Psychology.

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 12 February 2013
doi: 10.3389/fpsyg.2013.00047

Anticipatory regulation of action control in a Simon task: behavioral, electrophysiological, and fMRI correlates

Gamze Strack[1], **Christian Kaufmann**[1], **Stefanie Kehrer**[2], **Stephan Brandt**[2] **and Birgit Stürmer**[1]*

[1] *Department of Psychology, Humboldt-Universität zu Berlin, Berlin, Germany*
[2] *Department of Neurology, Berlin NeuroImaging Center, Berlin, Germany*

Edited by:
Rico Fischer, Technische Universität Dresden, Germany

Reviewed by:
Uta Wolfensteller, Technische Universität Dresden, Germany
Joseph A. King, Duke University, USA

Correspondence:
Birgit Stürmer, International Psychoanalytic University Berlin GmbH, Stromstr. 3, 10555 Berlin, Germany.
e-mail: birgit.stuermer@ipu-berlin.de

With the present study we investigated cue-induced preparation in a Simon task and measured electroencephalogram and functional magnetic resonance imaging (fMRI) data in two within-subjects sessions. Cues informed either about the upcoming (1) spatial stimulus-response compatibility (rule cues), or (2) the stimulus location (position cues), or (3) were non-informative. Only rule cues allowed anticipating the upcoming compatibility condition. Position cues allowed anticipation of the upcoming location of the Simon stimulus but not its compatibility condition. Rule cues elicited fastest and most accurate performance for both compatible and incompatible trials. The contingent negative variation (CNV) in the event-related potential (ERP) of the cue-target interval is an index of anticipatory preparation and was magnified after rule cues. The N2 in the post-target ERP as a measure of online action control was reduced in Simon trials after rule cues. Although compatible trials were faster than incompatible trials in all cue conditions only non-informative cues revealed a compatibility effect in additional indicators of Simon task conflict like accuracy and the N2. We thus conclude that rule cues induced anticipatory re-coding of the Simon task that did not involve cognitive conflict anymore. fMRI revealed that rule cues yielded more activation of the left rostral, dorsal, and ventral prefrontal cortex as well as the pre-SMA as compared to POS and NON-cues. Pre-SMA and ventrolateral prefrontal activation after rule cues correlated with the effective use of rule cues in behavioral performance. Position cues induced a smaller CNV effect and exhibited less prefrontal and pre-SMA contributions in fMRI. Our data point to the importance to disentangle different anticipatory adjustments that might also include the prevention of upcoming conflict via task re-coding.

Keywords: cognitive conflict, cueing, EEG, fMRI, pre-SMA, Simon task, anticipation, cognitive control

INTRODUCTION

Our ability to exert cognitive control in order to adjust ongoing performance to changing environmental conditions is essential for flexible behavior in everyday life. Whenever prior information or experience is available we attempt to avoid costs of inappropriate behavior (Kool et al., 2010). Such anticipatory processes are especially inevitable in settings calling for online control: the need for cognitive control elicited by preceding difficulties or errors engenders regulatory processes and people build subjective predictions about upcoming task demands on the basis of accumulating task knowledge. As such, cognitive control adjustments should be generally viewed along an anticipatory-online control continuum (Ullsperger and King, 2010). Yet, almost everything we know about action control refers to instantaneous online mechanisms like conflict detection and resolution. In contrast, little is known about the anticipation of these processes. Anticipatory regulation is a broad term and includes types of anticipatory control that may extensively differ in function. Recently, Ridderinkhof et al. (2010) conceptualized anticipatory regulation along two dimensions. The first dimension describes the point in time that triggers anticipatory regulation: *Reactive* anticipatory regulation is prompted by prior behavior or events such as preceding errors or cognitive conflicts. As well, anticipatory control regulation can be induced by prior information such as cues and is then of genuine *prospective* nature. The second dimension describes different types of adjustments that may accomplish both prospective as well as reactive anticipatory regulation: *Proactive*[1] adjustments may boost subsequent online conflict control by modifying the level of response or inhibition readiness. Or, alternatively, *preemptive* adjustments may to diminish or avoid the need for cognitive control by modifying the level or focus of selective attention. Such adjustments are obviously not only thinkable in the context of cognitive conflict but in all situations with high cognitive demand or increased error probability. Wühr and Kunde (2008) presented an excellent example of preemptive regulation in a Simon conflict task (Simon, 1967). In this task, spatially oriented responses are assigned to a non-spatial stimulus feature (e.g., stimulus figure). The task-irrelevant stimulus location, however, alters randomly and either matches or mismatches response location resulting in compatible and incompatible trials. Wühr and Kunde (2008) cued the compatibility of upcoming trials and showed that participants effectively used the actually task-irrelevant stimulus location instead of the

---

[1] Note that the term "proactive" has a different connotation in the dual routes framework of Braver et al. (2007).

task-relevant stimulus feature for response selection. Participants, therefore, changed their attentional focus. The aim of the present study was to further investigate prospective anticipatory regulation in a Simon task by using behavioral, electrophysiological, and hemodynamic measures.

Anticipatory regulation has previously been shown to improve behavioral performance in many situations involving cognitive conflict (Fassbender et al., 2006; Luks et al., 2007; Sohn et al., 2007; Aarts et al., 2008; Donohue et al., 2008; Alpay et al., 2009). Anticipatory processes can also be reflected by an electrophysiological measure, namely the contingent negative variation (CNV; Leuthold et al., 2004). This event-related potential (ERP) is observed during expectancy of an upcoming event (Walter, 1964). The terminal phase of the CNV prior to target onset reflects general preparation with sensory, motor, and cognitive shares depending on the particular task (e.g., Damen and Brunia, 1994; Fan et al., 2007). Yet, few CNV studies investigated the influence of higher-level processes like the anticipation of looming conflict. Fan et al. (2007) showed that cues eliciting higher unspecific response readiness enhanced the CNV amplitude in a conflict paradigm. We showed in a previous study that the CNV is susceptible for both cue-induced prospective anticipation and reactive anticipation due to the trial sequence (Alpay et al., 2009). More is known about post-target ERPs that indicate processes of online action control such as the N2 (Folstein and van Petten, 2008 for a review). This ERP deflection is a negative wave with a fronto-central distribution that usually peaks between 200 and 350 ms after onset of the imperative stimulus. The amplitude of the anterior N2 is magnified by processes involving cognitive control (Kopp et al., 1996; Heil et al., 2000; Nieuwenhuis et al., 2003; Falkenstein, 2006; Kehrer et al., 2009). The N2 has been associated with activation of the anterior cingular cortex (ACC), a ventrally located area within the posterior medial frontal cortex (pMFC; Nieuwenhuis et al., 2003; Ridderinkhof et al., 2004). The seminal conflict monitoring theory postulates that the ACC detects conflict and calls for top-down control processes to resolve concurrent response tendencies (Botvinick et al., 2004). There are several functional neuroimaging functional magnetic resonance imaging (fMRI) studies that provide evidence for an association of pMFC activation and online action control (Ridderinkhof et al., 2004 for a review). Rather few fMRI studies investigated *anticipatory* regulation of online action control. Some of them focused on the ACC and found respective anticipatory activation after explicit cueing of upcoming control demands (Sohn et al., 2007; Aarts et al., 2008). Anticipatory processes in these studies were of proactive nature and it remains an open question whether ACC activation can also be expected in preemptive anticipatory adjustments. The literature provides inconclusive results about the role of the ACC in anticipatory regulation since some studies did not find any preparatory ACC activation (MacDonald et al., 2000; Fassbender et al., 2006; Luks et al., 2007; Donohue et al., 2008). Another candidate region that might be also involved in anticipatory regulation of action control is the pre-SMA. Mars et al. (2009) and Neuhaus et al. (2010) assume that pre-SMA rather than ACC activation is associated with situations involving direct competition (Ullsperger and von Cramon, 2001), inhibition (Nachev et al., 2007), updating (Shima et al., 1996), or reprogramming (Isoda and Hikosaka, 2007) of actions. Using model-based fMRI that takes individual differences into account, Forstmann et al. (2008a) reported that the Response time (RT) distribution of response capture covaried with pre-SMA activation. Some researchers claim a key role for the pre-SMA in anticipatory control regulation. Hikosaka and Isoda (2010) concluded in their review that pre-SMA activation occurs when cues indicate a switch, whereas ACC activation occurs after error feedback. Ullsperger and King (2010) seized this idea, proposing that not only task switching but rather all processes of online action control can be more or less regulated by anticipation, and that underlying processes might be associated with pre-SMA activation.

In the present study we cued upcoming control demands in a Simon task in order to investigate how participants anticipate and which neural structures are associated with this anticipatory regulation. In particular, we were interested in whether these processes are performed by the ACC or the pre-SMA. Therefore, we employed a Simon task and presented one of three cue types prior to each Simon target that either (1) informed about the compatibility of the upcoming Simon target (rule cues), or (2) informed about the spatial position of the upcoming Simon stimulus (position cues), or (3) provided no information (non-informative cues). Rule cues were expected to induce prospective anticipatory regulation of action control and thus to be most beneficial for task completion. In contrast, position cues were assumed to trigger an anticipatory attentional shift to the correct stimulus location. Both rule and position cues reduced the stimulus set twofold while keeping the response set bivalent, i.e., no prediction of the response key was possible. Non-informative cues were introduced as a control condition that neither reduced the stimulus nor the response set. We additionally applied NoGo[2] trials in order to prevent preemptive adjustments such as the deduction of the correct response from the stimulus position (e.g., "compatible" means to press the key corresponding to the stimulus location). Such preemptive adjustments were indicated by behavioral measures in a cued Simon task in Wühr and Kunde (2008): participants shifted their attention from the task-relevant stimulus figure to the task-irrelevant stimulus position (e.g., a cue indicating an incompatible trial means a crossed response). We investigated additional measures that are indicative of conflict, e.g., the N2 and conditional accuracy functions (CAFs) in order come to a better understanding of the underlying adjustments in the present study. CAFs plot behavioral accuracy as a function of RT speed and usually show that fast responses tend to be more error-prone, especially in incompatible conditions (response capture, Ridderinkhof et al., 2010 for a review). Typically accuracy starts low and improves quickly within the fastest segment of RTs. According to Ridderinkhof et al. (2010) the slope between the first two bins in a CAF indexes the overcoming of response capture in incompatible trials (the steeper, the more response capture, that is, the more conflict). If rule cues lead to preemptive adjustments (circumvention of the original instruction) they should be associated with a lower N2 amplitude and less indication of

---

[2] Our previous study showed that NoGo-related processes did not modulate behavioral and electrophysiological cueing and Simon effects Alpay et al. (2009).

response capture. Non-informative cues, since being a measure of the unmodified Simon effect, were assumed to show the opposite pattern. Otherwise, if rule cues just modulate response readiness (proactive adjustments) the N2 should compare to that after non-informative cues. They were also expected to exhibit magnified N2 amplitudes for incompatible events (as compared to compatible events). By contrast, a N2 compatibility effect should be absent after rule cues if they trigger preemptive anticipatory regulation despite our NoGo manipulation. In order to get more insight into the neural basis of prospective anticipation we investigated electrophysiological and hemodynamic measures of pre-target processes. We expected rule cues to enhance the anticipatory pre-target CNV. At the hemodynamic level, we were interested in the neural networks that accomplish anticipatory regulation. We expected that proactive preparation to upcoming conflicts should be associated with ACC activation as reported in studies that investigated this type of anticipatory regulation (Sohn et al., 2007; Aarts et al., 2008). However, there is little research about networks involved in preemptive adjustments. One important candidate structure in such networks might be the pre-SMA because rule cues trigger the retrieval of relevant response contingencies and, thus, prospectively prepare for action selection (Rushworth, 2008; Ullsperger and King, 2010).

MATERIALS AND METHODS

PARTICIPANTS

Thirty-nine students participated in a first EEG session, whereas the data of eight participants had to be excluded afterward (six due to augmented ocular or movement-related artifacts, one due to extremely slow responses, one due to technical problems). The remaining 31 participants took part at a second session where fMRI was measured. One participant was excluded from further analysis because of incomplete behavioral data acquisition during the fMRI session. Two participants were excluded because of movement artifacts. All of the remaining 28 students (21 women, 7 men; 18–31 years old, mean age = 22.6 years) had normal or corrected-to-normal vision, were right-handed and reported no history of neurological, psychiatric, or major medical disorder. All participants were students from the Humboldt-Universität zu Berlin that either received course credits or were paid 8 EURO per hour for volunteering at the experimental procedures. They signed an informed consent prior to both experimental sessions. All procedures were previously approved by the ethical review board at the Charité university medical center.

PROCEDURE, STIMULI, AND DESIGN

Participants completed a first EEG session and, 4–6 weeks later, a second fMRI session. In the EEG session, participants were seated in a sound-attenuated chamber at a constant viewing distance of 100 cm to a 17″ TFT computer screen. They responded by pressing one of two response keys horizontally arranged on a table (distances: 20 cm to participant, 30 cm between keys). The timing of the task program that displayed stimuli and recorded behavioral responses was triggered every 2 s by simulated scanner main pulses, i.e., timing of the experiment was exactly the same during the EEG and fMRI sessions. In the fMRI session participants lay supine in the MR scanner. Imaging data were collected using a standard birdcage head coil. Vacuumed pillows were used to minimize head movements. Stimuli were projected on a screen mounted above the MRI head coil and could be viewed through an attached mirror. The main pulses of the scanner determined the timing of the task program that displayed stimuli and recorded behavioral responses (Presentation® software)[3]. Participants responded with their left or right index finger by pressing one of two optical response keys placed at their hands.

The trial procedure was identical in both sessions (**Figure 1**). Each trial consisted of a cue period (1 s), a delay period (5 s), a stimulus period (0.2 s), and a fixed time interval for the response (1.8 s, whereas responses later than 1 s were classified as too late). Stimulation was presented white on a dark gray background on a flat computer screen in the EEG chamber and via a back-projection screen in the MRI scanner. A white fixation dot (0.09° visual angle) marked the center of the screen as long as no cue stimulus was displayed.

A horizontal Simon task was combined with three different types of precues: rule cues (RULE), position cues (POS), and non-informative cues (NON). (1) RULE cues ("compatible," "incompatible") informed about the compatibility of the upcoming Simon task and, thus, were assumed to enable the anticipation of the subsequent control demand. (2) POS cues informed about the spatial location ("left," "right") of the upcoming Simon stimulus and enabled the anticipatory allocation of spatial attention. (3) A fifth cue was non-informative ("continue") and served as control condition that provides no information for preparatory processes. Cue stimuli were centrally displayed as German translations of the capitalized words compatible ("KOMPATIBEL"), incompatible ("INKOMPATIBEL"), left ("LINKS"), right ("RECHTS"), and continue ("WEITER"). The cues subtended a visual angle of $1° - 2.5°$ in horizontal and 0.23° in vertical orientation. Each Simon trial was randomly cued with one of these cues that were presented with equal probabilities. All cues were valid and the cue-target combinations were counterbalanced. Participants were asked to prepare for the subsequent Simon task by using the cue information as good as possible.

After the cue delay participants were shown one of three Simon stimuli that were randomly presented either left or right (0.5° visual angle) of fixation: one indicated a right-hand response, the second indicated a left-hand response, and the third was a NoGo. A white-filled square or diamond served as Simon Go stimuli, while a star indicated a NoGo (each 0.75° visual angle). The NoGo condition was randomly presented in 1/3 of all trials, resulting in equal probabilities for the occurrence the three stimuli. Participants were instructed to respond as fast and as accurate as possible.

Both experimental sessions consisted of blocks each containing 75 trials in a pseudorandom sequence (lasting approximately 13 min). Participants performed six blocks in the EEG session (450 trials, approximately 1.5 h recording time). In the fMRI session they completed four blocks (300 trials, approximately 55 min scanning time) each recorded as one run. That is, every cue ("KOMPATIBEL," "INKOMPATIBEL," "LINKS," "RECHTS,"

---

[3] www.neurobs.com

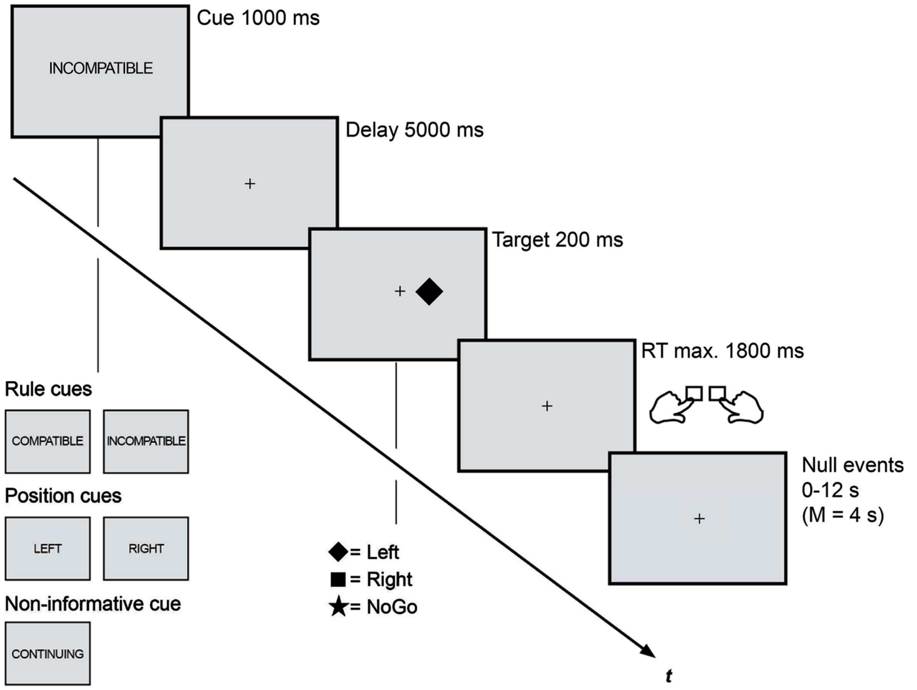

FIGURE 1 | Rule cues predicted compatibility and were for this reason expected to trigger anticipatory action control. These cues reduced the task set from four to two possible S-R assignments and were presented along two different control conditions: position cues that predicted the upcoming stimulus position and reduced the task set in the same amount enabled an attentional shift to the task-relevant visual half. Non-informative cues did not induce any anticipatory processes and lead not to a task set reduction. Trial procedure as well as timing was kept identical in the EEG and fMRI session.

"WEITER") was presented 90 (EEG) or 60 (fMRI) times while the subsequent stimulus type was either compatible, incompatible, or a NoGo with equal probabilities (no false RULE cues). We reduced the total amount of trials in the fMRI session to keep the scanning duration feasible. Opportunity for brief rest was given between blocks in both sessions. The trial order within every block was optimized with an algorithm designed to maximize the separability of the conditions in a rapid event-related fMRI design (optseq2; Dale, 1999). After the 2-s inter-trial interval, period of fixation lasting between 0 and 12 s, jittered in increments of 2 s (mean = 4 s), were interleaved with the experimental trials as determined by the optimization program. The order of runs and the stimulus-to-response assignment were counterbalanced among sessions and participants. Participants completed one 75-trial practice block prior to both sessions.

DATA ANALYSES

For all analyses trials with erroneous responses, trials immediately following errors and responses faster than 100 ms or slower than 1000 ms after target onset were discarded (RT, EEG) or modeled separately (fMRI). This reduced ERP data by 4.6% and fMRI data by 3.9%. NoGo targets were excluded from all analyses except for the behavioral analysis of false alarms. For RT and EEG analyses, ANOVAs are Huynh Feldt-corrected and *post hoc* comparisons Bonferroni-corrected. T-tests are two-tailed, if not mentioned otherwise.

BEHAVIORAL DATA

For RT distributional analysis RTs for each cue condition (RULE, POS, NON) and target condition (compatible, incompatible) were rank-ordered and divided into quartiles (four equal-sized bins). Mean RTs and accuracies for each condition and each quartile were computed. Conditional accuracy plots were created for each of the three cue conditions by plotting the accuracy of mean RTs for incompatible trials on the y-axis as a function of response speed on the x-axis (mean RTs for both compatibility conditions in quartiles). Slopes were calculated for the three delta plot segments determined by the data points of quartile 1 and 2 (slope 1), quartile 2 and 3 (slope 2), and quartile 3 and 4 (slope 3). ANOVAs conducted involved the factors cue (RULE, POS, NON) and slope (slope 1, slope 2, slope 3).

EEG RECORDING AND ANALYSIS

EEG was continuously recorded at 64 Ag/AgCl electrodes in an extended 10–20 system montage referenced to the participants' left mastoid. AFz served as ground electrode. The horizontal electrooculogram (EOG) was recorded from the outer canthi and vertical EOG was recorded from FP1 and below the left eye. All electrode impedances were kept below 5 kΩ. The EEGs and EOGs were recorded DC at a sampling rate of 1000 Hz and filtered online using a 250-Hz high cut-off. After recording the EEG was down-sampled offline to 250 Hz. Electrophysiological signals were recorded with Brain Vision Recorder and analyzed with Brain Vision Analyzer (Brain Products GmbH, Germany).

For the CNV analysis the signals were filtered offline with an additional low-pass filter of 5 Hz, 48 dB/oct. Artifacts with voltage steps exceeding 20 μV per sampling point were automatically removed. Cue-locked epochs of 9 s were created for each trial, starting 1 s before cue onset and ending 2 s after target onset. A time period of 1 s before cue onset was subtracted as baseline. Segments with amplitudes exceeding ±200 μV were automatically discarded from further analysis. In addition, trials were visually inspected and discarded if ocular artifacts occurred in the last 1 s before target onset (time interval of interest). The EEG epochs were averaged separately for each participant and cue condition. We analyzed the late CNV in a 1-s time interval immediately before target onset. An ANOVA was conducted containing 60 EEG electrodes and three cue types (RULE, POS, NON). A *post hoc* ANOVA additionally tested for compatibility after RULE cues (RULE prediction × electrode).

For the N2 analysis EEG was low-pass filtered with 30 Hz and high-pass filtered with 1 Hz, each with 48 dB/oct (time constant = 0.1592 s). Recorded signals were automatically removed when voltage steps exceeded 50 μV per sampling point, as well as when the difference between maximal and minimal activity fell below 0.50 μV within a 100-ms interval. All EEG channels were then submitted to an Infomax independent component analysis (ICA) algorithm for blink-correction. The ICA component reflecting an eye blink was identified for each subject excluded from signal synthesis of ICA sources to EEG channels. Other artifacts were eliminated semi-automatically. Target-locked segments of 9 s were created for each trial, starting 1 s before cue onset and ending 2 s after target onset. EEG epochs were averaged separately for each participant and each cue and compatibility condition. Since a pre-target baseline might be biased by cue-induced effects, we analyzed the N2 independently of a baseline following a peak-to-peak approach that has been introduced by Nieuwenhuis et al. (2003). Accordingly, peak-to-peak detection was determined for every condition in each participant in the Fz electrode (Nieuwenhuis et al., 2003): the N2 peak was automatically identified within time windows of 200–450 ms after stimulus onset. The N2 amplitude was then defined as the amplitude of this peak minus the immediately preceding positive peak (P2). Automatic peak detection was additionally visually inspected and corrected, if necessary. The ANOVA contained three cue types (RULE, POS, NON) and two Simon trial compatibilities (compatible, incompatible).

fMRI DATA ACQUISITION AND ANALYSES

Data were acquired at the Berlin NeuroImaging Center (Germany) on a 1.5-T MR scanner equipped with a circular-polarized head coil (Siemens Sonata, Erlangen, Germany) with an T2*-weighted single-shot gradient echo planar imaging sequence: 35 slices (interleaved), 3 mm isotropic resolution, 64 × 64 matrix, FOV = 192 mm, TE = 40 ms, TR = 2.00 s, flip angle = 90°, 1640 AC-PC oriented images for each run. Before functional runs, 176 anatomical T1-weighted slices were acquired (spatial resolution 1 mm × 1 mm × 1 mm, TR = 12.24 ms, TE = 3.56 ms, flip angle = 23°, 256 × 224 matrix; Deichmann, 2005). A vacuum head cushion was used to immobilize the participants' heads and necks in order to reduce movement artifacts. Earplugs were provided to attenuate background noise and additional headphones were used to communicate with subjects. Image preprocessing and analysis was carried out with SPM5 (Statistical Parametric Mapping)[4]. The first four volumes of each functional time series were discarded to avoid non-steady state effects caused by T1 saturation. Subsequently, motion correction estimation revealed that no subject showed more than 2 mm head movement (translation) and more than 1° of rotation during one run. After slice time correction of the functional data the anatomical data set was co-registered with the mean T2* image and T1-weighted images were segmented into gray matter, white matter, and cerebrospinal fluid. The gray matter of the co-registered structural image was spatially normalized to the standard template provided by the Montreal Neurological Institute (MNI) template using an automated spatial transformation (12-parameter affine transformation followed by non-linear iterations using 7 × 8 × 7 basis functions). The resulting transformation matrix was applied to the T2* data, and a resampling to a resolution of 3 mm × 3 mm × 3 mm voxel size was performed. Finally, the normalized images were smoothed with a Gaussian kernel (full width at half maximum) of 9 mm to create a locally weighted average of the surrounding voxels.

Statistical analyses were performed with a general-linear model approach. First, each subject was modeled independently. Five vectors of event onsets were created for model estimation, defining the experimental conditions of RULE cues (compatible, incompatible), POS cues (left, right), and NON-cues (continue). These pre-target effects were calculated locked to the cue onset. Additionally, error and post-error trials were modeled as one separate condition, although their quantity was insufficient for further analysis. The regressors were then convolved with a canonical hemodynamic response function (HRF) and employed as event-related regressors to model the BOLD responses within each experimental block. The HRF was combined with a temporal derivative as we assumed the peak response to vary in time. Six spatial realignment parameters served as additional regressors to remove signals correlated with head motion. Slow signal drifts were removed with a high-pass filter cut-off of 128 s. Model parameters were estimated using classical restricted maximum likelihood estimates. The estimation was made including a first-order autoregressive model in order to estimate temporal autocorrelations in the time series data and to correct for non-sphericity by adjusting the degrees of freedom appropriately. Voxelwise statistical parametric maps (SPM) were calculated for linear contrasts between regressors of interest for each subject. The resulting contrast images were submitted into a group analyses, treating subjects as random effects. Whole brain statistics were calculated for the contrasts RULE cue > NON-cue, POS cue > NON-cue, and RULE cue > POS cue (and their respective reversed contrasts) by performing one-sample t-tests. We additionally tested for effects of compatibility in RULE cues, RULE incompatible > RULE compatible. All results reported relate to activation averaged across voxels in clusters larger than 25 contiguous voxels meeting a threshold at $p < 0.05$, corrected for multiple measurements (FDR; Genovese et al., 2002). All resulting cluster maxima were converted to Talairach space[5] and entered into the

---

[4] www.fil.ion.ucl.ac.uk

[5] http://imaging.mrc-cbu.cam.ac.uk/imaging/MniTalairach

Talairach Damon (Lancaster et al., 2000) in order to determine the nearest anatomical loci. Pearson correlations between individual cluster activation in all frontal areas (Z-standardized beta-values) and the behavioral RULE cue benefit (Z-standardized NON-RULE RT difference) were calculated.

RESULTS

BEHAVIORAL RESULTS

Response times and accuracy data (error percentages) for all factor levels are displayed in **Table 1** and **Figure 2**. **Figure 3** shows the effect of accuracy as a function of RT dispersion.

Response times

An exploratory ANOVA containing both sessions revealed a non-significant trend between the sessions, $Fs < 3.68$, $ps > 0.06$. We, therefore, collapsed data across sessions. A typical Simon effect of 26 ms, $F(1,27) = 35.0$, $p < 0.001$, and a main effect of cue condition occurred, $F(2,54) = 70.78$, $p < 0.001$. RULE cues ($M = 498$ ms, SD $= 74$) enabled by 88 ms faster responses than NON-cues ($M = 586$ ms, SD $= 90$), $t(27) = 9.41$, $p < 0.001$, and by 53 ms faster than POS cues ($M = 552$ ms, SD $= 81$), $t(27) = 7.05$, $p < 0.001$. POS cues still triggered an RT benefit of 34 ms against NON-cues, $t(27) = 7.40$, $p < 0.001$. We calculated t-tests in order to examine whether these cueing benefits were present in both compatibility conditions: RULE against NON-was faster in compatible assignments, $t(27) = 9.44$, $p < 0.001$ as well as in incompatible assignments, $t(27) = 8.0$, $p < 0.001$, and POS versus NON-was also faster in both compatible, $t(27) = 4.67$, $p < 0.001$, and incompatible trials, $t(27) = 8.35$, $p < 0.001$. An overall interaction of compatibility and cue condition, $F(2,54) = 14.82$, $p < 0.001$, indicated a particularly pronounced RT difference between the compatibility conditions after RULE cues, $t(27) = 7.71$, $p < 0.001$. However, the compatibility effect was also significant for trials with POS cueing, $t(27) = 2.8$, $p < 0.05$, and for trials with NON-cueing, $t(27) = 2.66$, $p < 0.05$. In sum, RTs showed utilization of both RULE and POS cues since they were fastest after RULE cues, intermediate after POS cues, and slowest after NON-cues. In addition, compatibility effects were present in all cueing conditions while they were enhanced after RULE cues. However, this greater compatibility effect came about because RULE cues speeded up compatible assignments relatively more than incompatible trials (and not because incompatible responses were slowed down).

Table 1 | Means and standard deviations (SDs) of reaction times (RT) and percentages of error (PE) as a function of the factors compatibility (compatible, incompatible) and cue condition (RULE, rule cues; POS, position cues; NON, non-informative cues).

	RULE		POS		NON	
	RT	PE	RT	PE	RT	PE
COMPATIBLE						
Mean	474.9	3.2	545.3	5.8	577.0	5.6
SD	66.1	3.0	81.2	4.0	93.2	5.3
INCOMPATIBLE						
Mean	521.6	3.6	557.8	6.9	595.1	9.2
SD	83.9	2.8	83.1	4.0	90.4	5.2

Accuracy

The overall accuracy in the EEG session (6.12%) did not differ from that in the fMRI session (5.33%), $F(1,27) = 1.69$, $p = 0.20$. For this reason, data were collapsed across sessions. About 2.94% of NoGo trials were false alarms that were analyzed separately for cue effects. This analysis revealed that false alarms occurred more often after RULE cues than both after POS cues, $t(27) = 4.96$, $p < 0.001$, and after NON-cues, $t(27) = 5.13$, $p < 0.001$. The false alarm rate

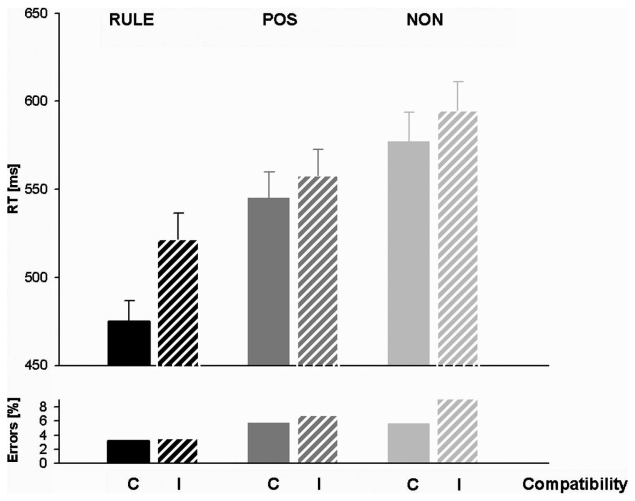

FIGURE 2 | Response times (RTs) and error percentages for rule cues (RULE), position cues (POS), and non-informative cues (NON) in compatible and incompatible trials. Note the particularly short RTs for rule cued compatible trials.

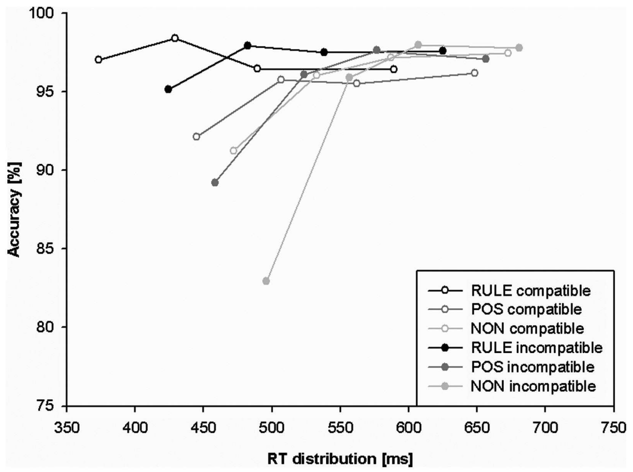

FIGURE 3 | Conditional accuracy function (CAF) plotting behavioral accuracy as a function of response speed for rule cues (RULE), position cues (POS), and non-informative cues (NON). The slope of the fastest portion (first segment) of RTs is conceived as a measure of response capture.

between POS and NON-cues did not differ, $t(27) = 1.65$, $p > 0.1$. Thus, the false alarm rate suggested that the NoGo manipulation that was introduced to prevent preemptive adjustments was less effective after RULE cues. Accuracy was further analyzed for all Simon go trials with a cue (RULE, POS, NON) × trial compatibility (C, I) within-subjects ANOVA. A main effect of cue condition, $F(2{,}54) = 20.98$, $p < 0.001$, indicated that trials with RULE cues entailed less errors (3.41%) than both trials with NON-cues (7.44%), $t(27) = 5.48$, $p < 0.001$, and trials with POS cues (6.32%), $t(27) = 6.03$, $p < 0.001$. The numerically higher accuracy for trials cued with POS cues against trials cued with NON-cues failed significance, $t(27) = 1.64$, $p = 0.11$. A main effect of compatibility, $F(1{,}27) = 5.70$, $p < 0.05$, was due to more errors in incompatible trials (6.56%) than in compatible ones (4.89%). The factors cue condition and compatibility interacted, $F(2{,}54) = 5.35$, $p < 0.01$, since compatibility affected accuracy only in trials with NON-cues, $t(27) = 3.14$, $p < 0.01$, but not in the other cue conditions, $t\text{s} < 1.1$. In sum, trials with RULE cues were accomplished most accurate and exhibited no compatibility effect in accuracy rates. The accuracy results therefore support the former notion that the higher compatibility effect in RTs for RULE cues might rather originate from a relatively greater cue benefit for compatible than for incompatible trials and not from a greater cognitive conflict (that would be associated with slower responses and lower accuracy). As opposed to this, trials with NON-cues showed typical conflict effects with lowered accuracy and a Simon effect in accuracy rates.

We analyzed accuracy as a function of RT dispersion in CAFs to investigate whether prospective anticipation induced by cues diminishes the impact of the misleading stimulus location. The strength of initial response capture in the Simon task is reflected in the frequency of fast errors that are thought to indicate stronger initial capture (Ridderinkhof et al., 2010 for a review). The slope value between the two fastest RT segments indexes the strength of initial response capture. We predicted that RULE cues should be associated with least response capture while NON-cues should be associated with most response capture. We submitted the slope between the first two fastest segments of the CAF to a repeated-measures ANOVA including the factors cue condition and compatibility. Overall main effects were present for cues, $F(2{,}54) = 21.04$, $p < 0.001$, as well as for compatibility, $F(1{,}27) = 22.32$, $p < 0.001$. The main effect for compatibility was induced by a steeper positive-going CAF slope for incompatible than compatible trials, $t(27) = 4.7$, $p < 0.001$ (indicating more response capture). The main effect of cue condition was due to a steeper positive slope for trials with NON-cues than for trials with RULE cues, $t(27) = 6.24$, $p < 0.001$, as well as for trials with NON-cues than trials with POS cues, $t(27) = 2.88$, $p < 0.05$. The slope of trials with POS cues was also more positive-going than the slope of events with RULE cues, $t(27) = 4.067$, $p = 0.001$. An overall interaction of compatibility and cue condition, $F(2{,}54) = 12.85$, $p < 0.001$, indicated that the latter differences between cue conditions did not occur in compatible trials, $t\text{s}(27) < 2.2$, $p\text{s} > 0.1$, but rather in incompatible trials: $t(27) > 3.39$, $p\text{s} < 0.01$. Compatibility effects in CAF slopes were strongest for NON-trials, $t(27) = 5.37$, $p < 0.001$, while weaker in POS trials, $t(27) = 2.57$, $p = 0.05$, and not significant in RULE trials, $t(27) = 1.36$, $p > 0.5$. Taken together, response capture and cognitive conflict, respectively, as indicated by CAF slopes were only present in incompatible trials that exhibited strongest effects for NON-cues, intermediate effects for POS cues, and smallest effects for RULE cues.

ELECTROPHYSIOLOGICAL RESULTS

An ANOVA of the cue-locked CNV amplitude (see **Figure 4**) including all electrodes and cue types revealed an interaction of electrode and cue condition for the time window 1 s before target onset, $F(118{,}3186) = 4.09$, $p < 0.001$. Additional ANOVAs each comparing two cue conditions showed that RULE cues elicited a greater CNV compared to NON-cues across all electrodes, $F(59{,}1593) = 5.62$, $p = 0.001$. RULE cues also generated overall more negativity compared to POS cues, $F(59{,}1593) = 3.70$, $p < 0.01$. The CNV amplitude for NON-was less pronounced than for POS, $F(59, 1593) = 2.26$, $p < 0.05$. A *post hoc t*-test for compatibility was calculated solely for RULE cues because only in this condition compatibility was predicted prior target onset: this test revealed no compatibility effect in the anticipatory CNV after RULE cues, $F < 1$.

An ANOVA of N2 at the Fz electrode (**Figure 5A**) resulted in a main effect of cue condition, $F(2{,}54) = 12.62$, $p < 0.001$. RULE cues reduced the magnitude of the N2 as compared to NON-cues, $t(27) = 4.30$, $p < 0.001$, and as compared to POS cues, $t(27) = 4.13$, $p < 0.001$. POS cues reduced the N2 amplitude as compared to NON-cues numerically, however, this effect failed significance, $t(27) = 1.83$, $p = 0.08$. The main effect of compatibility, $F(1{,}27) = 2.05$, $p = 0.16$, as well as the interaction between compatibility and cue condition were not significant, $F(2{,}54) = 2.29$, $p = 0.11$. In order to test our *a priori* hypothesis that incompatible trials should provoke a larger N2 amplitude than compatible trials, we ran a t-test. Such a compatibility effect was present in trials with NON-cues, $t(27) = 2.98$, $p < 0.01$, while it was absent in trials with both informative cue types, $F\text{s} < 1$ (see **Figure 5B**).

In sum, the pre-target CNV amplitude mirrored cue utilization since it was largest with RULE cues, intermediate with POS cues, and least pronounced with NON-cues. Interestingly, the results were reversed for the subsequent conflict-related N2 magnitude

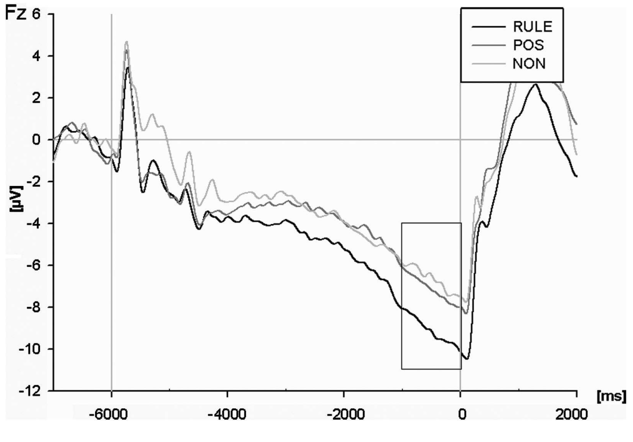

FIGURE 4 | The contingent negative variation (CNV) at the Fz electrode for rule cues (RULE), position cues (POS), and non-informative cues (NON). The analysis refers to the terminal second before target onset.

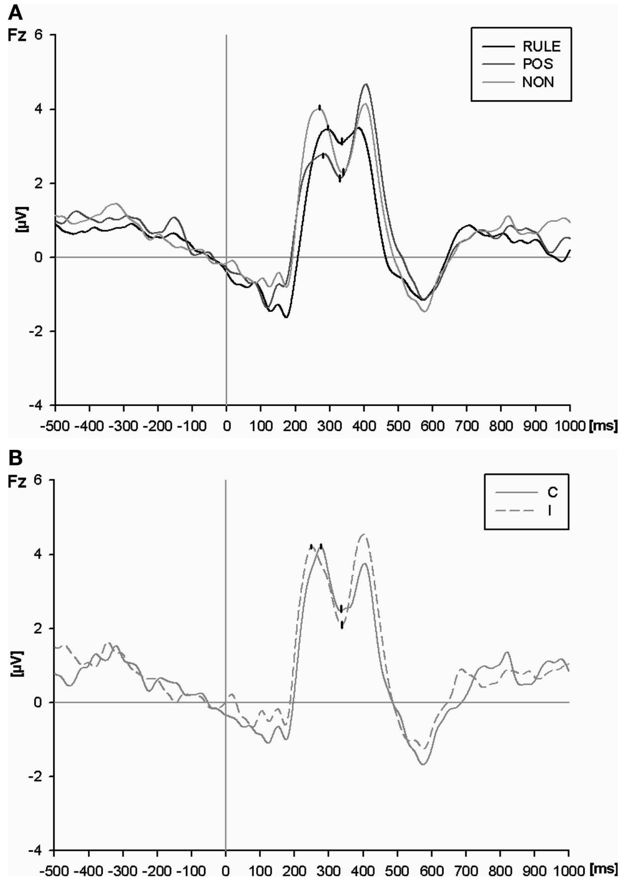

FIGURE 5 | (A) The N2 at the Fz electrode for rule cues (RULE), position cues (POS), and non-informative cues (NON) averaged across compatibility conditions. P2 and N2 peaks as determined by automatic peak detection are highlighted. **(B)** The P2 and N2 peaks at the Fz electrode for non-informative cues (NON) plotted separately for compatible and incompatible events.

(however, the N2 after POS cues was numerically intermediate but differed not significantly from NON-cues). In addition, a compatibility effect was only present in the N2 after NON-cues as expected. Thus, electrophysiological results show conflict indication for NON-cues but not for RULE cues.

fMRI RESULTS

Since we were interested in differentiating networks that are associated with particular cue information we ran pairwise t-tests for the activation following the cue onset. Contrasts comparing such anticipatory activation were *RULE cue > NON-cue* (results in **Table 2**), *POS cue > NON-cue* (**Table 3**), and *RULE cue > POS cue* (**Table 4**). We also calculated a contrast between rule cues predicting incompatible and those predicting compatible trials, *RULE incompatible > RULE compatible.*

RULE cue > NON-cue

The contrast of *RULE cue > NON-cue* highlighted a widespread fronto-posterior network of RULE cue-induced anticipation including frontal areas such as the left lateral rostral prefrontal cortex (rPFC), left posterior vlPFC, left dorsolateral prefrontal cortex (dlPFC), and the pre-SMA (see **Figure 6**; **Table 2** for a complete list that contains also temporal, occipital, basal ganglia, and thalamic activation). The reversed contrast *NON-cue > RULE cue* involved the medial rPFC, right dlPFC, and the ACC among other activations (see **Table 2**).

We tested whether the individual fMRI activation in the frontal areas resulting from the whole brain *RULE cue > NON-cue* contrast correlated with the individual RULE cue-induced behavioral benefit (calculated as the Z-standardized individual difference between RTs for NON-cues minus RTs for RULE cues). In fact, the success of RULE cue implementation correlated with the cluster activation of the pre-SMA (Talairach coordinates: 7, 5, 53), $r = 0.38$, $p < 0.05$, and the vlPFC (Talairach coordinates: −46, 0, 17), $r = 0.42$, $p < 0.05$, while no such correlations were found for rPFC and dlPFC.

POS cue > NON-cue

In contrast to the activation pattern revealed for RULE cue > NON-cue contrast, *POS cues* activated premotor regions and bilateral regions of posterior cortex and occipital cortex. No activated regions were found in the reversed *NON-cue > POS cue* contrast (**Table 3**).

RULE cue > POS cue

Like in the RULE > NON-contrast, areas more activated in RULE cues than in POS cues involved the left lateral rPFC, left posterior vlPFC, left dlPFC, and the pre-SMA (**Figure 6**; **Table 4**). As can be seen in **Table 4**, the activated neural network also involves temporal and parietal areas as well as basal ganglia and the insula. The reversed contrast, POS cue > RULE cue revealed anterior and posterior cingular and bilateral dlPFC activation among other premotor, motor, temporal, parietal, occipital, and cerebellar regions.

RULE incompatible > RULE compatible

No effect of anticipated compatibility was obtained for this or the reversed contrast.

GENERAL DISCUSSION

The present study aimed at investigating the neural underpinnings of anticipation of action selection within a conflict task. To this end, EEG and fMRI measures were recorded in a Simon task combined with cues that predicted compatibility (rule cues), stimulus location (position cues), or were non-informative. Behavioral results established that rule and position cues were utilized for both compatible and incompatible assignments. Compatible assignments were faster accomplished than incompatible assignments in all cue conditions. Rule cues yielded the strongest benefit for RTs and accuracy. Position cues entailed faster responses than non-informative cues for compatible and incompatible trials as well, although these cue benefits were smaller than those of rule cues. Moreover, accuracy as a function of response time (CAFs) measured response priming by the irrelevant stimulus location (response capture) was strongest after non-informative cues, intermediate after position cues, and weakest after rule cues. Rule cues, therefore, seemed to lead to anticipatory adjustments that prevented response capture. As a measure of anticipatory processes

Table 2 | Maxima of activation beginning with the onset of rule cues versus non-informative cues activation maxima.

Anatomical area	Cluster size	Hemisphere	*x*	*y*	*z*	*Z* value
RULE CUE > NON-CUE						
Middle frontal G. (BA 10)	47	L	−26	46	8	3.70
Middle frontal G. (BA 10)	37	L	−34	45	21	3.06
Middle frontal G. (BA 9)		L	−34	33	23	2.76
Precentral G. (BA 6)	1070	L	−27	−12	49	5.86
Precentral G. (BA 6)		R	23	−15	49	5.11
Pre-SMA/medial frontal G. (BA 6)		R	7	5	53	4.90
Putamen	456	L	−23	4	7	4.73
Inferior frontal G. (BA 44)		L	−46	0	17	4.42
Putamen	141	R	24	15	1	4.02
Caudate		R	15	2	18	3.14
Thalamus		R	10	−2	7	2.99
Thalamus	26	R	7	−24	−1	4.25
Middle temporal G.	28	L	−54	−35	−3	3.02
Superior temporal G. (BA 22)		L	−46	−38	5	2.81
Inferior parietal L. (BA 40)	795	L	−35	−44	37	5.30
Superior parietal L. (BA 7)		L	−30	−59	44	4.93
Superior parietal L. (BA 7)	437	R	32	−57	48	4.91
Supramarginal G. (BA 40)		R	40	−39	36	4.37
Inferior occipital G. (BA 17)	432	L	−15	−91	−10	4.61
lingual G. (BA 17)		R	15	−83	−1	4.04
Lingual G. (BA 18)		L	−15	−79	−9	4.02
Middle occipital G. (BA 19)	56	R	29	−87	7	3.87
NON-CUE > RULE CUE						
Medial frontal G. (BA 10)	91	R	2	56	18	4.07
Superior frontal G. (BA 9)		R	4	55	28	3.82
Superior frontal G. (BA 8)	66	R	21	28	48	5.28
Precuneus (BA 31)	359	R	7	−49	29	4.74
Cingulate G. (BA 31)		L	−2	−36	41	4.03
Anterior cingulate G. (BA 32)		R	2	43	3	3.18
Amygdala	31	R	21	−3	−12	4.88
Parahippocampal G. (BA 35)	96	L	−29	−28	−18	5.31

Results are reported in Talairach coordinates for peak voxel activations after a False Discovery Rate (FDR) correction ($p < 0.05$; minimum size of each cluster was 25 contingent voxels). Indented rows indicate subordinate clusters. Hemispheres: R (Right) or L (Left); Abbreviations: BA, Brodmann Area; G, Gyrus; L, Lobule.

the CNV was largest with rule cues. The post-target N2 indicates online action control (Folstein and van Petten, 2008) that was reduced by rule cues as compared to the other cue conditions. Both rule and position cues canceled out the difference between incompatible and compatible assignments in the conflict-related N2. Non-informative cues were the only condition that generated a compatibility effect in the N2. This N2 result resembles our behavioral accuracy data that also exhibited a compatibility effect solely after non-informative cues and in addition indicated a particularly enhanced response capture in distributional analyses. Taken together, these effects suggest the existence of cognitive conflict after non-informative cues. Viewed in this light, it is questionable whether compatibility effects after rule cues and position cues can be seen as typical Simon or conflict effects. For both informative cues the entire data pattern does not show any conflict-specific pre- or post-target effects besides the behavioral difference in response speed between the compatibility conditions. The overall short RTs and high accuracy (even in the fastest segment of response times) suggest that rule cues considerably simplified the task especially for compatible trials. Rule cues apparently provoked anticipatory adjustments that circumvented conflict in upcoming Simon trials. In our paradigm, rule cues reduced the task set from four to two possible stimulus-response (S-R) assignments. In particular, with rule cues participants knew whether to respond on the same or opposite direction of the stimulus location. Participants might thus have translated the instruction into more effective condition-action rules by first excluding NoGo trials and then responding according to the target position. Wühr and Kunde (2008) have previously shown that participants use such re-coding to circumvent conflict in two-choice Simon task. In fact, we introduced NoGos to get a three-choice task in order to avoid such a task reconfiguration. However, this manipulation was most probably not effective. This reasonable suspicion was further supported because false alarms occurred more often for NoGos after rule cues as compared to the other cue conditions indicating a higher readiness to respond after rule cues. The RT difference between compatible and incompatible Simon trials after rule cues may as such not be due to Simon conflict but might simply come about by differences in response selection complexity. In incompatible trials after rule cues one has to respond opposite to stimulus location which is a more complex response selection rule than selecting the response according to stimulus position in compatible trials. In a similar vein, position cues minimized the uncertainty about the stimulus position and allowed to anticipate the upcoming target location. Anticipatory attentional allocation to the target position might have reduced the cause for the Simon conflict, namely spatial uncertainty. However, position cues did not allow for a condition-action rule remapping like rule cues, because choice responses were still due to stimulus figure which was not known in advance. The latter fact might explain why the numerical reduction of the N2 induced by position cues against non-informative cues failed significance.

In contrast to present findings rule cues accelerated only compatible not incompatible assignments in our former study (Alpay et al., 2009). Two possible reasons can account for this difference: first, the 1.5-s cue-target interval in the former study (as compared to 6 s in the present study) may have been insufficient for preemptive anticipatory regulation in incompatible Simon trials, which need a translation into a more complex response selection rules as compared to compatible trials. Second, the complexity of the translation may have been additionally aggravated by the vertical design in the former study as compared to the more natural horizontal design that relates to bilateral body and brain symmetries in the present study (Vallesi et al., 2005). Most probably for the same reasons, our former study could also not reveal any rule cue-specific CNV modulation (Alpay et al., 2009) which we clearly observed in the present study.

The fMRI results indicated differences between anticipatory processes triggered by rule and position cues. A key finding was that rule cues induced more activation of the left lateral rPFC, left posterior vlPFC, left dlPFC, and the pre-SMA as compared to both position and non-informative cues. The rPFC has been proposed to enable other prefrontal regions to assist whenever rules have to be elaborated on a higher order level or when task management is in demand (Koechlin et al., 1999; Christoff and Gabrieli, 2000; Sakai and Passingham, 2003, 2006; Ramnani and Owen, 2004; Badre and D'Esposito, 2007; Wolfensteller and von Cramon, 2010, 2011). In the present study, rule cues were associated with more lateral rPFC activation than non-informative cues while the reversed contrast involved medial rPFC activation. Lateral activations of the rPFC are mostly associated with the maintenance and/or retrieval of task-relevant information while medial activation are present in studies investigating internal attending (Gilbert et al., 2006, for a meta-analysis). Rule-based response selection, especially when the task involves inhibitory or complex rules, has been consistently related to dlPFC activation (Sakai and Passingham, 2003; Bunge and Souza, 2008). Some studies also report anticipatory dlPFC activation (Fassbender et al., 2006; Luks et al., 2007) although the conflict monitoring theory posits that it is the resolution of cognitive conflict that takes place in the dlPFC (Botvinick et al., 2001).

Table 3 | Maxima of activation beginning with the presentation of position cues versus non-informative cues.

Anatomical area	Cluster size	Hemisphere	*x*	*y*	*z*	*Z* value
POS CUE > NON-CUE						
Putamen	27	L	−23	3	12	3.47
Precentral G. (BA 6)	182	L	−24	−12	49	4.24
Precuneus (BA 7)	164	R	18	−66	50	4.49
Superior parietal L. (BA 7)	255	L	−19	−63	52	4.34
Middle occipital G. (BA 19)	1620	R	26	−84	8	5.36
Lingual G. (BA 18)		L	−12	−85	−12	5.32
Cerebellum (declive)		L	−23	−76	−17	5.06
NON-CUE > POS CUE						
No suprathreshold clusters						

Results are reported in Talairach coordinates for peak voxel activations after a False Discovery Rate (FDR) correction ($p < 0.05$; minimum size of each cluster was 25 contingent voxels). Indented rows indicate subordinate clusters. Hemispheres: R (Right) or L (Left); Abbreviations: BA, Brodmann Area; G, Gyrus; L, Lobule.

Interestingly, the pre-SMA and vlPFC activation correlated with the behavioral benefit participants derived from rule cues. This finding points to the importance of these areas for the effective preparation of specific response contingencies. The left posterior vlPFC and pre-SMA have been previously associated with the maintenance of task-relevant knowledge that is used to guide subsequent behavior (for a review, see Bunge, 2004). Participants can mentally rehearse response contingencies using phonological codes while they can also prepare to respond with one or more effectors, by maintaining relevant high-level (i.e., relatively abstract) action representations (Bunge, 2004). Phonological rehearsal during task preparation has been claimed to be associated with activation of the left posterior vlPFC (Smith and Jonides, 1999; Bunge et al., 2003; Wagner et al., 2004) while abstract action representation rather involves the pre-SMA (Hazeltine et al., 2000). It is plausible that participants internally rehearsed their re-coded rules and at the same time maintained associated action codes in the present study. According to previous studies, pre-SMA activation seems to be involved in voluntary prospective action control (Sumner et al., 2007) representing action intentions (Lau et al., 2004) and initiating action sequences (Kennerley et al., 2004) rather than actual movements. As well, the pre-SMA is involved when S-R associations have to be re-learned or reversed (Nakamura et al., 1998), and when response competition is present (Milham et al., 2001; Derrfuss et al., 2004; Kennerley et al., 2004; Nachev et al., 2007; Taylor et al., 2007). Recently, it has been discussed whether the pre-SMA might be associated with prospective anticipatory regulation by selectively preparing the appropriate task set and triggering adaptation to conflict (Hikosaka and Isoda, 2010; King et al., 2010; Ullsperger and King, 2010). Importantly, the present data do not necessarily suggest a specific role of the pre-SMA and vlPFC in conflict control but rather indicate that these

Table 4 | Maxima of activation beginning with the onset of rule cues versus position cues.

Anatomical area	Cluster size	Hemisphere	*x*	*y*	*z*	*Z* value
RULE CUE > POS CUE						
Middle frontal G. (BA 9)	80	L	−35	21	30	4.34
Middle frontal G. (BA 10)		L	−37	37	18	3.49
Precentral G. (BA 9)	28	R	38	23	37	3.49
Middle frontal G. (BA 9)		R	32	33	24	3.40
Inferior frontal G. (BA 44)	270	L	−51	9	15	4.88
Claustrum		L	−29	18	5	4.71
Insula		L	−37	13	2	4.03
Claustrum	93	R	30	18	4	3.97
Putamen		R	21	10	3	3.93
Insula		R	38	18	7	3.90
Pre-SMA (BA 6)	60	L	−10	−2	66	3.71
Superior frontal G. (BA 6)		L	−4	−9	52	3.25
Middle temporal G. (BA 21)	56	L	−51	−41	5	3.56
Middle temporal G. (BA 22)		L	−54	−52	4	3.32
Inferior parietal L. (BA 40)	98	R	37	−54	43	4.74
Inferior parietal L. (BA 40)	475	L	−41	−50	42	5.86
Supramarginal G. (BA 40)		L	−54	−46	26	4.23
Precuneus (BA 7)	41	L	−10	−72	35	3.68
POS CUE > RULE CUE						
Anterior cingulate (BA 24)	650	L	−9	33	−1	4.00
Medial frontal G. (BA 9)		R	5	45	14	3.95
Medial frontal G. (BA 9)		L	−12	44	22	3.90
Superior frontal G. (BA 8)	145	R	23	27	51	4.60
Middle frontal G. (BA 8)		R	24	14	41	3.80
Superior frontal G. (BA 6)		R	12	33	51	3.11
Media frontal G. (BA 32)	61	L	−21	12	40	4.60
Posterior cingulate (BA 23)	663	R	7	−57	18	4.76
Precuneus (BA 7)		R	1	−48	43	4.61
Posterior cingulate (BA 30)		L	−15	−53	12	4.54
Middle temporal G. (BA 39)	69	R	48	−66	26	3.91
Middle occipital G. (BA 18)	458	R	26	−84	5	4.52
Cerebellum (declive)		R	27	−63	−9	4.40
Lingual G. (BA 18)		R	29	−74	−8	4.37
Cerebellum (declive)	382	L	−18	−84	−15	4.09
Fusiform G. (BA 19)		L	−29	−74	−11	4.07
Middle Occipital G. (BA 18)		L	−24	−86	1	3.80

Results are reported in Talairach coordinates for peak voxel activations after a False Discovery Rate (FDR) correction ($p < 0.05$; minimum size of each cluster was 25 contingent voxels). Indented rows indicate subordinate clusters. Hemispheres: R (Right) or L (Left); Abbreviations: BA, Brodmann Area; G, Gyrus; L, Lobule.

areas might be activated whenever prospective task reconfigurations are applied in order to reduce complexity or computational load during task implementation.

Furthermore, no indications of compatibility-specific processes within rule cues were found in the present study. The anticipation of an incompatible trial did not differ from the anticipation of a compatible trial (see also Forstmann et al., 2008b, for similar results in a Simon task fMRI study). This holds not only true for CNV and fMRI measures in the present study but also for the CNV results in our former study (Alpay et al., 2009). As mentioned before, participants re-coded the task by shifting their attentional focus on target position instead of target figure; hence the absence of anticipatory conflict-specific processes may not be surprising. It can rather be seen as an additional evidence for how effectively such preemptive mechanisms can prevent potentially effortful situations. The few studies that previously investigated anticipatory high-conflict versus low conflict effects mostly focused on predictions based on the conflict monitoring account. Some found anticipatory high versus low conflict ACC activation while others did not (Fassbender et al., 2006; Luks et al., 2007; Sohn et al., 2007; Aarts et al., 2008; Donohue et al., 2008). A reasonable explanation for these disparate findings is that anticipatory

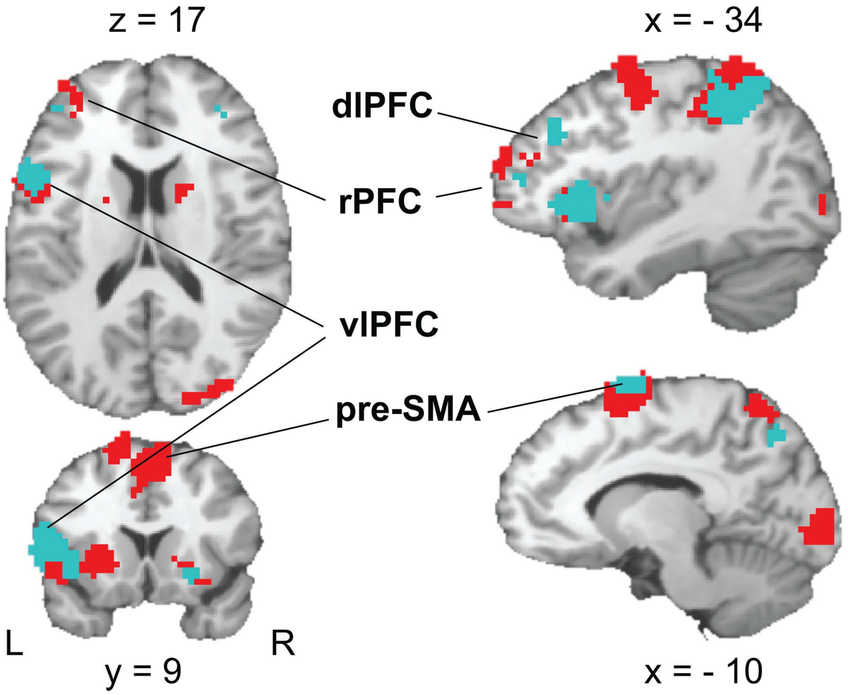

FIGURE 6 | Activation patterns in brain areas associated with rule cue-induced pre-target processes. A FDR-corrected ($p < 0.05$, $k \geq 25$ voxels) T-maps contrasting RULE > NON (red color) and RULE > POS (cyan color, superimposed) are plotted on a single subject Colin brain in MNI space (highlighted areas of interest correspond to Talairach peak voxel coordinates in **Tables 2** and **4**).

action regulation can be accomplished through different types of anticipatory adjustments. In particular, high versus low conflict conditions may trigger different usage of proactive or preemptive adjustments depending on the costs and benefits in terms of cognitive effort. In a similar vein prospective anticipatory adjustments are investigated in the field of task switching. A recent review of brain networks accomplishing task preparation showed a heterogeneous pattern of brain areas related to anticipatory regulation across studies (Ruge et al., 2013). The authors argue for different preparatory regulation modes that, e.g., focus on action-related or attention-related sets. The key for a better understanding of anticipatory control may lie in a careful separation of the actual underlying regulatory processes such as preemptive and proactive adjustments.

ACKNOWLEDGMENTS

This research was supported by a grant of the German Research Foundation to Birgit Stürmer (STU248/3-1).

REFERENCES

Aarts, E., Roelofs, A., and van Turennout, M. (2008). Anticipatory activity in anterior cingulate cortex can be independent of conflict and error likelihood. *J. Neurosci.* 28, 4671–4678.

Alpay, G., Goerke, M., and Stürmer, B. (2009). Precueing imminent conflict does not override sequence-dependent interference adaptation. *Psychol. Res.* 73, 803–816.

Badre, D., and D'Esposito, M. (2007). FMRI evidence for a hierarchical organization of the prefrontal cortex. *J. Cogn. Neurosci.* 19, 1–18.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., and Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychol. Rev.* 108, 624–652.

Botvinick, M. M., Cohen, J. D., and Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: an update. *Trends Cogn. Sci. (Regul. Ed.)* 8, 539–546.

Braver, T. S., Gray, J. R., and Burgess, G. C. (2007). "Explaining the many varieties of working memory variation: dual mechanisms of cognitive control," in *Variation in Working Memory*, eds A. R. A. Conway, M. J. Kane, A. Miyake, and J. N. Towse (New York: Oxford University Press), 76–106.

Bunge, S. A. (2004). How we use rules to select actions: a review of evidence from cognitive neuroscience. *Cogn. Affect. Behav. Neurosci.* 4, 564–579.

Bunge, S. A., Kahn, I., Wallis, J. D., Miller, E. K., and Wagner, A. D. (2003). Neural circuits subserving the retrieval and maintenance of abstract rules. *J. Neurophysiol.* 90, 3419–3428.

Bunge, S. A., and Souza, M. J. (2008). "Neural representations used to specify action," in *Neuroscience of Rule-Guided Behavior*, eds S. A. Bunge and J. D. Wallis (New York: Oxford University Press), 45–65.

Christoff, K., and Gabrieli, J. D. E. (2000). The frontopolar cortex and human cognition: evidence for a rostrocaudal hierarchical organization within the human prefrontal cortex. *Psychobiology* 28, 168–186.

Dale, A. M. (1999). Optimal experimental design for event-related fMRI. *Hum. Brain Mapp.* 8, 109–114.

Damen, E. J. P., and Brunia, C. H. M. (1994). Is a stimulus conveying task-relevant information a sufficient condition to elicit a stimulus-preceding negativity? *Psychophysiology* 31, 129–139.

Deichmann, R. (2005). Optimized RF excitation for anatomical brain imaging of the occipital lobe using the 3D MDEFT sequence and a surface transmit coil. *Magn. Reson. Med.* 53, 1212–1216.

Derrfuss, J., Brass, M., and von Cramon, D. Y. (2004). Cognitive control in the posterior frontolateral cortex: evidence from common activations in task coordination, interference control, and working memory. *Neuroimage* 23, 604–612.

Donohue, S. E., Wendelken, C., and Bunge, S. A. (2008). Neural correlates of preparation for action selection as a function of specific task demands. *J. Cogn. Neurosci.* 20, 1–13.

Falkenstein, M. (2006). Inhibition, conflict and the Nogo-N2. *Clin. Neurophysiol.* 117, 1638–1640.

Fan, J., Kolster, R., Ghajar, J., Sh, M., Knight, R. T., Sarkar, R., et al. (2007). Response anticipation and response conflict: an event-related potential and functional magnetic resonance imaging study. *J. Neurosci.* 27, 2272–2282.

Fassbender, C., Foxe, J. J., and Garavan, H. (2006). Maping the functional anatomy of task preparation: priming task appropriate brain networks. *Hum. Brain Mapp.* 27, 819–827.

Folstein, J. R., and van Petten, C. (2008). Influence of cognitive control and mismatch on the N2 component of the ERP: a review. *Psychophysiology* 45, 152–170.

Forstmann, B. U., van den Wildenberg, W. P. M., and Ridderinkhof, K. R. (2008a). Neural mechanisms, temporal dynamics, and individual differences in interference control. *J. Cogn. Neurosci.* 20 , 1854–1865.

Forstmann, B. U., Jahfari, S., Scholte, H. S., Wolfensteller, U., van den Wildenberg, W. P. M., and Ridderinkhof, K. R. (2008b). Function and structure of the right inferior frontal cortex predict individual differences in response inhibition: a model-based approach. *J. Neurosci.* 28, 9790–9796.

Genovese, C. R., Lazar, N. A., and Nichols, T. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *Neuroimage* 15, 870–878.

Gilbert, S. J., Spengler, S., Simons, J. S., Steele, J. D., Lawrie, S. M., and Frith, C. D, Burgess, P. W. (2006). Functional specialization within rostral prefrontal cortex (BA 10): a meta-analysis. *J. Cogn. Neurosci.* 18, 932–948.

Hazeltine, E., Poldrack, R., and Gabrieli, J. D. (2000). Neural activation during response competition. *J. Cogn. Neurosci.* 2, 118–129.

Heil, M., Osman, A., Wiegelmann, J., Rolke, B., and Hennighausen, E. (2000). N200 in the Eriksen-task: inhibitory executive processes? *J. Psychophysiol.* 14, 218–225.

Hikosaka, O., and Isoda, M. (2010). Switching from automatic to controlled behavior: cortico-basal ganglia mechanisms. *Trends Cogn. Sci. (Regul. Ed.)* 14, 154–161.

Isoda, M., and Hikosaka, O. (2007). Switching from automatic to controlled action by monkey medial frontal cortex. *Nat. Neurosci.* 10, 240–248.

Kehrer, S., Kraft, A., Irlbacher, K., Koch, S. P., Hagendorf, H., Kathmann, N., et al. (2009). Electrophysiological evidence for cognitive control during conflict processing in visual spatial attention. *Psychol. Res.* 73, 751–761.

Kennerley, S. W., Sakai, K., and Rushworth, M. F. S. (2004). Organization of action sequences and the role of the pre-SMA. *J. Neurophysiol.* 91, 978–993.

King, J. A., Korb, F. M., von Cramon, D. Y., and Ullsperger, M. (2010). Post-Error Behavioral adjustments are facilitated by activation and suppression of task-relevant and task-irrelevant information processing. *J. Neurosci.* 30, 12759–12769.

Koechlin, E., Basso, G., Pietrini, P., Panzer, S., and Grafmann, J. (1999). The role of the anterior prefrontal cortex in human cognition. *Nature* 399, 148–151.

Kool, W., McGuire, J. T., Rosen, Z. B., and Botvinick, M. M. (2010). Decision making and the avoidance of cognitive demand. *J. Exp. Psychol. Gen.* 139, 665–682.

Kopp, B., Mattler, U., Goertz, R., and Rist, F. (1996). N2, P3 and the lateralized readiness potential in a nogo task involving selective response priming. *Electroencephalogr. Clin. Neurophysiol.* 99, 19–27.

Lancaster, J. L., Woldorff, M. G., Parsons, L. M., Liotti, M., Freitas, C. S., Rainey, L., et al. (2000). Automated Talairach atlas labels for functional brain mapping. *Hum. Brain Mapp.* 10, 120–131.

Lau, H. C., Rogers, R. D., Haggard, P., and Passingham, R. E. (2004). Attention to intention. *Science* 303, 1208–1210.

Leuthold, H., Sommer, W., and Ulrich, R. (2004). Preparing for action: inferences from CNV and LRP. *J. Psychophysiol.* 18, 77–88.

Luks, T. L., Simpson, G. V., Dale, C. L., and Hough, M. G. (2007). Preparatory allocation of attention and adjustments in conflict processing. *Neuroimage* 35, 949–958.

MacDonald, A. W. III, Cohen, J. D., Stenger, V. A., and Carter, S. C. (2000). Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science* 288, 1835–1838.

Mars, R. B., Klein, M. C., Neubert, F.-X., Olivier, E., Buch, E. R., Boorman, E. D., et al. (2009). Short-latency influence of medial frontal cortex on primary motor cortex during action selection under conflict. *J. Neurosci.* 29, 6926–6931.

Milham, M. P., Banich, M. T., Webb, A., Barad, V., Cohen, N. J., Wszalek, T., et al. (2001). The relative involvement of anterior cingulate and prefrontal cortex in attentional control depends on nature of conflict. *Brain Res. Cogn. Brain Res.* 12, 467–473.

Nachev, P., Wydell, H., O'Neill, K., Husain, M., and Kennard, C. (2007). The role of the pre-supplementary motor area in the control of action. *Neuroimage* 36(Suppl. 2), T155–T163.

Nakamura, K., Sakai, K., and Hikosaka, O. (1998). Neuronal activity in medial frontal cortex during learning of sequential procedures. *J. Neurophysiol.* 80, 2671–2687.

Neuhaus, A. H., Urbanek, C., Opgen-Rhein, C., Hahn, E., Tam Ta, T. M., Koehler, S., et al. (2010). Event-related potentials associated with Attention Network Test. *Int. J. Psychophysiol.* 76, 72–79.

Nieuwenhuis, S., Yeung, N., van den Wildenberg, W., and Ridderinkhof, K. R. (2003). Electrophysiological correlates of anterior cingulate function in a go/no-go task: effects of response conflict and trial type frequency. *Cogn. Affect. Behav. Neurosci.* 3, 17–26.

Ramnani, N., and Owen, A. M. (2004). Anterior prefrontal cortex: insights into function from anatomy and neuroimaging. *Nat. Rev. Neurosci.* 5, 184–194.

Ridderinkhof, K. R., Forstmann, B., Wylie, S. A., Burle, B., and van den Wildenberg, W. P. M. (2010). Neurocognitive mechanisms of action control: resisting the call of the sirens. *Wiley Interdiscip. Rev. Cogn. Sci.* 2, 174–192.

Ridderinkhof, K. R., Ullsperger, M., Crone, E. A., and Nieuwenhuis, S. (2004). The role of the medial frontal cortex in cognitive control. *Science* 306, 443–447.

Ruge, H., Jamadar, S., Zimmermann, U., and Karayanidis, F. (2013). The many faces of preparatory control in task switching: reviewing a decade of fMRI research. *Hum. Brain Mapp.* 34, 12–35.

Rushworth, M. F. (2008). Intention, choice, and the medial frontal cortex. *Ann. N. Y. Acad. Sci.* 1124, 181–207.

Sakai, K., and Passingham, R. E. (2003). Prefrontal interactions reflect future task operations. *Nat. Neurosci.* 6, 75–81.

Sakai, K., and Passingham, R. E. (2006). Prefrontal set activity predicts rule-specific neural processing during subsequent cognitive performance. *J. Neurosci.* 26, 1211–1218.

Shima, K., Mushiake, H., Saito, N., and Tanji, J. (1996). Role for cells in the presupplementary motor area in updating motor plans. *Proc. Natl. Acad. Sci. U.S.A.* 93, 8694–8698.

Simon, J. R. (1967). Choice reaction time as a function of auditory S-R correspondence, age and sex. *Ergonomics* 10, 559–564.

Smith, E. E., and Jonides, J. (1999). Storage and executive processes in the frontal lobes. *Science* 283, 1657–1661.

Sohn, M.-H., Albert, M. V., Jung, K., Carter, C. S., and Andersen, J. R. (2007). Anticipation of conflict monitoring in the anterior cingulate cortex and the prefrontal cortex. *Proc. Natl. Acad. Sci. U.S.A.* 104, 10330–10334.

Sumner, P., Nachev, P., Morris, P., Peters, A. M., Jackson, S. R., Kennard, C., et al. (2007). Human medial frontal cortex mediates unconscious inhibition of voluntary action. *Neuron* 54, 697–711.

Taylor, S. F., Stern, E. R., and Gehring, W. J. (2007). Neural systems for error monitoring: recent findings and theoretical perspectives. *Neuroscientist* 13, 160–172.

Ullsperger, M., and King, J. A. (2010). Proactive and reactive recruitment of cognitive control: comment on Hikosaka and Isoda. *Trends Cogn. Sci. (Regul. Ed.)* 14, 191–192.

Ullsperger, M., and von Cramon, D. Y. (2001). Subprocesses of performance monitoring: a dissociation of error processing and response competition revealed by event-related fMRI and ERPs. *Neuroimage* 14, 1387–1401.

Vallesi, A., Mapelli, D., Schiff, S., Amodio, P., and Umilta, C. (2005). Horizontal and vertical Simon effect: different underlying mechanisms? *Cognition* 96, B33–B43.

Wagner, A. D., Bunge, S. A., and Badre, D. (2004). "Cognitive control, semantic memory, and priming: contributions from prefrontal cortex," in *The Cognitive Neurosciences*, 3rd Edn, ed. M. S. Gazzaniga (Cambridge, MA: MIT Press), 709–725.

Walter, W. G. (1964). Slow potential waves in the human brain associated with expectancy, attention and decision. *Eur. Arch. Psychiatry Clin. Neurosci.* 206, 309–322.

Wolfensteller, U., and von Cramon, D. Y. (2010). Bending the rules: strategic behavioral differences are reflected in the brain. *J. Cogn. Neurosci.* 22, 278–291.

Wolfensteller, U., and von Cramon, D. Y. (2011). Strategy-effects in prefrontal cortex during learning of higher-order S-R rules. *Neuroimage* 57, 598–607.

Wühr, P., and Kunde, W. (2008). Precueing spatial S-R correspondence: is there regulation of expected response conflict? *J. Exp. Psychol. Hum. Percept. Perform.* 34, 872–883.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 16 April 2012; accepted: 22 January 2013; published online: 12 February 2013.

Citation: Strack G, Kaufmann C, Kehrer S, Brandt S and Stürmer B (2013) Anticipatory regulation of action control in a Simon task: behavioral, electrophysiological, and fMRI correlates. Front. Psychology ***4****:47. doi: 10.3389/fpsyg.2013.00047*

This article was submitted to Frontiers in Cognition, a specialty of Frontiers in Psychology.

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 24 September 2012
doi: 10.3389/fpsyg.2012.00358

Model-based analysis of context-specific cognitive control

Joseph A. King[1], Christopher Donkin[2], Franziska M. Korb[1] and Tobias Egner[1]*

[1] Center for Cognitive Neuroscience, Duke University, Durham, NC, USA
[2] School of Psychology, University of New South Wales, Kensington, NSW, Australia

Edited by:
Rico Fischer, Technische Universität Dresden, Germany

Reviewed by:
Bruce Milliken, McMaster University, Canada
Tom Verguts, Ghent University, Belgium
Gilles Dutilh, University of Basel, Switzerland

***Correspondence:**
Tobias Egner, Center for Cognitive Neuroscience, Duke University, Box 90999, Durham, NC 27708, USA.
e-mail: tobias.egner@duke.edu

Interference resolution is improved for stimuli presented in contexts (e.g., locations) associated with frequent conflict. This phenomenon, the context-specific proportion congruent (CSPC) effect, has challenged the traditional juxtaposition of "automatic" and "controlled" processing because it suggests that contextual cues can prime top-down control settings in a bottom-up manner. We recently obtained support for this "priming of control" hypothesis with functional magnetic resonance imaging by showing that CSPC effects are mediated by contextually cued adjustments in processing selectivity. However, an equally plausible explanation is that CSPC effects reflect adjustments in response caution triggered by expectancy violations (i.e., prediction errors) when encountering rare events as compared to common ones (e.g., incongruent trials in a task context associated with infrequent conflict). Here, we applied a quantitative model of choice, the linear ballistic accumulator (LBA), to distil the reaction time and accuracy data from four independent samples that performed a modified flanker task into latent variables representing the psychological processes underlying task-related decision making. We contrasted models which differentially accounted for CSPC effects as arising either from contextually cued shifts in the *rate* of sensory evidence accumulation ("drift" models) or in the *amount* of evidence required to reach a decision ("threshold" models). For the majority of the participants, the LBA ascribed CSPC effects to increases in response threshold for contextually infrequent trial types (e.g., congruent trials in the frequent conflict context), suggesting that the phenomenon may reflect more a prediction error-triggered shift in decision criterion rather than enhanced sensory evidence accumulation under conditions of frequent conflict.

Keywords: cognitive control, conflict, evidence accumulation models, interference, mathematical modeling, priming, prediction error, response threshold

INTRODUCTION

The ability to focus attention on information relevant to the task at hand while simultaneously ignoring myriad potential sources of distraction in the environment is critical for purposeful, goal-directed behavior. The efficiency at which the brain supports this ability to filter relevant stimuli from irrelevant noise can be gauged by "interference" effects in performance of classic selective attention/response conflict paradigms such as the Stroop color-word naming task (Stroop, 1935; MacLeod, 1991) or the Eriksen flanker task (Eriksen and Eriksen, 1974). In the flanker task, for instance, interference effects are expressed as reliably slower reaction times (RT) and decreased accuracy on trials in which a central target stimulus is flanked by incongruent distracters (e.g., HHSHH or (< > >) relative to trials in which the target is flanked by congruent ones (e.g., HHHHH or < < < < <). Interference (or "conflict"; defined as concurrent activation of mutually incompatible stimulus or response representations) is commonly thought to arise from involuntary, "automatic" processing of irrelevant information based on well-learned stimulus-response associations that are triggered in bottom-up fashion. Accordingly, the ability to resolve interference/conflict is thought to be dependent on effortful, "controlled" processing that employs internal goal representations to intentionally overcome habitual associations in a top-down manner (Cohen et al., 1990; Botvinick et al., 2001).

Recent research using selective attention/response conflict tasks has challenged the traditional distinction between automatic and controlled processing, however, implying that this juxtaposition may in fact represent a false dichotomy. Specifically, several studies have suggested a melding of bottom-up associative processing and top-down attentional control settings by showing that when stimuli are presented in contexts (e.g., locations, colors, or sensory modalities) paired with frequent conflict, interference resolution is significantly improved (i.e., congruency effects are reduced; Corballis and Gratton, 2003; Crump et al., 2006, 2008; Lehle and Hübner, 2008; Wendt et al., 2008; Vietze and Wendt, 2009; Wendt and Kiesel, 2011; D'Angelo and Milliken, 2012; for review, see Bugg and Crump, 2012). Interestingly, these so-called context-specific proportion congruent (CSPC) effects occur even though observers are unaware of any systematic contextual variation in conflict frequency (Crump et al., 2008; Heinemann et al., 2009; Sarmiento et al., 2012). For example, using a modified Stroop task, Crump et al. (2006) showed that interference effects were reduced for stimuli presented in contexts (e.g., above central fixation) in which 75% of trials were incongruent (i.e., low proportion congruent/frequent conflict context) relative to those for stimuli presented in contexts (e.g., below fixation) in which 75% of trials were congruent (i.e., high proportion congruent/infrequent conflict context). The context-specificity and implicit nature of CSPC effects suggests that they are driven by bottom-up stimulus features. A purely associative explanation can be ruled out, however, because the context-specific improvement in interference resolution generalizes to frequency-unbiased stimuli (Crump and Milliken, 2009; Heinemann et al., 2009). Building on these previous findings, King et al. (2012) obtained neural evidence of bottom-up contextual priming of top-down control in a functional magnetic resonance imaging (fMRI) experiment. In particular, we found that the behavioral expression of CSPC effects in a flanker task variant using trial-unique stimuli (**Figure 1A**) was mirrored in contextual variation of hemodynamic activity associated with conflict processing in a region of the medial superior parietal lobule (mSPL) broadly implicated in top-down attentional selection (Yantis, 2008; Chiu and Yantis, 2009; Esterman et al., 2009; Greenberg et al., 2010; Shomstein, 2012) and that this activity explained modulation of stimulus-driven processing in task-relevant regions of sensory cortex.

Extant data pertaining to CSPC effects support the hypothesis that they reflect contextually cued adjustments in perceptual processing selectivity (e.g., Crump et al., 2006; Lehle and Hübner, 2008; Wendt et al., 2008; Crump and Milliken, 2009). That is, presentation of a stimulus in a context associated with frequent conflict appears to promote more efficient segregation of relevant from irrelevant stimulus information, facilitating faster responses to incongruent stimuli (but slower responses to congruent ones) relative to a context of infrequent conflict. However, an equally plausible alternative explanation is that the phenomenon is attributable to adjustments in response caution triggered by the relative frequency of events within each stimulus context. Specifically, a rare, contextually unlikely stimulus may induce a shift toward a more conservative response criterion, granting the observer more time for reaching a reliable perceptual decision. Thus, the characteristic pattern of CSPC effects (**Figure 1B**) could either reflect enhanced processing selectivity for stimuli presented in the frequent conflict context as suggested by several behavioral studies (e.g., Crump et al., 2006; Lehle and Hübner, 2008; Wendt et al., 2008; Crump and Milliken, 2009) and corroborated by our neuroimaging findings (King et al., 2012), or instead indicate a relative increase in response threshold when encountering unexpected, rare events (e.g., incongruent trials in the infrequent conflict context) as compared to expected or common ones (e.g., incongruent trials in the frequent conflict context). Neither conventional analyses of mean RT and error rates, nor our fMRI analyses could clearly disambiguate between these two possibilities. The purpose of the current study was to use a formal quantitative model of decision making to adjudicate between competing accounts of CSPC effects which differentially attribute the phenomenon to (1) contextually cued enhancement in processing selectivity or (2) shifts in response caution triggered by violations of expectancy regarding stimulus congruency (i.e., prediction error) within each context.

Quantitative sequential sampling models of decision making are increasingly being used to decompose the cognitive processes and neural mechanisms underlying choice RTs (for review, see Forstmann et al., 2011; Mars et al., 2012), such as those made in selective attention/response conflict paradigms (e.g., White et al., 2011). Several "evidence accumulation" models of choice have been developed (e.g., Smith and Vickers, 1988; Ratcliff and Rouder, 1998; van Zandt, 2000; Usher and McClelland, 2001; Brown and Heathcote, 2008), all of which vary in their assumptions regarding the precise nature of the constituent cognitive processes involved in rapid decision making and computational efficiency. Nonetheless, these models share the same basic notion that when participants make a decision about a stimulus, they continuously sample information from the environment and that this information serves as evidence for one of the possible responses. When evidence in favor of a potential response reaches a threshold, a decision is made and the associated response is given. In predicting performance, evidence accumulation models take into account the interaction between response speed and accuracy to estimate four central parameters: (1) an *a priori* bias for one or the other decision ("starting-point"), (2) the rate of evidence accumulation in favor of a particular decision ("drift rate"), (3) the amount of evidence that is necessary for triggering a particular decision ("response threshold"), and (4) the time involved in stimulus encoding and response execution ("non-decision time"). Here, we applied an established model of decision making, the linear ballistic accumulator (LBA) model (Brown and Heathcote, 2008; Donkin et al., 2009, 2011b; for examples of recent applications with comparable trial-per-condition numbers as in the current experiments, see e.g., Forstmann et al., 2008, 2010; Ho et al., 2009; Ludwig et al., 2009; van Maanen et al., 2011; McVay and Kane, 2012) to behavioral data collected from four independent samples during performance of the flanker task depicted in **Figure 1A**. Our objective was to test whether CSPC effects can be better accounted for as contextually cued shifts in the rate of evidence accumulation about the target stimulus (i.e., drift rate) or in the amount of evidence required to reach a decision (i.e., response threshold). **Figure 1C** illustrates how decisions regarding targets in this task are represented in the LBA. We predicted that if CSPC effects reflect contextually cued adjustments in processing selectivity, a model in which the rate of evidence accumulation (drift rate parameters) was allowed to vary across context and congruency conditions would provide the most parsimonious account of the empirical data ("drift" model). Alternatively, we expected that if CSPC effects reflect shifts in response caution triggered by unexpected, contextually unlikely stimuli, a model in which response threshold was allowed to vary across conditions would provide the best fit to the observed performance ("threshold" model).

MATERIALS AND METHODS

PARTICIPANTS

The data reported here were collected from a total of 87 participants belonging to four independent samples that performed the identical flanker task (**Figure 1A**) in (1) the fMRI experiment described in King et al. (2012; $n = 25$; data set I), (2) a behavioral pilot study designed to test the adequacy of the paradigm for the magnetic resonance scanner environment ($n = 19$; data set II) and two follow-up behavioral studies designed to explore whether CSPC effects in this task, (3) are mediated by spatial stimulus-response compatibility effects ($n = 25$; data set III), or (4) vary as a function of awareness regarding the contextual

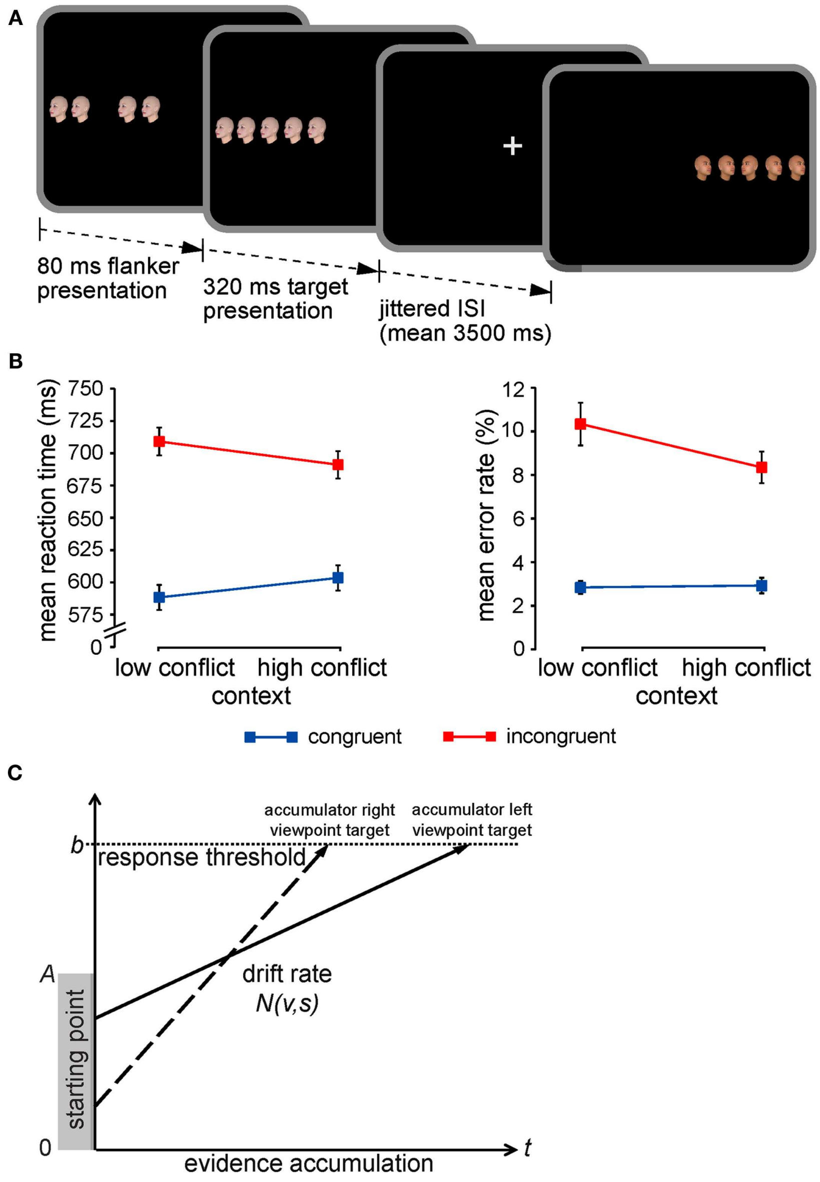

FIGURE 1 | Experimental paradigm, CSPC effects, and the LBA model. **(A)** The face-viewpoint flanker task used to collect all four data sets was identical. Each trial began with the presentation of four novel (trial-unique) flanker faces, followed by an identical target face in the center of the array. Participants had to classify the viewpoint direction of the target face with a button press. Target and flanker face-viewpoint direction was congruent in half of all trials (shown here in the first trial) and incongruent in the other half (shown here in the second trial). The proportion of congruent to incongruent stimuli (conflict frequency) was manipulated in a context-specific manner according to stimulus location: one side of fixation was associated with 75% congruent trials (low-conflict context) and the other side with 75% incongruent trials (high-conflict context). For further details, see Section "Materials and Methods." **(B)** Mean RTs and error rates (±SEM) are plotted for flanker congruent and incongruent trials as a function of the contextual conflict-frequency manipulation, illustrating the critical context × congruency interactions (i.e., CSPC effects). **(C)** The LBA model as applied to a typical decision in the face-viewpoint flanker task. One accumulator corresponds to the response that the target face is pointing left (solid arrow), while the other accumulator corresponds to a rightward response (dashed arrow). A response is triggered as soon as an accumulator reaches the response threshold, *b* (horizontal dotted line). Each accumulator begins with a starting amount of evidence drawn randomly from the range indicated by the gray-shaded rectangle (between 0 and *A*), and the accumulation rate (i.e., drift) for each response is drawn from a normal distribution with an appropriate mean, *v*, and SD, *s*.

conflict-frequency manipulation ($n = 18$; data set IV), respectively. All studies were conducted according to protocols approved by the Duke University Health System Institutional Review Board. For a detailed description of the sample contributing to data set I, see King et al. (2012). For data set II, a total of 21 volunteers with normal or corrected-to-normal vision participated. The data of two participants were excluded from further analysis due to chance level performance. The final sample (10 females, 9 males; mean age = 27.3 years; range = 22–37 years) included nine members of the Duke University Center for Cognitive Neuroscience (two research assistants, three doctoral students, three post-doctoral researchers, and one assistant professor) and 10 individuals recruited from the greater Durham, NC community by an advertisement on the Duke University Center for Cognitive Neuroscience Research Participation website who received $10 their participation. For data set III, a total of 26 undergraduates participated for class credit. The data of one participant was excluded from further analysis due to chance level performance. The final sample consisted of 17 females and 8 males (mean age = 20 years; range = 18–24 years). For data set IV, a total of 20 undergraduates participated for class credit. The data of two participants were excluded for chance level performance. The final sample consisted of 11 females and 7 males (mean age = 19.7 years; range = 18–23 years).

APPARATUS, STIMULI, AND PROCEDURE

Task programming, stimulus presentation, and behavioral recording were carried out with Presentation software (Neurobehavioral Systems; Albany, CA, USA). Face stimuli for the flanker experiment were generated with FaceGen software (Singular Inversions; Toronto, ON, Canada) to produce an equal number of left- and right-looking male and female faces (137 each; viewpoint angle: ~45–50°) with unique identities from various age- and ethnic-groups. A total of 448 face images were used, one for each face trial of the experiment. Further details regarding stimulus generation are provided in King et al. (2012). For the fMRI experiment (data set I), stimuli were presented against a black background on a back projection screen, which participants viewed in a mirror mounted to the head coil; simulating a viewing distance of ~80 cm. Given these viewing conditions, individual face stimuli within flanker arrays extended ~0.72° horizontally and 1.1° vertically and were presented at ~2.9, 3.8, 4.7, 5.6, and 6.5° horizontal visual angle to the left and right of central fixation. For the behavioral experiments (data sets II–IV), participants sat in a dimly lit room and viewed stimuli displayed against a black background on a 19″ LCD monitor at a distance of ~80 cm, approximating the viewing conditions in the scanner.

In each trial of the flanker task (**Figure 1A**), a novel stimulus array (row of five identical trial-unique face images) was presented pseudorandomly either to the left or right of fixation. Participants were instructed to rapidly and accurately classify with a button press the viewpoint direction of the face in the center of the array (target) and ignore the flanker faces (distracters). For data sets I, II, and IV, responses were given with a right-hand index or middle finger button press. For data set III, responses were given with the index fingers of both hands. Stimulus-response mapping was counterbalanced across participants for all experiments. The target face was presented for 320 ms; its onset was delayed by 80 ms from the onset of the flanker faces, which were shown for 400 ms. Target and flanker face-viewpoint direction was congruent in half of all trials and incongruent in the other. Proportion congruency (i.e., conflict frequency) was manipulated according to stimulus location by defining one side of fixation as a high-conflict context (i.e., low proportion congruent; 25% congruent/75% incongruent trials) and the other as a low-conflict context (i.e., high proportion congruent; 75% congruent/25% incongruent trials; counterbalanced across participants). Inter-stimulus intervals were jittered between 3 and 5 s as randomly drawn from a pseudoexponential distribution, where 50% of intervals lasted 3 s, 25% lasted 3.5 s, 12% lasted 4 s, 6% lasted 4.5 s, and 6% lasted 5 s, resulting in a mean interval of ~3.5 s. To counteract potential spatial stimulus-response compatibility effects in the fMRI experiment (data set I), participants responded on a response box (Current Designs, Philadelphia, PA, USA) that was vertically oriented on the participant's chest (i.e., in plane with the length of their body). For the same reason, responses were given on the ↑ (8) and ↓ (2) buttons of the numeric keypad on a QWERTY US keyboard in the behavioral pilot experiment (data set II) and the experiment designed to test the influence of awareness of contextual variation in conflict frequency on CSPC effects (data set IV). Given that the purpose of data set III was to test whether CSPC effects might be mediated by potential stimulus-response compatibility effects, we asked participants to respond in a lateralized manner using the z and m keys on a QWERTY US keyboard. We explored the influence of awareness of the contextual conflict-frequency manipulation on CSPC effects in data set IV by informing the participants which side of fixation was associated with mostly congruent stimuli (low-conflict context) and mostly incongruent stimuli (high-conflict context) and encouraging them to use this information to their advantage. This manipulation was successful in that all 18 subjects that contributed to this data set reported that they noticed the location-based variation in congruency frequency in a post-test questionnaire, while only one out of 25 participants that contributed to data set I (King et al., 2012) reported explicit knowledge of the contextual conflict-frequency manipulation. Trials occurred in four blocks in the fMRI experiment (data set I; 112 trials each) and in seven blocks in the behavioral experiments (data sets II, III, and IV; 64 trials each). Participation in fMRI experiment lasted ~75 min including a 64-trial training session, anatomical scanning, performance of an independent localizer task, and completion of a post-test survey (see King et al., 2012, for further details). Participation in the behavioral experiments lasted ~40 min, including a 64-trial training session.

CONVENTIONAL ANALYSIS OF RESPONSE LATENCY AND ACCURACY

Prior to exploring the performance data with conventional analyses, we excluded the first trial of each block and all trials with excessively fast or slow responses (<150 ms/>2000 ms; 1.2% of all trials). We tested for contextual variation in interference effects [i.e., CSPC effects; $(\text{incongruent-congruent})_{\text{low-conflict context}} - (\text{incongruent-congruent})_{\text{high-conflict context}}$] and their possible modulation as a function of spatial stimulus-response compatibility and/or awareness of the contextual conflict-frequency manipulation by submitting mean correct trial RTs (excluding post-error correct trials) and error rates to 2 (context: high conflict vs. low-conflict) × 2 (spatial stimulus-response compatibility: compatible vs. incompatible) × 2 (flanker congruency: congruent vs. incongruent) repeated-measures ANOVAs, using experimental session (data sets I–IV) as a between-subjects factor. Our previous fMRI study (data set I) revealed that CSPC effects varied as a function of context transitions. Specifically, they were only present for context repetitions, but absent for switches between contexts (e.g., from the low- to the high-conflict context; King et al., 2012). A supplementary 2 (context transition: repetition vs. switch) × 2 (context) × 2 (congruency) ANOVA using experimental session as a between-subjects factor explored whether this pattern was stable across data sets.

MODEL FITTING

The primary objective of the current study was to explore whether the LBA model attributes CSPC effects to contextually cued adjustments in processing selectivity (as indexed by the *rate* of evidence accumulation, i.e., drift) or to shifts in response caution triggered by unexpected stimuli within each context (as indexed by the *amount* of evidence required to make a decision, i.e., response threshold). These hypotheses were tested by fitting the performance data from each individual participant from each of the four data sets with models whose parameterizations reflected these differing assumptions about the influence of implicit contextual information on conflict processing. Support for each of the hypotheses comes from how well the respective models can fit the data. Readers unfamiliar with the methods involved in fitting sequential sampling models to choice RT data or the techniques involved in model selection (see the following section) are referred a tutorial paper which focuses specifically on the LBA, but is generally applicable to other evidence accumulation models (Donkin et al., 2011a).

We report the results of four models of CSPC effects in detail. The first two models (Models V1 and V2) assumed that CSPC effects arise from the influence of context on evidence accumulation rate (v; see **Figure 1C**). Both of these "drift" models accounted for CSPC effects by predicting the difference in v for congruent and incongruent stimuli to be larger in the low-conflict context than in the high-conflict context, but they did so in different ways. In Model V1, there was no constraint placed on v, and a separate parameter was estimated for each of the four experimental conditions (i.e., $v_{\text{Incon-Low}}$, $v_{\text{Con-Low}}$, $v_{\text{Incon-High}}$, and $v_{\text{Con-High}}$). In contrast, Model V2 assumed that the increase in v (i.e., faster rate) as we move from low- to high-conflict contexts for incongruent stimuli (recall that people get better at responding to incongruent stimuli in high-conflict contexts) is of the same magnitude as the decrease in v (i.e., slower rate) from low- to high-conflict contexts for congruent stimuli (people are worse for congruent stimuli in high-conflict contexts). As such, three rate parameters were estimated: $v_{\text{Incon-Low}}$, $v_{\text{Con-Low}}$, and Δv, while accumulation rates in the high-conflict context were $v_{\text{Incon-Low}} + \Delta v$ for incongruent trials and $v_{\text{Cong-Low}} - \Delta v$ for congruent trials. In other words, Model V2 assumed the absolute difference in v resulting from a shift between contexts to be equal for congruent and incongruent trials. To illustrate, whereas a shift from the low- to the high-conflict context should lower v for congruent trials, it should increase v for incongruent trials to the same degree. In both drift models, response threshold was held constant across the high- and low-conflict contexts.

The latter two models (Models B1 and B2) assumed that CSPC effects arise from the influence of context on response threshold (b; see **Figure 1C**). Both of these "threshold" models accounted for CSPC effects by predicting that there would be differences in the distance from the top of the start-point distribution to response threshold, $b - A$. In particular, it was assumed that the difference between thresholds in the congruent and incongruent stimuli would be larger in the low-conflict than in the high-conflict context. However, as in the drift models outlined above, the threshold models also accounted for CSPC in different ways. In Model B1, as in Model V1, no constraint was placed on the way that response thresholds would change according to the context and congruency conditions, and so four threshold parameters were estimated ($b_{\text{Incon-Low}} - A$, $b_{\text{Con-Low}} - A$, $b_{\text{Incon-High}} - A$, and $b_{\text{Con-High}} - A$). In contrast, Model B2 was constrained in manner equivalent to Model V2 such that the absolute difference in response threshold for congruent and incongruent stimuli was equal between the low- and high-conflict contexts. That is, the reduction in thresholds as we move from low- to high-conflict contexts for incongruent stimuli is of the same magnitude as the increase in thresholds from low- to high-conflict contexts for congruent stimuli. In particular, three threshold parameters were estimated: $b_{\text{Incon-Low}} - A$, $b_{\text{Con-Low}} - A$, and Δb, while thresholds in the high conflict were $b_{\text{Incon-Low}} - A + \Delta b$ for incongruent trials and $b_{\text{Incon-Low}} - A - \Delta b$ for congruent trials. In both threshold models, evidence accumulation rate was allowed to vary as a function of stimulus congruency, but not across the two contexts.

In all models, the SD of the distribution of drift rate across trials, s, the maximum of the uniform between-trial distribution of start-point, A, and non-decision time, t_0, were fixed across the congruency and context conditions. Though no restrictions were made about the sign of Δb and Δv parameters, Models V2 and B2 were parameterized such that positive values of Δb and Δv would produce the standard CSPC effects.

Models were fit to each of the individual participants from each of the four data sets. The likelihood of the response time and response choice on each trial (the number of valid trials per participant after excluding response omissions ranged from 408 to 448; mean = 445 trials; SD = 5.8 trials) was calculated using the formulas outlined in Brown and Heathcote (2008). Particle swarm optimization was used to find best-fitting parameters by searching for the maximum of the sum of the likelihoods across all trials for each individual.

In addition to the models outlined above, we fit a number of other model parameterizations following standard practice (Donkin et al., 2011a) that we do not report here. For example, we fit one model in which both evidence accumulation rate and response thresholds were allowed to vary concurrently and another in which the CSPC effect was assumed to reflect a shift in non-decision time, t_0. None of these models outperformed any of the models we report in detail, with probabilities generally not greater than about 5%. As such, we refrain from further discussion of these models.

MODEL SELECTION

The four models were compared using the commonly employed Akaike and Bayesian Information Criterion (AIC, Akaike, 1974; BIC, Schwarz, 1978, respectively). BIC was calculated using the standard formula

$$\text{BIC} = k \ln N - 2 \ln L,$$

where L is the likelihood of the parameters given the data, N is the number of data points used to calculate the likelihood value, and

k is the number of free parameters used to fit the data. Similarly, AIC was calculated using

$$\text{AIC} = 2k - 2 \ \ln \ L.$$

Note that for our data, AIC has a smaller complexity term whenever $\ln N > 8$.

To aid interpretability, AIC and BIC values were converted into AIC and BIC weights using the method outlined in Wagenmakers and Farrell (2004). In short, the information criterion (IC) values are transformed in ΔIC values by subtracting the smallest IC value from the IC for each model. ΔIC are then turned into weights using the following

$$w_i\,(\text{IC}) = \frac{e^{-\frac{1}{2}\Delta_i(\text{IC})}}{\sum_k e^{-\frac{1}{2}\Delta_k(\text{IC})}}$$

where w_i is the weight for the ith model. AIC and BIC weights reflect the probability that a particular model is true.

RESULTS

CONVENTIONAL ANALYSES OF RESPONSE LATENCY AND ACCURACY

For the combined sample ($n = 87$), overall performance was high (93.5% correct) and characterized by typical flanker interference effects. RTs were slower for incongruent stimuli (700 ms) than for congruent arrays [596 ms; $F(1,83) = 668.8$; $p < 0.0001$]. Similarly, error rates were elevated on incongruent (9.2%) relative to congruent trials [3.0%; $F(1,83) = 80.0$; $p < 0.0001$]. Interestingly, a reversed spatial stimulus-response compatibility effect emerged in RTs. Responses were generally slower when the viewpoint direction of target faces (e.g., left) corresponded (i.e., were compatible) with the location of stimulus array (e.g., left of fixation; 655 ms) relative to when the viewpoint direction of targets did not correspond (i.e., were incompatible) with the stimulus position (641 ms). The magnitude of this effect varied across experimental sessions [$F(3,83) = 2.9$; $p < 0.05$] such that it was most pronounced in data set IV (22 ms), but virtually absent in data set I (3 ms). In any event, spatial stimulus-response compatibility effects did not interact with flanker congruency, stimulus context, or their combination [all $F(3,83) < 3.3$; n.s.] and therefore have no implications for the interpretation of the CSPC effects at the focus of interest in this study. The contextual manipulation of flanker conflict frequency did not have any general effect on RTs [$F(1,83) = 0.3$; n.s.], but error rates were elevated in the low- (6.5%) vs. high-conflict context [5.6%; $F(1,83) = 9.1$; $p < 0.005$]. A main effect of experimental session was present in RTs [$F(3,83) = 11.3$; $p < 0.0001$], with responses being slower in the fMRI session (729 ms) than those in the three other experiments combined (620 ms).

More importantly, CSPC effects were clearly evident both in RTs [$F(1,83) = 53.9$; $p < 0.0001$] and error rates [$F(1,83) = 11.4$; $p < 0.001$] and were of comparable magnitude across experimental sessions [both $F(3,83) < 1.8$; n.s.]. Flanker interference effects were reduced for stimuli presented in the high-conflict location (RTs: 88 ms; error rates: 5.1%) relative to those in the low-conflict context (RTs: 121 ms; error rates: 7.2%; **Figure 1B**). Indicating that neither the lateralized response procedure introduced in data set III, nor informing participants about the contextual conflict-frequency manipulation in data set IV had any effect on CSPC effects, context × flanker congruency effects did not interact with the stimulus-response compatibility factor, experimental session, or their combination either in the RT or error rate data [all $F(3,83) < 1.7$; n.s.]. Replicating the finding that CSPC effects vary as a function of context transitions (King et al., 2012), they clearly occurred in context repetitions (47 ms), but were absent in context switches [19 ms; $F(3,83) = 17.5$; $p < 0.0001$]. Indicating the reliability of this effect, it did not vary across data sets [$F(3,83) = 1.2$; n.s.], even after excluding the data of our previous study [data set I; $F(2,59) = 0.01$; n.s.].

Together, the results of the conventional analyses of RTs and accuracy rates illustrate the robustness of CSPC effects on the one hand and an important boundary condition on the other, namely, that they appear to occur only in context repetitions. Additionally, they show that they are not confounded by spatial stimulus-response compatibility effects in the current paradigm and occur independently of participants' awareness of the contextual conflict-frequency manipulation. However, these data do not speak to our motivating question whether CSPC effects reflect contextually cued adjustments in processing selectivity or shifts in response caution triggered by the infrequent events within each stimulus context. To address this issue, we turn now to the modeling data.

MODELING DATA

The average parameter values for each of the four data sets for each of the four LBA models are shown in **Table 1**. Looking at the parameter values, it is apparent that, in general, the drift Models V1 and V2 accounted for CSPC effects by assuming that evidence accumulation rates were larger (i.e., faster) in the high-conflict context than in the low-conflict context on incongruent trials and smaller (i.e., slower) in the high-conflict context than in low-conflict contexts on congruent trials. The threshold Models B1 and B2, on the other hand, accounted for CSPC effects through the equivalent setting of response thresholds: larger thresholds in the high-conflict context than in the low-conflict context on congruent trials, and vice versa for incongruent trials.

OBSERVED AND PREDICTED RT DISTRIBUTIONS

The quality of agreement between the models and the data from each of the four data sets are plotted in **Figures 2A–D** (one figure per data set). The figure shows RT distributions for correct and erroneous responses on congruent and incongruent trials in the high- and low-conflict contexts (columns), along with model predictions from the four models (rows), as cumulative distribution function plots. Each plot is made up of quantile estimates from correct and incorrect RT distributions. The quantile estimates show the RT below which 10, 30, 50, 70, and 90% of the responses in that distribution fall. The positions of the quantiles on the x-axis reflect the speed at which responses are made, so that slower distributions stretch further to the right. The heights of the quantiles indicate, separately for correct and incorrect trials, the absolute cumulative proportion of responses with RTs below the quantile cutoff.

The plots in **Figure 2** demonstrate that the predictions from all four models (circles) match the observed data (diamonds) well.

Table 1 | Parameter values for Models V1 and V2 ("drift" models) and Models B1 and B2 ("threshold" models) averaged across participants in data sets I, II, III, and IV.

	Data set	s	A	$b_{\text{C-L}}$	$b_{\text{I-L}}$	$b_{\text{C-H}}$	$b_{\text{I-H}}$	Δb	t_0	$v_{\text{C-L}}$	$v_{\text{I-L}}$	$v_{\text{C-H}}$	$v_{\text{I-H}}$	Δv
Model V1	I	0.18	0.14		0.48			–	0.14	0.79	0.64	0.77	0.66	–
	II	0.15	0.09		0.36			–	0.13	0.78	0.63	0.75	0.65	–
	III	0.13	0.09		0.37			–	0.08	0.73	0.58	0.71	0.60	–
	IV	0.17	0.15		0.46			–	0.08	0.77	0.62	0.75	0.64	–
Model V2	I	0.18	0.14		0.48			–	0.14	0.79	0.64	–	–	0.02
	II	0.15	0.09		0.36			–	0.13	0.78	0.63	–	–	0.02
	III	0.13	0.10		0.37			–	0.08	0.73	0.59	–	–	0.02
	IV	0.17	0.15		0.46			–	0.08	0.77	0.62	–	–	0.02
Model B1	I	0.19	0.16	0.43	0.48	0.41	0.47	–	0.19	0.77	0.70	–	–	–
	II	0.17	0.10	0.32	0.36	0.33	0.35	–	0.17	0.76	0.70	–	–	–
	III	0.14	0.10	0.36	0.37	0.37	0.37	–	0.11	0.73	0.61	–	–	–
	IV	0.19	0.18	0.42	0.47	0.43	0.47	–	0.13	0.77	0.69	–	–	–
Model B2	I	0.19	0.17	0.43	0.48	–	–	0.01	0.20	0.78	0.70	–	–	–
	II	0.17	0.10	0.32	0.37	–	–	0.01	0.17	0.76	0.69	–	–	–
	III	0.14	0.11	0.36	0.37	–	–	0.01	0.12	0.75	0.62	–	–	–
	IV	0.18	0.20	0.43	0.47	–	–	0.01	0.14	0.79	0.70	–	–	–

C, Con, I, Incon; H, High, L, Low; s, standard deviation; A, upper limit of the start-point distribution; b, response threshold; t_0*, non-decision time; v, drift rate. In Model B2,* $b_{\text{Incon-High}} = b_{\text{Incon-Low}} + \Delta b$ *and* $b_{\text{Con-High}} = b_{\text{Con-Low}} - \Delta b$*. In Model V2,* $v_{\text{Incon-High}} = v_{\text{Incon-Low}} - \Delta v$ *and* $v_{\text{Con-High}} = v_{\text{Con-Low}} + \Delta v$*.*

The LBA model appears to give a good account of the full RT distributions for correct responses. All models appear to struggle somewhat to account for the speed of incorrect responses, especially in data sets III and IV (the lower function in each row of the second and fourth columns in **Figures 2C,D**). Differences between the models in their ability to fit the data are small, but perhaps most pronounced in their account of correct responses in low-conflict incongruent trials (the second column in each panel of **Figure 2**), particularly for data sets I and II. Models V1 and V2 (drift models) appear to systematically predict faster correct responses than were observed, while Models B1 and B2 (threshold models) also show this misfit, particularly for data sets III and IV, but to a lesser degree.

To help distinguish between the models, we turn to their predictions for mean RT. **Figure 3** contains the predictions for mean RT for Models V2 and B2 (the predictions of Models V1 and B1 are very similar, and the overall pattern of misfits the same). Model predictions (open circles) are close to the observed data (filled squares) for both models. Model B2 does appear to outperform Model V2 for all but Data Set IV, for which both models appear to provide an equivalent account.

MODEL SELECTION

Table 2 contains AIC and BIC weights for each of the models for each of the four data sets. Additionally, the table presents the number of participants best fit by each model (in parentheses). The AIC weights suggest that Model B1 is most often the true model across participants and data sets (42.5% of the time), followed by Model B2 (38% of the time; threshold models), then Model V2 and finally V1 (the drift models "won" only roughly 20% of the time). The results are different using BIC, where we see that Model B2 is preferred more often than Model B1 (roughly 47 vs. 14% of the time). Notice also, however, that for data sets III and IV, the difference in model probabilities for Model B2 and V2 is less clear. The differences in conclusions drawn from BIC and AIC reflect the fact that BIC has a larger penalty for complexity, and that Model B2 has one fewer free parameter than Model B1, and because the response threshold models use one more free parameter than their respective drift models.

The AIC and BIC weights can be used to compare the "B" model class, the response threshold models, to the models assuming that CSPC effects are due to changes in the evidence accumulation rate, the "V" model class (i.e., Models B1 and B2 vs. Models V1 and V2). The rows labeled "B vs. V" in **Table 2** report how much more likely a response threshold model is the true model than a drift model. Averaged across data sets, a model assuming a response threshold-based explanation for CSPC effects is 1.87 times more likely to be the true model than the drift model according to BIC and four times more likely according to AIC. Thus, contrary to the hypothesis that CSPC reflect contextually cued adjustments in perceptual processing selectivity (e.g., Crump et al., 2006; Lehle and Hübner, 2008; Wendt et al., 2008; King et al., 2012), the current results obtained with the LBA model suggest that this phenomenon might be better attributed to shifts in response caution primed by infrequent events within each stimulus context.

DISCUSSION

We applied the LBA model to performance from four independent flanker task data sets to adjudicate between (1) the hypothesis that CSPC effects reflect adjustments in processing selectivity cued by contextual information associated with conflict frequency (Corballis and Gratton, 2003; Crump et al., 2006, 2008; Lehle and Hübner, 2008; Wendt et al., 2008; Crump and Milliken, 2009; Heinemann et al., 2009; Vietze and Wendt, 2009; Wendt and Kiesel, 2011; Bugg and Hutchison, 2012; D'Angelo and Milliken, 2012; King et al., 2012; Sarmiento et al., 2012; for review, see Bugg and

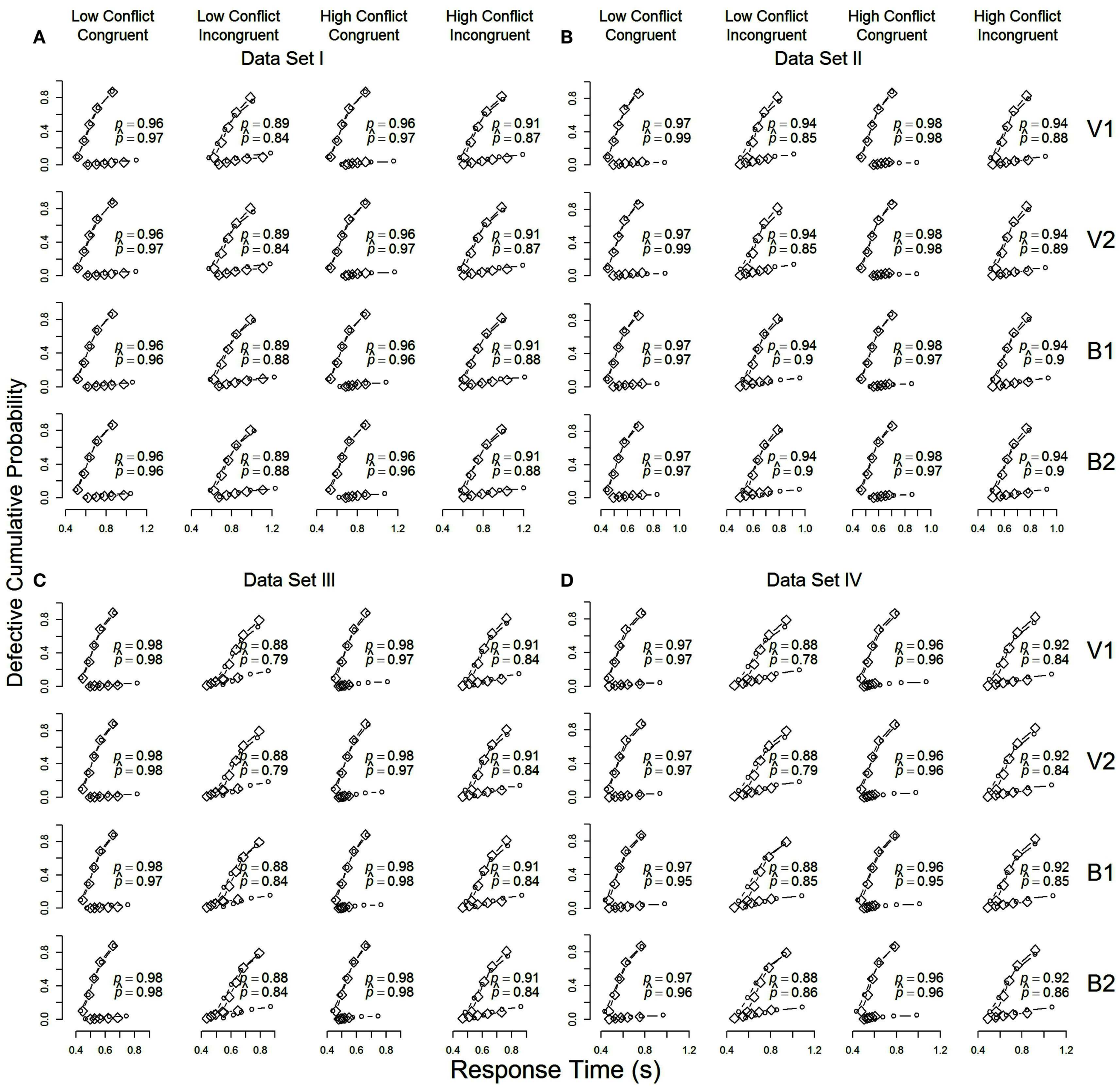

FIGURE 2 | Cumulative distribution function plots for data averaged over participants in each of the four data sets (A–D). Observed data (diamonds) and model predictions (circles) from Models V1 and V2 ("drift" models) and B1 and B2 ("threshold" models) are shown in the rows of each panel. For each condition (low- vs. high-conflict context, congruent vs. incongruent stimulus), the upper function presents results for correct response, and the lower function presents results for incorrect responses. For each condition, the observed and predicted proportion of correct responses are shown using p and $\hat{p}$, respectively.

Crump, 2012) and (2) an alternative account which attributes the phenomenon to shifts in response caution triggered by the occurrence of contextually unexpected events (e.g., incongruent trials in the low-conflict context). We predicted that if context-specific improvements in interference resolution index priming of attentional focus in favor of target stimuli, a model in which the *rate* of evidence accumulation (i.e., drift) was allowed to vary across context and congruency conditions would provide the best fit to the observed performance. In contrast, if contextual variation in the efficiency of conflict-control is attributable to shifts in response caution triggered by violations of expectancy about stimulus congruency (i.e., prediction errors), we assumed a model in which the *amount* of sensory evidence required to reach a decision (i.e., response threshold) varied according to event frequency

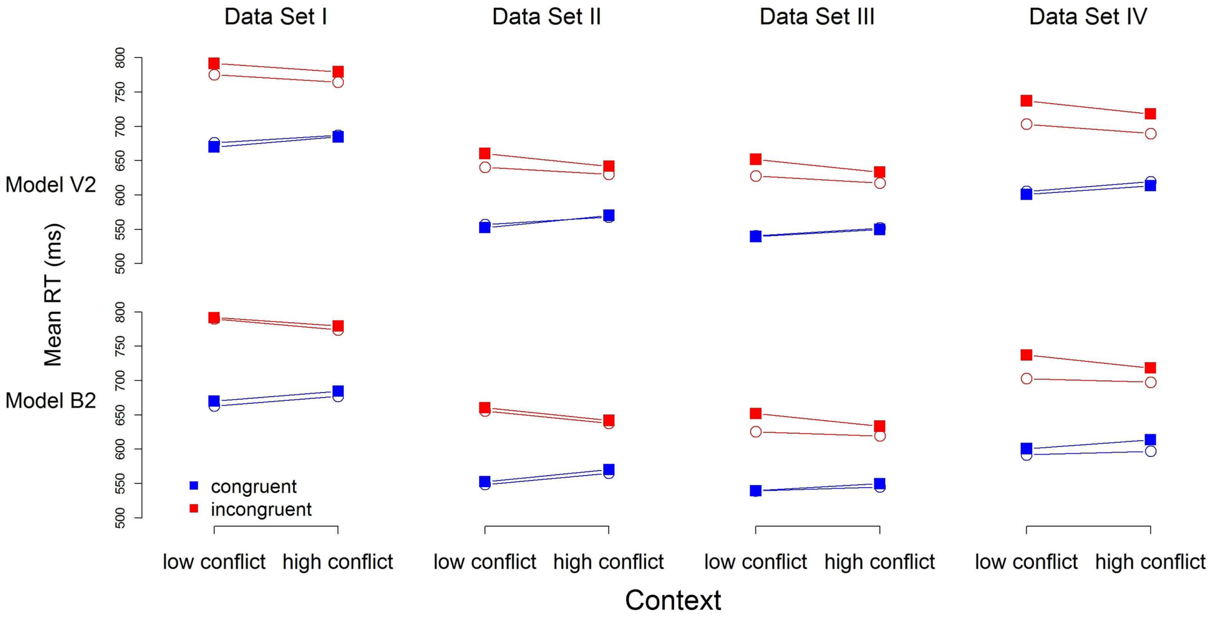

FIGURE 3 | Observed (filled squares) and predicted (open circles) mean RT for each of the four data sets. Note: RTs were calculated in a manner similar to that in **Figure 1B**, with the exception that the first trial of each block was not excluded.

Table 2 | AIC and BIC weights for Models V1 and V2 ("drift" models) and B1 and B2 ("threshold" models) for each of the four data sets.

		Data sets				Σ
		I	**II**	**III**	**IV**	
AIC	Model V1	0.059 (0)	0.050 (0)	0.185 (4)	0.152 (2)	6.9%
	Model V2	0.106 (4)	0.094 (2)	0.119 (3)	0.116 (2)	12.6%
	Model B1	0.412 (8)	0.498 (9)	0.468 (12)	0.419 (8)	42.5%
	Model B2	0.422 (13)	0.358 (8)	0.228 (6)	0.313 (6)	38.0%
	B vs. V	5.06	5.94	2.29	2.73	
BIC	Model V1	0.027 (0)	0.015 (0)	0.129 (3)	0.085 (1)	4.6%
	Model V2	0.313 (8)	0.221 (5)	0.389 (10)	0.337 (7)	34.5%
	Model B1	0.123 (3)	0.255 (4)	0.216 (4)	0.153 (1)	13.8%
	Model B2	0.537 (14)	0.509 (10)	0.267 (8)	0.425 (9)	47.1%
	B vs. V	1.94	3.24	0.934	1.37	

The row labeled B vs. V shows how much more likely that either of Models B1 or B2 is the true model compared to Model V1 or V2 according to AIC and BIC. The values in parentheses represent the number of participants best fit by each model in each data set. The sum column (Σ) shows the percentage of participants for which each model provided best fit.

within each context would deliver the best explanation of the empirical data. We implemented two "drift" and "threshold" models to account for CSPC effects. The results showed that while both classes of models captured the observed performance well (at least for correct trials), models explaining CSPC effects as resulting from shifts in response caution (i.e., the threshold models) accounted for the data better than those attributing the phenomenon to adjustments in processing selectivity (i.e., the drift models). Although evidence indicating that the threshold models provided better fit than the drift models ranged from only relatively weak (as expressed in BIC) to moderately strong (as expressed in AIC), the differences in the model selection parameters between the two classes of models were fairly consistent across the four data sets. Together, these findings provide preliminary evidence that the currently dominant view of the mechanisms underlying CSPC effects may need to be reconsidered.

Previous behavioral investigations of CSPC effects have suggested that the presentation of a stimulus in a context associated with frequent conflict primes the retrieval and execution of contextually appropriate conflict-control settings, facilitating interference resolution by enhancing processing selectivity (Corballis and Gratton, 2003; Crump et al., 2006, 2008; Lehle and Hübner, 2008;

Wendt et al., 2008; Crump and Milliken, 2009; Heinemann et al., 2009; Vietze and Wendt, 2009; Wendt and Kiesel, 2011; D'Angelo and Milliken, 2012; Sarmiento et al., 2012). Our recent fMRI study corroborated this "priming of control" hypothesis by showing that CSPC effects were mediated by activity in a region of the mSPL demonstrated to be involved in attentional control (e.g., Yantis, 2008; Shomstein, 2012) and that this activity explained top-down modulation of task-related sensory processing in visual cortex (King et al., 2012). The current modeling results qualify these previous interpretations, however, because they suggest that CSPC effects may reflect more a consequence of a shift in decision criterion triggered by contextually unexpected events than adjustments in attentional focus driven by conflict frequency.

Further insight into the putative origin of CSPC effects can be gained by considering the current results vis-à-vis those obtained by analyzing CSPC effects as a function of context transitions (i.e., context switches vs. repetitions) in our previous fMRI study. Specifically, we found in that study that while CSPC effects were observable in mSPL activation immediately upon a switch between contexts, they were observable in behavior only after context repetitions (a finding replicated here across all four data sets), suggesting that contextually appropriate control settings are rapidly retrieved in a highly flexible manner and mediate behavioral adaptation on the following trial(s) in that context (King et al., 2012). Although an analogous analysis with the LBA could not be conducted here due to an inadequate number of trials after splitting up the context and congruency conditions according to the context transition factor, it can be assumed that the mechanism suggested to mediate CSPC effects by the current modeling results (i.e., prediction error-triggered adjustments in response threshold) is also driven by context repetitions and not by context changes, in particular because CSPC effects were present only when context repeated in all experiments. It thus seems reasonable to speculate that adjustments in response threshold triggered by unexpected events would require at least one context repetition in order for a prediction regarding upcoming stimulus congruency to be in place. This view implies that contextual conflict-control settings entail predictions regarding upcoming congruency such that the relative performance gain for contextually likely stimuli (e.g., incongruent trials in the high-conflict context) and the relative performance decrement for contextually unlikely stimuli (e.g., incongruent trials in the low-conflict context) which comprise CSPC effects reflect fulfillment and violation of expectations, respectively. Such a proposition would be generally in line with the notion that the cognitive system promotes processing efficiency and goal-directed performance by continuously generating models of the environment according to current contextual demands and information stored in memory to predict future stimulus input (Friston, 2005). In any event, this novel perspective on the putative origin of CSPC effects would not have been possible from traditional analyses of behavior or functional neuroimaging alone.

It should be noted, however, that the present results provide only tentative evidence for notion that CSPC effects reflect prediction error-triggered adjustments in response caution, and some caveats should be kept in mind when interpreting our data. First, even though the LBA provided reliably good fit to the current empirical data, the model was not originally conceived to account for behavior on tasks in which the information being accumulated changes in quality over time. Many of the current theories for the flanker task assume, however, that an attentional window narrows in on the target stimulus either gradually (Eriksen and St James, 1986; Cohen et al., 1992) or abruptly (Hübner et al., 2010) over the course of a trial, thus improving the quality of evidence as time progresses. This is in direct contrast to the fundamental assumption of the LBA that evidence accumulation rate is constant over time (Brown and Heathcote, 2008). Second, the current modeling effort is at odds with another basic assumption of sequential sampling models, namely, that response threshold is already set prior to evidence accumulation. By contrast, both of the favored Models B1 and B2 captured CSPC effects by letting response threshold be adjusted according to stimulus congruency. This leads to the theoretically problematic proposition that congruency is already "known" before the start of evidence accumulation. Nevertheless, it could be argued that these models are in principle feasible, in particular because the stimuli in the present studies do not need to be analyzed to a high degree in order to distinguish congruent from incongruent trials, given the pronounced perceptual difference between congruent and incongruent arrays at a Gestalt level (cf. Baylis and Driver, 1992). Moreover, the detection of stimulus congruency (or of a stimulus as being perceptually surprising) as such is of no help in deciding whether the target face is oriented to the left or right. Therefore, the assumption that congruency or stimulus Gestalt can be detected (and affect threshold settings) *before* the decision-making process regarding target face orientation has been completed is not implausible. In sum, shifts in response threshold could feasibly be driven by a fast perceptual oddball detection occurring immediately following initial encoding of lower-level stimulus attributes, but prior to any in-depth stimulus processing or categorization according to a higher-level criterion such as target face orientation.

In future research, we aim to explore whether results similar to those reported here would be delivered by recent adaptations of sequential sampling models that were designed specifically to accommodate decision making in flanker tasks and avoid the issues outlined above, such as the spotlight diffusion model (White et al., 2011). We did not use White et al.'s model in the current analysis simply because it was not practically possible for us to achieve optimal model fits to the near 100 individual participants for all model parameterizations within a reasonable time period. The advantage of a time-varying rate of evidence accumulation in White et al.'s model is clear, but since it must be simulated (involving under optimal computing conditions several hours per model per subject), we opted to use the more computationally efficient LBA model (requiring less than a minute per model per subject) for the current project. One might speculate that a diffusion-like model in which drift rate can rapidly accelerate or decelerate within-trials as a function of fulfillment or violation of contextual expectancies regarding stimulus congruency would provide a better account of CSPC effects than the favored threshold models as revealed here with the LBA.

It is promising nonetheless that despite the LBA's possible misspecification, the model provided good fit to the observed RTs across data sets, at least on correct trials. Although we cannot rule out that the relative misfit for error trial RT distributions was not a consequence of the violations outlined above, we speculate that it may be attributable to the overall high performance/relatively low error rates and fast error RTs in the current experiments. That is, since the fast error RTs did not occur in all subjects or data sets, it is unclear to what extent they are reliable and should be used to discount the applicability of a model like the LBA. Future studies using similar protocols might create conditions that are more error prone, for instance, by reducing the stimulus presentation time. Such data would help determine whether the misfits here are simply an artifact of the fitting method, or reveal a true misfit of the model.

If we take the present results at face value, however, they provide initial support for an intriguing alternative account of CSPC effects. According to this new hypothesis, subjects implicitly encode the stimulus statistics (i.e., frequency of different trial types) associated with each context, just like in the currently dominant view of the phenomenon. However, instead of selectively enhancing their attentional focus to stimuli presented in the high-conflict context, they may form perceptual expectations for the frequent trial types in both the high- and low-conflict contexts (presumably to optimize perceptual inference and/or response selection). When expectations in a given context are violated, perceptual prediction errors in visual cortex may then be used as a control signal, indicating the need to raise response thresholds, such that sufficient evidence can be accumulated about the unexpected stimulus and a correct response can be selected. The notion that visual processing underlying perceptual inference is strongly driven by expectations and prediction error signals has garnered much empirical support in recent years (Summerfield and Koechlin, 2008; Summerfield et al., 2008; Egner et al., 2010; Jiang et al., 2012), thus supporting the basic neural feasibility of this hypothesis. Convergent electroencephalographic and fMRI evidence suggests that a subcortical-frontomedial network including the anterior mid-cingulate cortex, a region traditionally thought to be centrally involved conflict- and error monitoring (Botvinick et al., 2001; Ridderinkhof et al., 2004), may drive the type of adaptation investigated here by responding, more generally than to conflicts or errors *per se*, to any unexpected event and evaluating whether adjustments are needed (Wessel et al., 2012), confirming the core predictions of recent computational modeling work (Alexander and Brown, 2011; Silvetti et al., 2011; see also Egner, 2011). Note that, under this new perspective, CSPC effects can still be argued to constitute a reflection of "priming of control" (Spapé and Hommel, 2008; Verguts and Notebaert, 2008, 2009; King et al., 2012), but the nature of the primes and control processes differ from previous assumptions, in that they represent a shift in response caution primed by contextually surprising stimuli rather than shifts in attentional focus primed by contextual cues.

In a related literature on item-specific proportion congruent (ISPC) effects (Jacoby et al., 2003; Blais et al., 2007), there has been some debate about whether improved interference resolution for mostly incongruent items reflects a selective conflict-control mechanism that enhances processing for specific items, or merely an associative, contingency-based process by which participants learn associations between salient distracter features and responses (Schmidt and Besner, 2008; Bugg et al., 2011; Bugg and Hutchison, 2012). Crump and Milliken (2009) and Heinemann et al. (2009) both demonstrated that CSPC effects are immune to a purely associative account, because they generalize to frequency-unbiased "transfer" items. The CSPC effects in the current experiments underline these previous findings and provide further support for a control account, because they were obtained using trial-unique stimuli (i.e., the identity of the faces in the flanker array was novel on each trial) and neither stimulus congruency nor conflict-frequency context were predictive of a specific response.

Validation of the current results and their potential impact on theories of conflict-control will involve various lines of future research. First, although our modeling results were more or less consistent across all four data sets, it remains to be seen whether a "threshold" model would also provide a better account for CSPC effects than a "drift" model in other interference paradigms, such as the Stroop task. Second, as detailed above, both the experimental tasks and computational modeling approaches have scope for additional optimization for further addressing the question asked here. Additionally, new empirical protocols could be developed to provide a direct test of the notion that CSPC effects reflect increased response caution elicited by prediction errors.

In conclusion, this study suggests that CSPC effects may not necessarily reflect contextually cued attentional focus as commonly conceived, but rather shifts in response caution triggered by contextually surprising stimuli. While generally in line with the "priming of control" hypothesis (Spapé and Hommel, 2008; Verguts and Notebaert, 2008, 2009; King et al., 2012), it should be reiterated that this is the first attempt to use a model of choice and RT distributions to account for CSPC effects and more research with specialized modeling techniques that avoid the potential drawbacks of our LBA-based approach is needed to corroborate this novel perspective. If valid, the notion that expectancy violations can drive conflict adaptation effects, regardless of whether they are context-specific as in the current study, or on an item-specific (e.g., Blais et al., 2007) or sequential level (e.g., Egner, 2007), would bring important insight on the mechanisms underlying conflict-control.

ACKNOWLEDGMENTS

The work of the first and third authors was funded by National Institute of Mental Health Award R01MH087610 to the fourth author. We thank Hannah Gold, Daniel Fishman, and Matthew Nemoy for assistance in data collection.

REFERENCES

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Trans. Automat. Control* 19, 716–723.

Alexander, W. H., and Brown, J. W. (2011). Medial prefrontal cortex as an action-outcome predictor. *Nat. Neurosci.* 14, 1338–1344.

Baylis, G. C., and Driver, J. (1992). Visual parsing and response competition: the effect of grouping factors. *Percept. Psychophys.* 51, 145–162.

Blais, C., Robidoux, S., Risko, E. F., and Besner, D. (2007). Item-specific adaptation and the conflict-monitoring hypothesis: a computational model. *Psychol. Rev.* 114, 1076–1086.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., and Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychol. Rev.* 108, 624–652.

Brown, S. D., and Heathcote, A. (2008). The simplest complete model of choice response time: Linear ballistic accumulation. *Cogn. Psychol.* 57, 153–178.

Bugg, J. M., and Crump, M. J. C. (2012). In support of a distinction between voluntary and stimulus-driven control: a review of the literature on proportion congruent effects. *Front. Psychol.* 3:367. doi: 10.3389/fpsyg.2012.00367

Bugg, J. M., and Hutchison, K. A. (2012). Converging evidence for control of color–word Stroop interference at the item level. *J. Exp. Psychol. Hum. Percept. Perform.* doi:10.1037/a0029145. (in press).

Bugg, J. M., Jacoby, L. L., and Chanani, S. (2011). Why it is too early to lose control in accounts of item-specific proportion congruency effects. *J. Exp. Psychol. Hum. Percept. Perform.* 37, 844–859.

Chiu, Y. C., and Yantis, S. (2009). A domain-independent source of cognitive control for task sets: shifting spatial attention and switching categorization rules. *J. Neurosci.* 29, 3930–3938.

Cohen, J. D., Dunbar, K., and McClelland, J. L. (1990). On the control of automatic processes: a parallel distributed processing account of the Stroop effect. *Psychol. Rev.* 97, 332–361.

Cohen, J. D., Servan-Schreiber, D., and McClelland, J. L. (1992). A parallel distributed-processing approach to automaticity. *Am. J. Psychol.* 105, 239–269.

Corballis, P. M., and Gratton, G. (2003). Independent control of processing strategies for different locations in the visual field. *Biol. Psychol.* 64, 191–209.

Crump, M. J., Gong, Z., and Milliken, B. (2006). The context-specific proportion congruent Stroop effect: location as a contextual cue. *Psychon. Bull. Rev.* 13, 316–321.

Crump, M. J., and Milliken, B. (2009). The flexibility of context-specific control: evidence for context-driven generalization of item-specific control settings. *Q. J. Exp. Psychol. (Hove)* 62, 1523–1532.

Crump, M. J., Vaquero, J. M., and Milliken, B. (2008). Context-specific learning and control: the roles of awareness, task relevance, and relative salience. *Conscious. Cogn.* 17, 22–36.

D'Angelo, M. C., and Milliken, B. (2012). Context-specific control in the single-prime negative-priming procedure. *Q. J. Exp. Psychol. (Hove)* 65, 887–910.

Donkin, C., Averell, L., Brown, S., and Heathcote, A. (2009). Getting more from accuracy and response time data: methods for fitting the linear ballistic accumulator. *Behav. Res. Methods* 41, 1095–1110.

Donkin, C., Brown, S., and Heathcote, A. (2011a). Drawing conclusions from choice response time models: a tutorial using the linear ballistic accumulator. *J. Math. Psychol.* 55, 140–151.

Donkin, C., Brown, S., Heathcote, A., and Wagenmakers, E. J. (2011b). Diffusion versus linear ballistic accumulation: different models but the same conclusions about psychological processes? *Psychon. Bull. Rev.* 18, 61–69.

Egner, T. (2007). Congruency sequence effects and cognitive control. *Cogn. Affect. Behav. Neurosci.* 7, 380–390.

Egner, T. (2011). Surprise! A unifying model of dorsal anterior cingulate function? *Nat. Neurosci.* 14, 1219–1220.

Egner, T., Monti, J. M., and Summerfield, C. (2010). Expectation and surprise determine neural population responses in the ventral visual stream. *J. Neurosci.* 30, 16601–16608.

Eriksen, B. A., and Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Percept. Psychophys.* 16, 143–149.

Eriksen, C. W., and St. James, J. D. (1986). Visual attention within and around the field of focal attention: a zoom lens model. *Percept. Psychophys.* 40, 225–240.

Esterman, M., Chiu, Y. C., Tamber-Rosenau, B. J., and Yantis, S. (2009). Decoding cognitive control in human parietal cortex. *Proc. Natl. Acad. Sci. U.S.A.* 106, 17974–17979.

Forstmann, B. U., Anwander, A., Schafer, A., Neumann, J., Brown, S., Wagenmakers, E. J., Bogacz, R., and Turner, R. (2010). Cortico-striatal connections predict control over speed and accuracy in perceptual decision making. *Proc. Natl. Acad. Sci. U.S.A.* 107, 15916–15920.

Forstmann, B. U., Dutilh, G., Brown, S., Neumann, J., Von Cramon, D. Y., Ridderinkhof, K. R., and Wagenmakers, E. J. (2008). Striatum and pre-SMA facilitate decision-making under time pressure. *Proc. Natl. Acad. Sci. U.S.A.* 105, 17538–17542.

Forstmann, B. U., Wagenmakers, E. J., Eichele, T., Brown, S., and Serences, J. T. (2011). Reciprocal relations between cognitive neuroscience and formal cognitive models: opposites attract? *Trends Cogn. Sci. (Regul. Ed.)* 15, 272–279.

Friston, K. (2005). A theory of cortical responses. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 360, 815–836.

Greenberg, A. S., Esterman, M., Wilson, D., Serences, J. T., and Yantis, S. (2010). Control of spatial and feature-based attention in frontoparietal cortex. *J. Neurosci.* 30, 14330–14339.

Heinemann, A., Kunde, W., and Kiesel, A. (2009). Context-specific prime-congruency effects: on the role of conscious stimulus representations for cognitive control. *Conscious. Cogn.* 18, 966–976.

Ho, T. C., Brown, S., and Serences, J. T. (2009). Domain general mechanisms of perceptual decision making in human cortex. *J. Neurosci.* 29, 8675–8687.

Hübner, R., Steinhauser, M., and Lehle, C. (2010). A dual-stage two-phase model of selective attention. *Psychol. Rev.* 117, 759–784.

Jacoby, L. L., Lindsay, D. S., and Hessels, S. (2003). Item-specific control of automatic processes: Stroop process dissociations. *Psychon. Bull. Rev.* 10, 638–644.

Jiang, J., Schmajuk, N., and Egner, T. (2012). Explaining neural signals in human visual cortex with an associative learning model. *Behav. Neurosci.* 126, 575–581.

King, J. A., Korb, F. M., and Egner, T. (2012). Priming of control: implicit contextual cueing of top-down attentional set. *J. Neurosci.* 32, 8192–8200.

Lehle, C., and Hübner, R. (2008). On-the-fly adaptation of selectivity in the flanker task. *Psychon. Bull. Rev.* 15, 814–818.

Ludwig, C. J., Farrell, S., Ellis, L. A., and Gilchrist, I. D. (2009). The mechanism underlying inhibition of saccadic return. *Cogn. Psychol.* 59, 180–202.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: an integrative review. *Psychol. Bull.* 109, 163–203.

Mars, R. B., Shea, N. J., Kolling, N., and Rushworth, M. F. (2012). Model-based analyses: promises, pitfalls, and example applications to the study of cognitive control. *Q. J. Exp. Psychol. (Hove)* 65, 252–267.

McVay, J. C., and Kane, M. J. (2012). Drifting from slow to "d'oh!": working memory capacity and mind wandering predict extreme reaction times and executive control errors. *J. Exp. Psychol. Learn. Mem. Cogn.* 38, 525–549.

Ratcliff, R., and Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychol. Sci.* 9, 347–356.

Ridderinkhof, K. R., Ullsperger, M., Crone, E. A., and Nieuwenhuiss, S. (2004). The role of the medial frontal cortex in cognitive control. *Science* 306, 443–447.

Sarmiento, B. R., Shore, D. I., Milliken, B., and Sanabria, D. (2012). Audiovisual interactions depend on context of congruency. *Atten. Percept. Psychophys.* 74, 563–574.

Schmidt, J. R., and Besner, D. (2008). The Stroop effect: why proportion congruent has nothing to do with congruency and everything to do with contingency. *J. Exp. Psychol. Learn. Mem. Cogn.* 34, 514–523.

Schwarz, G. (1978). Estimating the dimension of a model. *Ann. Stat.* 6, 461–464.

Shomstein, S. (2012). Cognitive functions of the posterior parietal cortex: top-down and bottom-up attentional control. *Front. Integr. Neurosci.* 6:38. doi:10.3389/fnint.2012.00038

Silvetti, M., Seurinck, R., and Verguts, T. (2011). Value and prediction error in medial frontal cortex: integrating the single-unit and systems levels of analysis. *Front. Hum. Neurosci.* 5:75. doi:10.3389/fnhum.2011.00075

Smith, P. L., and Vickers, D. (1988). The accumulator model of 2-choice discrimination. *J. Math. Psychol.* 32, 135–168.

Spapé, M. M., and Hommel, B. (2008). He said, she said: episodic retrieval induces conflict adaptation in an auditory Stroop task. *Psychon. Bull. Rev.* 15, 1117–1121.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *J. Exp. Psychol.* 18, 643–661.

Summerfield, C., and Koechlin, E. (2008). A neural representation of prior information during perceptual inference. *Neuron* 59, 336–347.

Summerfield, C., Trittschuh, E. H., Monti, J. M., Mesulam, M. M., and Egner, T. (2008). Neural repetition suppression reflects fulfilled perceptual expectations. *Nat. Neurosci.* 11, 1004–1006.

Usher, M., and McClelland, J. L. (2001). The time course of perceptual choice: the leaky, competing accumulator model. *Psychol. Rev.* 108, 550–592.

van Maanen, L., Brown, S. D., Eichele, T., Wagenmakers, E. J., Ho, T., Serences, J., and Forstmann, B. U. (2011). Neural correlates of trial-to-trial fluctuations in response caution. *J. Neurosci.* 31, 17488–17495.

van Zandt, T. (2000). How to fit a response time distribution. *Psychon. Bull. Rev.* 7, 424–465.

Verguts, T., and Notebaert, W. (2008). Hebbian learning of cognitive control: dealing with specific and nonspecific adaptation. *Psychol. Rev.* 115, 518–525.

Verguts, T., and Notebaert, W. (2009). Adaptation by binding: a learning account of cognitive control. *Trends Cogn. Sci. (Regul. Ed.)* 13, 252–257.

Vietze, I., and Wendt, M. (2009). Context specificity of conflict frequency-dependent control. *Q. J. Exp. Psychol. (Hove)* 62, 1391–1400.

Wagenmakers, E. J., and Farrell, S. (2004). AIC model selection using Akaike weights. *Psychon. Bull. Rev.* 11, 192–196.

Wendt, M., and Kiesel, A. (2011). Conflict adaptation in time: foreperiods as contextual cues for attentional adjustment. *Psychon. Bull. Rev.* 18, 910–916.

Wendt, M., Kluwe, R. H., and Vietze, I. (2008). Location-specific versus hemisphere-specific adaptation of processing selectivity. *Psychon. Bull. Rev.* 15, 135–140.

Wessel, J. R., Danielmeier, C., Morton, J. B., and Ullsperger, M. (2012). Surprise and error: common neuronal architecture for the processing of errors and novelty. *J. Neurosci.* 32, 7528–7537.

White, C. N., Ratcliff, R., and Starns, J. J. (2011). Diffusion models of the flanker task: discrete versus gradual attentional selection. *Cogn. Psychol.* 63, 210–238.

Yantis, S. (2008). The neural basis of selective attention: cortical sources and targets of attentional modulation. *Curr. Dir. Psychol. Sci.* 17, 86–90.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 06 June 2012; accepted: 03 September 2012; published online: 24 September 2012.

*Citation: King JA, Donkin C, Korb FM and Egner T (2012) Model-based analysis of context-specific cognitive control. Front. Psychology **3**:358. doi: 10.3389/fpsyg.2012.00358*

This article was submitted to Frontiers in Cognition, a specialty of Frontiers in Psychology.

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 27 September 2012
doi: 10.3389/fpsyg.2012.00367

In support of a distinction between voluntary and stimulus-driven control: a review of the literature on proportion congruent effects

Julie M. Bugg[1]† and Matthew J. C. Crump[2]*†

[1] Department of Psychology, Washington University in St. Louis, St. Louis, MO, USA
[2] Department of Psychology, Brooklyn College of CUNY, Brooklyn, NY, USA

Edited by:
Rico Fischer, Technische Universität Dresden, Germany

Reviewed by:
Tobias Egner, Duke University, USA
Alexander Heinemann, Julius-Maximilians-Universität Würzburg, Germany
Caroline Gottschalk, TU Dresden, Germany

***Correspondence:**
Matthew J. C. Crump, Department of Psychology, Brooklyn College of CUNY, 2900 Bedford Avenue, Brooklyn, NY 11210, USA.
e-mail: mcrump@brooklyn.cuny.edu

† Julie M. Bugg and Matthew J. C. Crump have contributed equally to this work.

Cognitive control is by now a large umbrella term referring collectively to multiple processes that plan and coordinate actions to meet task goals. A common feature of paradigms that engage cognitive control is the task requirement to select relevant information despite a habitual tendency (or bias) to select goal-irrelevant information. At least since the 1970s, researchers have employed proportion congruent (PC) manipulations to experimentally establish selection biases and evaluate the mechanisms used to control attention. PC manipulations vary the frequency with which irrelevant information conflicts (i.e., is incongruent) with relevant information. The purpose of this review is to summarize the growing body of literature on PC effects across selective attention paradigms, beginning first with Stroop, and then describing parallel effects in flanker and task-switching paradigms. The review chronologically tracks the expansion of the PC manipulation from its initial implementation at the list-wide level, to more recent implementations at the item-specific and context-specific levels. An important theoretical aim is demonstrating that PC effects at different levels (e.g., list-wide vs. item or context-specific) support a distinction between voluntary forms of cognitive control, which operate based on anticipatory information, and relatively automatic or reflexive forms of cognitive control, which are rapidly triggered by the processing of particular stimuli or stimulus features. A further aim is to highlight those PC manipulations that allow researchers to dissociate stimulus-driven control from other stimulus-driven processes (e.g., S-R responding; episodic retrieval). We conclude by discussing the utility of PC manipulations for exploring the distinction between voluntary control and stimulus-driven control in other relevant paradigms.

Keywords: cognitive control, proportion congruent, Stroop, flanker, voluntary control, stimulus-driven control

INTRODUCTION

Selective attention paradigms such as Stroop and flanker tasks contrast performance on incongruent (i.e., incompatible) trials where multiple responses are activated by a stimulus (e.g., naming the ink color of the word RED in blue ink; responding to the central arrow in <<<><<<) to congruent (i.e., compatible) trials where a single response is activated by a stimulus (e.g., naming the ink color of the word RED in red ink; responding to the central arrow in >>>>>>>). Interference effects emerge in such tasks with slowed (and sometimes more errant) responding on incongruent/incompatible trials relative to congruent/compatible trials. Although interference effects are routinely observed, their magnitude varies substantially as a function of theoretically important factors (e.g., working memory capacity, age, and clinical status). Of current interest is a factor termed proportion congruent (PC), referring to the proportion of trials that are congruent. PC dramatically modulates the size and even the direction (Logan and Zbrodoff, 1979) of the interference effect. Paradigms with mostly congruent trials (typically 67–80%) produce significantly larger interference effects than paradigms with mostly incongruent trials.

This review summarizes the growing literature on PC effects and examines the theoretically important question of what these effects signify about cognitive control. Part of the answer rests on careful consideration of the various ways that PC is manipulated. PC has been manipulated on three distinct levels: the list-wide level (e.g., separate blocks of trials are mostly congruent or incongruent); the item level (e.g., particular words are mostly congruent or incongruent); and the context level (e.g., items presented in one context are mostly congruent, but mostly incongruent in a different context). One goal of the review is to convince the reader that manipulations of PC at each level shed light on qualitatively different cognitive control processes. List-level control operates based on anticipatory information whereas item- and context-level control are rapidly triggered by the occurrence of particular stimuli or stimulus features. List-level PC manipulations index a more voluntary form of cognitive control, whereas item- and context-level PC manipulations index a reflexive or stimulus-driven form of cognitive control. In light of findings in the PC literature, a second over-arching goal of the review is to reconsider definitions of cognitive control, and we propose one that blends conventionally separate notions of controlled and automatic processes. Redefining cognitive control in this fashion suggests a need for new terminology to better describe the processes and representations affording control.

THE MANY FACES OF COGNITIVE CONTROL

Cognitive control is by now a large umbrella term referring collectively to multiple processes that plan and coordinate behavior to meet task goals. According to convention, controlled processes are contrasted with automatic processes (Posner and Snyder, 1975; Shiffrin and Schneider, 1977). Controlled processes are voluntary, effortful, slow, and flexible. They prepare plans or task-sets that configure attention to selectively process task-relevant information during task performance. Automatic processes are involuntary, effortless, fast, and inflexible. They operate independently from controlled processes and may cause stimuli to capture attention or to retrieve associated responses. Strongly automatic processes are said to be cognitively impenetrable, or not under control. The controlled vs. automatic dichotomy has productively guided research in attention and performance over several decades. The PC literature has benefited from this distinction, but it has also produced new evidence challenging the dichotomy and conventional terminology by demonstrating that attentional control can occur in an automatic fashion. The oxymoronic term "automatic control" was coined by Jacoby et al. (2003) to describe these effects (p. 643). A contention of this review is that the controlled vs. automatic dichotomy should be abandoned and replaced by terminology that better characterizes the continuum between controlled and automatic processing (Bugg et al., 2008; Egner, 2008).

The terminology that we suggest here takes the general theory of attention and action (Norman and Shallice, 1986; Cooper and Shallice, 2000) as a starting point. Before elaborating on the major points we first describe some considerations about the concept of cognitive control that led us to adopt the terminology. The word control has different connotations for different researchers. For example, consider how cognitive vs. motor control differ.

Cognitive control refers to anticipatory, preparatory, endogenous, proactive, strategic, or voluntary processes that create, maintain, or adjust plans, task-sets, and attentional filters during performance. The spirit of this kind of control is top-down, supervisory, or executive in the sense that goals for performance are planned, monitored, and adjusted for success. In everyday life these control processes aid people in planning, thinking, and deciding on actions that will help them obtain their goals. For example, planning a driving route to run errands, focusing on a conversation with a friend in a crowded room, or surveying the field and choosing to pass to a teammate rather than an opponent in sports all rely on cognitive control processes.

Motor control refers to the processes and representations that coordinate actions. Current theories of motor control assume that motor schemas provide plans for action that are carried out by the motor system, and that online feedback from the environment and from internal simulations of the ongoing action can update and adjust movements to keep them in line with the action plan (Jordan and Rumelhart, 1992; Miall and Wolpert, 1996). Some aspects of motor control overlap with cognitive control. For example, like cognitive control, actions are planned, monitored, and adjusted by the motor system. As well, people have voluntary control of their actions. By contrast, other aspects of motor control, like the development of highly trained motor skills, overlap with the concept of automaticity. For example, the motor skills involved in driving a car are a common example of learned automatic routines. Many drivers have experienced arriving at an unintended destination like their place of work (when they originally planned a trip to the grocery store) as if they were driving on auto-pilot.

The differences between cognitive and motor control do not fit neatly into the controlled vs. automatic dichotomy, but instead speak to different levels of control. One important difference between levels is proximal vs. distal control. Proximal control refers to the representations, such as motor schemas, stimulus-response associations, and task-set representations that directly coordinate attention and action. Distal control refers to the control of proximal control; for example, by voluntary processes that select among motor schemas or task-sets, or as will be further developed in this review by exogenous cuing of proximal control representations.

Proximal control most closely resembles automatic processing. Automatic processes are commonly thought to be exogenous, involuntary, implicit, ballistic, reactive, cue/stimulus-driven, and cognitively impenetrable, or not under voluntary control; regardless, automatic processes are a fundamental component of control (cf. Hommel, 2007): they directly coordinate complex routine behaviors, and in this sense reflect proximal control of attention and action.

The interplay between proximal and distal control is insightfully framed by general theories of attention and action (Norman and Shallice, 1986; Cooper and Shallice, 2000). We outline the theory and consider its use for characterizing the multiple levels of control newly evidenced by the PC literature. The theory posits the supervisory and contention scheduling systems. The supervisory system is the distal controller, executive, or homunculus. This system "knows" current goals for action, and monitors the output of actions to ensure that goals are achieved. The system provides course correction and signals adjustments or new actions when the direction of performance has gone astray. The contention scheduling system is the workhorse and houses the proximal control representations or action schemas for familiar routines. Schemas refer to task and attentional sets, stimulus-response associations, and motor plans that provide the recipes for action needed to accomplish performance goals and sub-goals.

A fundamental assumption of the theory is that proximal representations are themselves controlled by either exogenous or endogenous means. Exogenous control refers to stimulus or cue-driven activation of associated proximal representations. For example, a coffee cup can trigger the motor movements needed to reach and grasp for the cup. Endogenous control refers to the supervisory system superseding ongoing activation of proximal representations that may lead performance astray. For example, a yellow traffic light could trigger a braking schema, but viewing an oncoming tailgater in the rear-view mirror could initiate supervisory intervention to inhibit the braking operation, and activate the schema for driving through an intersection to prevent an accident. Exogenous and endogenous forms of control are both distal in the sense that they act on the proximal representations that directly coordinate attention and action.

The theory highlights the terms proximal vs. distal control and exogenous vs. endogenous control and in doing so, preserves much of the spirit of the controlled vs. automatic distinction. Voluntary processes are capable of monitoring and adjusting attention and action, and stimuli are capable of triggering associated responses on a non-voluntary basis. In some sense the terminology simply recasts "automatic" processes as proximal control, and "controlled" processes as those involved distally in the control of control. However, the theory also captures important nuances needed to explain emerging findings in the PC literature. For example, the conventional controlled vs. automatic dichotomy does not aptly describe situations where stimuli in the environment trigger adjustments to attentional filtering that occur in a rapid-online fashion and without awareness. Here proximal control is achieved through an attentional set that directly enacts attentional filtering operations; however, the activation of this attentional set is triggered exogenously by associated cues in the environment. Multiple lines of evidence for this kind of stimulus-driven control, which is subserved by stimulus-attention associations rather than stimulus-response associations, have emerged from the PC literature, and we advance the terms proximal vs. distal and exogenous vs. endogenous as tools for describing these effects in a common terminology.

Although these terms accommodate important themes in the controlled vs. automatic dichotomy, and do not "throw out the baby with the bathwater," they also completely redefine automatic processes as being fundamental units of control (cf. Hommel, 2007). Automatic processes directly enact control over attention and action, and are distally controlled by endogenous and exogenous means. After reviewing the PC literature, we clarify these terms by distinguishing further between low (stimulus-response) and high (stimulus-attention) levels of proximal control, connecting the terminology to the range of PC phenomena, and discussing relations between levels of control more generally.

ROADMAP OF THE REVIEW

A general aim of the PC literature has been to better understand the nature of the representations and processes controlling attention and action in selective attention tasks. Progress has been made in clarifying the nature of voluntary strategic processes that influence attentional selection, and stimulus-driven processes that control attentional selection and action. We review PC findings first in the Stroop literature, and then describe parallel developments in the flanker and task-switching literatures. We focus on list-wide proportion congruent (LWPC), item-specific proportion congruent (ISPC), and context-specific proportion congruent (CSPC) manipulations. Then we discuss processes and models that could explain the findings, and connect insights from the PC literature for understanding the many faces of cognitive control to the broader attention and performance literature.

STROOP: LWPC, ISPC, AND CSPC

The Stroop task involves naming the ink-color of a color word (Stroop, 1935). Identification times are faster for congruent trials (e.g., the word red in RED ink) than incongruent trials (e.g., the word red in Blue ink). The RT difference, termed the Stroop effect, reflects a failure of attention to filter out information from the distracting word. The size of the Stroop effect can measure the effectiveness of the attentional filter. A small Stroop effect indicates strong filtering of distracting information, whereas a large Stroop effect indicates weak filtering of distracting information. PC manipulations at the list-wide, item-, and context-specific levels modulate the size of Stroop effects and provide useful tools for measuring control-based attentional adjustments.

LIST-WIDE PROPORTION CONGRUENT MANIPULATIONS

Many Stroop tasks present 50% congruent and 50% incongruent trials mixed at random. Consequently, participants are unable to accurately predict whether the next trial will be congruent or incongruent. LWPC manipulations vary the ratio of congruent and incongruent trials within a block. A mostly congruent block might be 75% congruent and 25% incongruent, and a mostly incongruent block the reverse. Stroop effects are larger for mostly congruent than mostly incongruent blocks, a finding termed the LWPC effect (e.g., Shor, 1975; Logan and Zbrodoff, 1979; Lowe and Mitterer, 1982; Logan et al., 1984; Cheesman and Merikle, 1986; Lindsay and Jacoby, 1994; West and Baylis, 1998; Kane and Engle, 2003).

Early accounts of LWPC effects assumed a role for strategic control. For example, using an ABOVE/BELOW spatial Stroop paradigm, Logan and Zbrodoff (1979) posited that participants strategically divide their attention between relevant and irrelevant dimensions, weighting the irrelevant dimension more heavily than the relevant in mostly congruent than mostly incongruent blocks (see also Logan, 1980; Logan et al., 1984, for evidence with color-word Stroop; Lowe and Mitterer, 1982). This is because the irrelevant dimension tends to validly cue the "value" of the relevant dimension (i.e., the response) in a mostly congruent list. In a similar vein, the dual-mechanisms of control account posits that participants develop expectancies about upcoming trials and modulate control proactively (e.g., Braver et al., 2007). When participants expect a congruent trial, as in a mostly congruent list, they may voluntarily pay more attention to the word, which usually corresponds to the correct response. Such a strategy would speed processing of congruent items, create strong interference for incongruent items, and increase the size of the Stroop effect. When incongruent trials are expected, participants may double-down on their attempt to filter out word information (i.e., avoid word reading). This strategy would slow down identification for congruent items (as the word would have less of a facilitating effect), speed up identification for incongruent items (less interference because of better attentional filtering), and decrease the size of the Stroop effect. Such predictions have been confirmed in some studies (e.g., Logan and Zbrodoff, 1979; Kane and Engle, 2003, Experiment 4; West and Baylis, 1998). Moreover, Lindsay and Jacoby (1994) have provided evidence from a process-dissociation procedure showing that the color-naming process (representing attention to the relevant dimension) does not vary as a function of LWPC. Rather, manipulating PC in this fashion produces a selective effect on the word reading process (representing attention to the irrelevant dimension). These data also point to a strategy that filters words differentially for mostly incongruent and mostly congruent lists.

Although strategic explanations of the LWPC effect are both parsimonious and intuitive, there has been much recent debate over the kinds of processes that may account for PC modulations to Stroop, including the LWPC effect. Not all accounts suggest use of a mechanism that relies on information about the list (i.e., the likelihood that the irrelevant dimension will be valid; the likelihood that trials will be incongruent) to strategically alter attention in advance of stimulus presentation. A competing account attributes LWPC effects to item-specific mechanisms (Bugg et al., 2008; Blais and Bunge, 2010) that operate only after a stimulus has been presented, and rely on information about particular stimuli. Before fully considering this account, we describe such item-specific mechanisms.

ITEM-SPECIFIC PROPORTION CONGRUENT MANIPULATIONS

A formative innovation was to manipulate PC at the level of individual items, rather than at the list-wide level (Jacoby et al., 2003). An ISPC manipulation assigns different PC levels to different sets of items. In the seminal study, Jacoby et al. (2003) assigned particular words to be mostly congruent or mostly incongruent. For example, the words RED and WHITE could be 80% congruent and 20% incongruent, whereas the words BLACK and GREEN could be 20% congruent and 80% incongruent. The mostly congruent and mostly incongruent items were randomly intermixed, resulting in a LWPC of 50/50 congruent and incongruent trials. Thus, participants were unable to predict whether an upcoming trial would be congruent or incongruent. That is, there was no basis for participants to form a list-wide strategy to increase word reading or filter out words. Still, a PC effect was observed indicating significantly less interference for mostly incongruent than mostly congruent items. Jacoby et al. termed this the ISPC effect, and firmly established that not all PC effects depend on having advance information about PC such as list-level information. In the ISPC paradigm, a participant could not know whether the word on a given trial was from the mostly congruent or mostly incongruent set until it was presented. As such, implementing a list-wide strategy to increase or prevent word reading would have been non-optimal (indeed, the fact that similar Stroop effects were *not* obtained for both item-types shows that such a strategy was not used).

Jacoby et al. (2003) suggested that ISPC effects may reflect rapid, online, stimulus-driven control over attentional filtering – a kind of oxymoronic "automatic control" (p. 643). On this view, individual items become associated with the attentional filters that are frequently employed for their respective item-types during the experimental session. For example, mostly congruent items become associated with an attentional filter that weakly filters word information, and mostly incongruent items become associated with an attentional filter that strongly filters word information (cf. Trainham et al., 1997; Jacoby et al., 1999). When an item appears as a stimulus on-screen it reflexively triggers the retrieval of its associated attentional filter, and this filter rapidly adjusts current attention settings to provide online control over processing of the Stroop item. Using the sample stimuli above, the idea is that when the word BLACK is presented, processing of the word is quickly attenuated. By contrast, when WHITE is presented, it triggers fuller processing of the word. In other words, the influence of the word is controlled at the item level, with the item itself acting as the environmental cue to enact a particular attentional set. Consistent with this view, process-dissociation estimates indicated that, like the LWPC manipulation, the ISPC manipulation was associated with a change in the contribution of the word process, and no change in the contribution of the color process across mostly incongruent and mostly congruent items (Jacoby et al., 2003).

An alternative view of the ISPC effect centers on an item-specific associative learning mechanism that capitalizes on the frequency with which particular words and colors are paired in ISPC designs (Jacoby et al., 2003). ISPC manipulations introduce item-frequency as a confound and ISPC effects could reflect that participants learn to respond faster to high than low frequency word-color pairs (Logan, 1988). Mostly congruent item-types repeat specific congruent items frequently and specific incongruent items infrequently (sometimes never repeated in a single block). By contrast, mostly incongruent item-types may repeat specific incongruent items frequently and specific congruent items infrequently. As such, the ISPC effect may reflect speeded responding for high-frequency items. In a similar vein, Schmidt and Besner (2008) suggested that because PC is confounded with contingency, a stimulus-response contingency-learning process may account for the ISPC effect. By their contingency account, the reason participants are faster in responding to congruent trials for mostly congruent than mostly incongruent items, and in responding to incongruent trials for mostly incongruent than mostly congruent items, is not due to item-specific control. Rather, they purport that participants learn the correlations between particular words and colors (cf. Musen and Squire, 1993; Dishon-Berkovits and Algom, 2000; Melara and Algom, 2003) and use the word to predict high contingency responses (colors). Again using the sample stimuli above, the idea is that participants learn to say "green" whenever BLACK is presented and "white" whenever WHITE is presented. The contingency account contends that the ISPC effect is entirely due to these contingency-learning processes and attentional modulation based on PC (i.e., item-specific control) plays no role in the effect.

Disentangling item-specific control and contingency learning

Item-specific control and contingency-learning accounts of the ISPC effect both assume a stimulus-driven control process, however they differ on the nature of the proximal representations enacting control. The contingency-learning account assumes that stimulus-response associations are the representation controlling action. The item-specific control account further assumes stimulus-attention associations: stimuli are associated with self-tailored attentional sets triggering rapid-online filtering of irrelevant information. There has been much debate in the literature over the contribution of item-specific control and item-specific contingency learning to ISPC effects. This question has been addressed in three ways: (a) by providing direct empirical evidence for the contingency account, and by crafting (b) item-specific designs, and (c) higher-order context-specific designs that rule out, or control for the influence of learning stimulus-response contingencies.

The primary evidence in favor of the contingency-learning account stems from Schmidt and Besner (2008, but see also Schmidt et al., 2007; Hutchison, 2011; Atalay and Misirlisoy, 2012) who re-analyzed Jacoby et al. (2003) to de-confound PC and contingency. Instead of conducting the standard analysis that compares Stroop interference for mostly congruent items (i.e., incongruent-congruent) to mostly incongruent items (i.e., incongruent-congruent), they used a contingency analysis to contrast interference for items that were equated in contingency (e.g., high contingency trials: mostly incongruent-incongruent – mostly congruent–congruent; low-contingency trials: mostly congruent incongruent – mostly incongruent-congruent). They predicted and confirmed that the contingency analysis would yield main effects of trial type and contingency but no interaction. According to Schmidt and Besner, the absence of the interaction was a key piece of evidence countering the item-specific control account, because accounts emphasizing modulation of word reading would predict "incongruent trials should be more affected by attention, given that the majority of the Stroop effect is interference with little or no facilitation from congruent trials" (p. 516).

Although Schmidt and Besner (2008) provided strong evidence in favor of the contingency account, Bugg et al. (2011a) questioned the ubiquity of the account and whether ISPC effects are always dominated by contingency learning. Their design de-confounded PC and contingency and permitted examination of the ISPC effect using the standard analysis approach. The key design feature was designating the relevant (to-be-named) dimension as the signal of ISPC rather than the irrelevant dimension, which was used in prior studies (e.g., Jacoby et al., 2003; Schmidt and Besner, 2008). When the irrelevant word dimension predicts ISPC, words signal *both* information that could be used to modulate word reading, and the most frequently paired response. When the relevant color dimension signals ISPC, contingency is equated across all four cells (combining PC and trial type) because the relevant dimension is 100% predictive of the correct response in each cell. Per a contingency account, an ISPC effect should not be obtained in this design because only PC (and not contingency) differentiates mostly congruent and mostly incongruent items. According to the item-specific control account (Bugg et al., 2011a), an ISPC effect should be obtained because participants use information signaling PC to modulate reliance on the word dimension.

In the critical experiment providing support for the item-specific control account, Bugg et al. (2011a, Experiment 2) found a significant ISPC effect using the above design in a picture-word Stroop task ("Name animal in picture, ignore word"). Moreover, the ISPC manipulation had a selective influence on incongruent trial performance with RTs speeded for the mostly incongruent than mostly congruent items, a finding consistent with Schmidt and Besner's (2008) prediction that a control mechanism would have a stronger influence on incongruent trials. In addition, Bugg et al. examined whether participants would transfer the control settings associated with mostly incongruent and mostly congruent items to a new set of stimuli. Importantly, these stimuli were new exemplars from the four animal categories that comprised the relevant dimension for training trials in the first two blocks of the task. For example, pictures of birds and cats were mostly congruent during training and pictures of dogs and fish were mostly incongruent during training. During the third block new pictures of birds, cats, dogs, and fish were presented as transfer trials and importantly these transfer trials were 50% congruent. Thus, if an ISPC effect was obtained for the transfer trials, it would suggest that participants had applied the control settings they associated with the training trials to these new transfer items. Indeed, transfer was shown. These findings are theoretically important because they challenge the contingency account, and other frequency-based accounts (e.g., Logan, 1988) that predict a RT advantage not only for mostly incongruent-incongruent trials (as was found) but also for mostly congruent-congruent trials, which was not observed.

So, where does that leave us? There is clearly evidence supporting both the contingency account and the item-specific control account. Such patterns mirror the original conclusion of Jacoby et al. (2003) who suggested a role for both processes. While such a conclusion is reasonably satisfying, it is important to understand the conditions under which one vs. the other dominates. For example, it would be prudent for researchers interested in stimulus-driven *control* to employ the design used by Bugg et al. (2011a, Experiment 2) rather than Jacoby et al. (2003). Bugg et al. proposed the basis of the ISPC signal as a design principle to differentiate ISPC designs producing effects reflecting cognitive control vs. contingency learning. When the relevant dimension signals ISPC, ISPC effects are control-based (see Bugg et al., Experiments 1 and 2 for support), but when the irrelevant dimension signals ISPC, effects are contingency-based. In support of the latter, it was found that when the exact same design was used as in Experiment 2, but words were designated mostly congruent or mostly incongruent, an ISPC effect was obtained but all of the action was in the congruent trials, consistent with predictions of a contingency account (Bugg et al., Experiment 3).

Color-word Stroop purists might contend that evidence for item-specific control in picture-word Stroop does not imply item-specific control in color-word Stroop. Picture-word and color-word Stroop effects may tap different processes (but see van Maanen et al., 2009). For example, Dell'Acqua et al. (2007) examined the locus of the interference effect in both Stroop tasks using a psychological refractory period paradigm, and found that picture-word Stroop interference arises earlier than color-word Stroop interference. In picture-word Stroop, the locus is the perceptual encoding stage whereas in color-word Stroop the locus is the response selection stage. Given that interference may serve as a trigger for item-specific control (e.g., Blais et al., 2007; Braver et al., 2007), it is possible interference arises too late in color-word Stroop paradigms for item-specific control to effectively modulate the influence of the distracting word. Countering this concern, Bugg and Hutchison (2012) replicated the critical patterns supporting the role of item-specific control in the ISPC effect using a color-word Stroop paradigm. That is, they showed that when the relevant dimension (here, color) signaled ISPC, effectively eliminating the confound between PC and contingency, an ISPC effect was still obtained contrary to the contingency account. In addition, like the patterns observed in picture-word Stroop (Bugg et al., 2011a, Experiment 2), the ISPC effect selectively influenced performance on the incongruent trials and transfer of item-specific control settings was observed for novel 50% congruent trials that consisted of "old" mostly congruent and mostly incongruent colors paired with new words. These findings provided further support for the idea that the locus of the ISPC signal (relevant vs. irrelevant dimension) is an important factor moderating use of item-specific control vs. item-specific contingency learning. However, in another experiment, Bugg and Hutchison showed that this view may be overly simplified; signaling ISPC via words can also produce control-dominated effects.

The goal of that experiment was to return to the original design of Jacoby et al. (2003) where words signal ISPC and the confound between PC and contingency is present to determine whether there are limitations on use of contingency learning in such designs. Bugg and Hutchison (2012) hypothesized that evidence favoring the contingency account would be limited to a two-item set design. In a two-item set design, the design that was used by Jacoby et al. (2003, Experiments 2a, 2b, and 3) and Schmidt and Besner (2008) in formulating the contingency account, a single high contingency response exists for both the mostly congruent and mostly incongruent word sets. In the mostly congruent set, it is the congruent response and in the mostly incongruent set, it is the incongruent response associated with the opposite color in that set. Contrast this with a four-item set where a single high contingency response option exists for the mostly congruent set but does not exist for the mostly incongruent set. There is no high contingency incongruent response. Rather, there are three equally probable responses on incongruent trials. This means that participants cannot predict with high accuracy the response that is mostly likely on any incongruent trial during a task that employs a four-item set. Given these differences, Bugg and Hutchison predicted that although the word signals ISPC in both a two- and four-item set, contingency-learning mechanisms would dominate only in the two-item set.

Two approaches were used to determine the underlying mechanism(s) responsible for the ISPC effect in the two- and four-item sets. The first was to examine the ISPC pattern (Bugg and Hutchison, 2012). For the two-item set, a symmetrical pattern was obtained reflecting speeding of RT on congruent trials from the mostly congruent set and on incongruent trials from the mostly incongruent set, the two trial types for which a high contingency response existed. By contrast, a stronger effect of the ISPC manipulation was found for incongruent trials than congruent trials in the four-item set. In particular, the RT speeding on incongruent trials in the mostly incongruent as compared to the mostly congruent set was larger than the speeding on congruent trials in the mostly congruent as compared to the mostly incongruent set, a pattern that is similar to the control-based ISPC pattern obtained in prior studies (Bugg and Hutchison, 2012, Experiments 1 and 2; Bugg et al., 2011a, Experiments 1 and 2).

The second approach was to examine transfer performance in the two- and four-item sets (Bugg and Hutchison, 2012). Transfer was assessed by presenting "old" mostly congruent and mostly incongruent words paired with new colors in a final block of trials, and these transfer items were 50% congruent. Per a contingency account, transfer should not be obtained because participants have no prior experience predicting/naming the new transfer colors. Per an item-specific control account, transfer should be obtained if participants have learned to use the word to modulate attentional settings because the old mostly congruent and mostly incongruent words still appear on transfer trials. For the two-item set, no evidence of transfer was obtained. That is, the magnitude of interference was similar for the mostly congruent and mostly incongruent words presented in new colors. By contrast, an ISPC effect was observed for the transfer items in the four-item set. Here, less interference was observed when responding to new colors that were paired with words from the mostly incongruent set than with words from the mostly congruent set. The selective effect of transfer in the four-item set, in conjunction with the ISPC pattern itself, is consistent with the view that item-specific control dominated in the four-item set. Participants utilized the word as a signal of control, quickly attenuating its influence when the word was mostly incongruent, and more fully processing the word when the word was mostly congruent. These findings suggest an update to the item-specific control account in showing that contingency-learning mechanisms do not always dominate when words are the signal of ISPC. Rather, contingency learning appears to dominate under select conditions, such as when a two-item set is used and high contingency responses can be learned for both congruent and incongruent trials.

List-wide proportion congruent manipulations: revisited

Previously, we mentioned that some folks have posited accounts of the LWPC manipulation that are not based on a strategic control process that prepares attention in advance of stimuli, but instead reflect the operation of stimulus-driven mechanisms such as item-specific control and item-specific contingency learning (e.g., Bugg et al., 2008; Schmidt and Besner, 2008; Blais and Bunge, 2010). The possibility that LWPC manipulations trigger use of item-specific mechanisms is bolstered by the fact that LWPC is perfectly confounded with ISPC in the standard design used in LWPC studies. Mostly congruent lists are composed from stimuli that are mostly congruent at the item level. For example, if four stimuli are used, each one is presented 75% of the time in a congruent color making for an ISPC of 75% congruent. In a mostly incongruent list each of the four stimuli are presented 25% of the time in a congruent color, such that the stimuli have an ISPC level of 25% congruent. Thus, participants could be modulating word reading on an item-by-item basis rather than employing a global and sustained word reading (or word avoiding) strategy. Similarly, participants could rely on item-specific contingency learning, predicting the responses that are mostly likely for particular words upon their presentation.

An initial hint in the literature that the latter type of mechanism may be contributing to the LWPC effect was evident in one of the earliest studies on the effect. Logan et al. (1984) found that the LWPC effect was robust when two word-color contingencies were present in the lists (Experiments 1 and 2); however, the LWPC effect was absent when four colors/words were used (Experiment 3). In Experiment 3, each word was paired with only two possible colors such that four separate word-color contingencies were present in each list, and high contingency responses could be predicted on the most frequent trial type within the mostly congruent and mostly incongruent lists. Logan et al. suggested that the manipulation exceeded capacity limitations; participants could not keep in mind the four word-color contingencies that existed within the list and so they abandoned the strategy. Such a finding is unanticipated by accounts that posit a list-wide strategy of filtering out words in the mostly incongruent list and fuller processing of (e.g., reading) words in the mostly congruent list. A word-filtering strategy should minimize interference in a mostly incongruent list even when a large number of color-word contingencies are present.

A number of recent studies have examined whether the LWPC effect reflects global, list-level modulation of word reading or the (possibly strategic) learning of contingencies or item-specific control. Bugg et al. (2008) determined whether a LWPC effect is observed when item-specific influences are controlled. They created two sets of items (words/colors). One set of items (e.g., GREEN and WHITE) established LWPC. For example, in the mostly incongruent list, these two items were presented 75% of the time in the incongruent color associated with the set. In the mostly congruent list, these two items were presented 75% of the time in the congruent color. Critically, a second set of items (e.g., RED and BLUE) was presented 50% of the time as congruent and 50% of the time as incongruent in *both* lists. Thus, these items were 100% identical and presented equally frequently in the mostly congruent and mostly incongruent lists. The key comparison for evaluating whether the LWPC effect reflected non-item-level processes was the magnitude of Stroop interference for the 50% congruent items in the mostly incongruent vs. mostly congruent list. Contrary to list-level control or strategic accounts, the LWPC effect was limited to the biased set of items (GREEN and WHITE) and was not obtained for the 50% congruent items that controlled for item-level influences.

Blais and Bunge (2010) used an almost identical design as Bugg et al. (2008) and replicated their primary result, again showing no evidence of list-level control. Moreover, Blais and Bunge had participants perform the Stroop task while in an fMRI scanner. They found that the anterior cingulate and dorsolateral prefrontal cortex, two regions previously implicated in top-down (e.g., list-wide) control (Botvinick et al., 2001), were selectively activated under conditions where item-specific control was presumed to operate (i.e., in contrasts involving the biased set of items). There were no differences in activation of these regions of interest in contrasts involving the 50% congruent items across the mostly congruent and mostly incongruent blocks. Like the findings of Bugg et al., these findings strongly challenged the view that list-level control is a mechanism underlying the LWPC effect.

Bugg and Chanani (2011) pursued the issue further by investigating whether the use of small stimulus sets precluded list-level control. When PC is defined by two-item sets, high contingency responses exist for congruent trials from the mostly congruent condition and incongruent trials from the mostly incongruent condition. Bugg and Chanani speculated that participants may not have engaged list-level control, a putatively more resource demanding process (cf. Braver et al., 2007), because they were capable of quickly and accurately performing the task using associative learning (prediction of high contingency responses) on the majority of trials within the mostly congruent and mostly incongruent lists. So they increased the number of items defining the PC lists, yet maintained the set size of the 50% congruent items at two. Using a picture-word Stroop task, birds, dogs, cats, and fish comprised the biased set and pigs and seals the 50% congruent set. Again, the key question centered on whether a LWPC effect would be obtained for the items in the 50% congruent set. In this study, unlike previous studies, that effect was in fact found, as was the LWPC effect for the biased set of items. Interestingly, the size of the LWPC effect (MC interference–MI interference) was larger for the biased set of items (62 ms), for which both item-specific and list-level mechanisms could be contributing, than the 50% congruent items (39 ms), for which only list-level control could be contributing. This suggests that the confound between ISPC and LWPC in most LWPC studies could inflate the size of the LWPC effect. The obtainment of an LWPC effect for 50% congruent items was also theoretically important in revealing conditions under which the LWPC effect reflects at least some contribution of list-level control and in developing a measure (i.e., the LWPC effect on 50% congruent trials) that permits researchers to selectively gage this control strategy.

A similar conclusion emerged from the work of Hutchison (2011). Like Bugg and Chanani (2011), Hutchison used items that were matched in congruency, however, the congruency was not 50%; instead he examined items that were 67% (mostly congruent) or 33% congruent (mostly incongruent) and which were embedded in mostly congruent or mostly incongruent lists. Additionally, for the mostly incongruent items, he varied whether or not items were associated with a single high contingency response. LWPC effects emerged when comparing interference across mostly congruent and mostly incongruent lists for each item type, but were strongest for mostly congruent items. This LWPC effect cannot be accounted for by item-specific influences and highlights the interaction of global control strategies with contingency learning processes. Collectively, the findings of Hutchison, and those of Bugg and Chanani, reinvigorated the list-level control account of LWPC effects and the idea that global strategies are sometimes used in resolving Stroop interference.

Following suit, Bugg et al. (2011b), sought converging evidence for the operation of list-level control using a slightly different method involving neutral trials (non-color words presented in different ink colors) that might be analogized to 50% congruent items in prior methods (Bugg et al., 2008; Blais and Bunge, 2010; Bugg and Chanani, 2011) in that they have no item-specific bias. Neutral trials are 100% neutral regardless of the overall bias of the list in which they reside. In their first experiment, neutral trials were embedded in mostly congruent, mostly incongruent, and mostly neutral lists, and six color-word stimuli were used so as to bias participants away from relying on associative/contingency learning (Bugg and Chanani). Mostly neutral lists were included to gain leverage on the question of what factors potentially trigger engagement of a list-level control strategy. Some models propose that the presence of a high degree of response conflict is a key determinant of top-down control processes used to minimize interference (e.g., Botvinick et al., 2001). It is possible, however, that list-level control is engaged whenever the irrelevant (word) dimension has little utility to responding, even when that dimension creates negligible response conflict (e.g., when most trials are neutral; cf. Melara and Algom, 2003). On this view, evidence for list-level control would be present in both the mostly incongruent and mostly neutral lists relative to the mostly congruent list. Here, such evidence would be a speeding of response times on neutral trials [in addition to reduced interference (i.e., incongruent-congruent RTs) in the same conditions]. Indeed, this is precisely what was found suggesting that a list-level strategy for attenuating interference such as word filtering was engaged in both the mostly incongruent and mostly neutral lists.

In a second experiment, Bugg et al. (2011b) used a third method for assessing the contributions of list-level control to the LWPC effect. Participants performed a Stroop task with an LWPC manipulation. However, they were also asked to perform a secondary prospective memory task during the Stroop task. Participants had to remember to press a response key (Stroop responses were vocal) whenever they encountered the word HORSE. In a control condition, participants pressed the response key whenever they encountered a particular pattern surrounding the Stroop stimulus. If participants implement a list-level word-filtering strategy in the mostly incongruent list, then performance on the secondary task should be impaired but only when the secondary task requires responding to a particular word and not when it requires responding to a particular pattern. As expected, less Stroop interference was observed in the mostly incongruent vs. mostly congruent list. This finding, however, did not adjudicate between item-specific and list-level processes because item-specific control, for example, could produce a similar pattern. Critically, the reduction in interference in the mostly incongruent list was accompanied by impairment in secondary task performance (relative to the mostly congruent list), and the impairment was specific to the word HORSE condition. The fact that the impairment was observed only for the word condition and not for the pattern condition was important in ruling out accounts of the impairment based on the difficulty of the ongoing Stroop task, which some might argue is higher when most trials are incongruent. These results further support the role of a list-level control strategy that modulates word reading even prior to stimulus onset; it is unclear how an item-specific mechanism that acts post-stimulus onset would account for the pattern of findings on neutral trials across the two experiments (Bugg et al., 2011b).

CONTEXT-SPECIFIC PROPORTION CONGRUENT MANIPULATIONS

A different approach to evaluating the item-specific control and contingency learning accounts stems from a third category of PC manipulations, termed CSPC manipulations. Here, PC is varied between different contexts in which the same items are presented. If features of an item can rapidly trigger attentional filters tailored to processing of particular items, then environmental cues that are associated with particular items, such as the location context in which an item appears, may also act as stimuli for triggering rapid-online control over attentional filtering. In a seminal study, Crump et al. (2006) used a prime-probe version of Stroop. A word (prime) was presented at fixation followed by a congruent or incongruent color patch (probe) that appeared randomly above or below fixation. The location of the color patch defined the context for the PC manipulation. For example, probes appearing above fixation were mostly congruent (75%) and probes appearing below fixation were mostly incongruent (75%). As with ISPC procedures, LWPC was 50/50 congruent and incongruent. Here again, Stroop effects were larger for probes appearing in the mostly congruent than mostly incongruent locations. Such CSPC effects have also been observed in a more traditional Stroop paradigm using font, rather than location, as the contextual cue (Bugg et al., 2008). Importantly, in these designs, all word-color pairs were presented with equal frequency and rule out accounts based on stimulus-response learning (e.g., associative/contingency). However, even in these designs an event-frequency learning process sensitive to unique word-color-location (or word-color-font) compounds could account for the observed CSPC effects.

Crump and Milliken (2009) addressed the event-frequency confound by manipulating PC both at the context and item-level. Two item-types were defined: context and transfer items. Context items carried the PC manipulation and were necessarily frequency biased. For example, red and green Stroop items were 100% congruent when they appeared above fixation and 100% incongruent when they appeared below fixation. Transfer items were not frequency biased. For example, yellow and blue items were 50% congruent and 50% incongruent in both locations. Both context and transfer items were mixed together and presented randomly in both locations within each block of trials. The question of interest was whether the attentional filter applied to the context items in their respective locations would generalize to the frequency-unbiased transfer items appearing in those locations. Indeed, CSPC effects were observed for the transfer items, with larger Stroop effects for transfer items when they appeared in the mostly congruent vs. mostly incongruent contexts. This transfer effect provides a clear example of rapid, online, context-triggered control adjustments.

PARALLEL DEVELOPMENTS IN THE FLANKER LITERATURE

The issues raised by PC research in the Stroop literature apply across selective attention tasks. A common feature of attention tasks is that they present participants with information selection problems. The Stroop task measures ability to select word from color information. The flanker task measures spatial attention ability to select central and ignore distracting peripheral information (Eriksen and Eriksen, 1974). Typically participants identify a target (e.g., a T or L) that is flanked by compatible (T T T) or incompatible (LTL) distractors. Responses are faster for compatible than incompatible trials indicating a failure of distractor suppression. As in the Stroop task, modulations to the size of the flanker effect measure processes that adjust attention filters (albeit spatial filters) to enhance or suppress processing of peripheral information. Thus, the flanker paradigm offers an opportunity to investigate how various forms of control coordinate spatial attention.

List-wide proportion congruent, ISPC-like and CSPC manipulations similar to those applied in Stroop have been shown to control the size of the flanker effect (e.g., Miller, 1987; Cohen et al., 1999; Corballis and Gratton, 2003; Lehle and Hübner, 2008; Wendt et al., 2008). As is the case for the LWPC manipulation in Stroop, the same manipulation in flanker tasks produces a similar pattern with larger compatibility effects for mostly compatible than mostly incompatible lists (e.g., Gratton et al., 1992; Lehle and Hübner, 2008, Experiment 2 Training Block performance; Taylor, 1977, Experiment 2; Wendt and Luna-Rodriguez, 2009). Unlike in Stroop, however, researchers have yet to examine whether a list-level strategy makes any contribution independent of item-specific influences (e.g., item-specific control or contingency learning). One possible reason this line of investigation has not been pursued is because ISPC designs and effects have not received much focus in flanker tasks and there has not been an empirical challenge to list-level explanations. This is a ripe area for future investigation.

In the realm of CSPC manipulations, by contrast, much work has been done exploring the flanker task. For example, Corballis and Gratton (2003) presented flanker items in different location contexts correlated with different levels of PC. Larger compatibility effects were observed for mostly congruent than mostly incongruent contexts. The locations were to the left and right of fixation because they were interested in determining whether cognitive control processes could become lateralized across hemispheres. Their hemispheric control hypothesis assumes that each hemisphere is capable of representing distinct attentional sets for controlling information specific to the processing demands of information presented to each hemisphere. CSPC effects in flanker tasks have been shown for up to four unique locations (Wendt et al., 2008), indicating a rapid and flexible engagement of attentional settings depending upon the location in which an item appears.

An important question that has been addressed in CSPC studies concerns the types of contextual cues that are effective in producing context-triggered control adjustments. In the initial study of Crump et al. (2006), a striking asymmetry was observed such that location but not shape-based contextual cues (i.e., whether the color patch probe was a square or circle) produced CSPC effects (for a similar pattern, see Crump et al., 2008). The fact that location-based cues may be processed automatically (Logan, 1998) offers one explanation for the asymmetry. Context-triggered control adjustments are presumed to occur very rapidly post-stimulus onset, and subtle differences in the speed with which the context is identified (i.e., location is faster than shape) could drive which cues are useful signals of PC. Crump et al. (2008) tested another explanation, the relevance hypothesis, which proposes that a particular contextual cue will be effective to the extent that is relevant to the current task, and thus attended (Nissen and Bullemer, 1987). Crump et al. speculated that location-based information, although nominally as irrelevant as shape in their prime-probe Stroop paradigm, might generally receive greater attention due to the importance of orienting to location in order to identify other stimulus attributes (e.g., name color). To test their hypothesis, shape was made relevant by asking participants to count the number of probes that were squares (or circles) while performing the Stroop task. Initial support for the relevance hypothesis was obtained. The shape-based cue, which was previously ineffective in triggering context-specific adjustments in control, produced a CSPC effect when attention was directed to the shape dimension.

Lehle and Hübner (2008) examined whether another identity-based cue, color, would produce a CSPC effect. They used a flanker task that included numerals as stimuli, and participants judged whether the central target was odd or even. Stimuli presented in one color (e.g., green) were 80% congruent while stimuli presented in the second color were 20% congruent. A 50 ms compatibility effect was obtained but the magnitude of the compatibility effect did not vary as a function of whether the stimuli appeared in the mostly congruent or mostly incongruent color. The results of their first experiment, thus, supported those of Crump et al. (2006, 2008), Experiment 1b) in showing an absence of an identity-based CSPC effect. In a second experiment, however, Lehle and Hübner obtained the effect. The primary change was that participants initially completed a set of training blocks wherein they experienced a fixed association between stimulus color and PC. The goal was for participants to learn the association between green and 80% congruent, for example, and red and 20% congruent prior to performing the CSPC task, where green and red stimuli were randomly intermixed. The size of the compatibility effect was similar to Experiment 1 (54 ms); however, the effect was modulated by CSPC with a smaller effect observed for stimuli presented in the mostly incongruent color.

Color-based CSPC effects in the flanker task could depend on the existence of pre-learned associations between color and PC; however, the findings of Vietze and Wendt (2009) challenge this view. A letter-based version of the flanker task (e.g., SSHSS) in which stimuli were presented in yellow or green was used. One color was associated with a high (or low, in separate blocks) level of PC and the other was 50% congruent. Interference was reduced for the stimuli associated with a low as compared to high PC. This suggests that a color-based CSPC effect can be obtained without any prior training on the associations between color and PC. An alternative explanation is that context-level and list-level control were both at play. When only one color within a block is biased (high or low PC level) and the other is 50% congruent, the overall list has a slight bias (e.g., 64% congruent or incongruent), unlike the lists in typical CSPC paradigms (e.g., Crump et al., 2006, 2008; Lehle and Hübner, 2008; Crump and Milliken, 2009). As such, it is possible that part of the reduction in interference that was observed for mostly incongruent colors involves use of advance information (to alter attention) regarding the likelihood of interference within the list. Of course, if list-level control were making a robust contribution, one would have expected a reduction in interference for the 50% congruent color when paired with a mostly incongruent color as well (cf. Bugg and Chanani, 2011), and that was not found.

Finally, researchers have also examined whether temporal information can serve as a contextual cue for cognitive control adjustments. At least one study has found that temporal information, such as the duration of a fore period (200 vs. 1200 ms) preceding stimulus onset, is an effective cue for carrying the CSPC manipulation (Wendt and Kiesel, 2011).

As in the Stroop task, researchers using priming procedures similar to the flanker task have attempted to rule out explanations of the CSPC effect pertaining to stimulus-frequency. Heinemann et al. (2009, Experiment 1) employed a similar procedure as Crump and Milliken (2009), examining whether context-specific control adjustments would be observed for a set of frequency-unbiased items (presented equally often in both contexts). They used a prime-target paradigm in which the "flanker" preceded the target, appearing in the same location as the target, rather than flanking it. Judgments of whether the target was smaller or larger than five were made on compatible (e.g., prime = 7 and target = 6) and incompatible (e.g., prime = 7 and target = 4) trials in the presence of a colored rectangle. The color of the rectangle was associated with the PC level of the accompanying stimulus. Contrary to a frequency-based account, the compatibility effect was smaller for frequency-unbiased stimuli that were accompanied by the mostly incongruent color than the mostly congruent color.

Heinemann et al. (2009) examined the role of conscious awareness in the obtainment of the CSPC effect by weakly (Experiment 1) vs. strongly (Experiment 2) masking the prime. The context manipulation did not produce differential compatibility effects when the prime was strongly masked. The authors concluded that conscious access to the incompatible prime stimuli is necessary for context-triggered adjustments in cognitive control, possibly because access allows participants to determine the prime's (distractor's) utility to processing the target information, which allows modulation of attention to the prime on subsequent trials. Note that at first blush this may seem discrepant with findings that show participants do not have conscious access to the PC manipulation (i.e., cannot report the approximate proportion congruence for each context, Crump et al., 2006). However, a subtle but important difference is that even when primes are only barely visible, this information may be sufficient for participants to develop an implicit sense of PC even if they cannot consciously report the identity of the prime.

Most recently, in a fMRI study, King et al. (2012) examined the neurophysiological underpinnings of CSPC effects in a variant of flanker that used face-stimuli as targets and distractors, and location as the cue to signal PC. Their task used unique faces on every trial and thus ruled out S-R learning as an explanation for their observed CSPC effects. Context-specific modulation of flanker interference was tied to activity in the medial superior parietal lobule that displayed functional coupling with visual regions processing the flanker stimuli. They also showed that CSPC effects depended on context repetitions across trials, and suggested that context cues may not trigger online retrieval of attention settings, but instead may instantiate or prime attentional sets that apply forward when the context repeats.

PARALLEL DEVELOPMENTS IN THE TASK-SWITCHING LITERATURE

Task-switching costs – the finding that performance costs ensue when switching rather than repeating a task– are influenced by a range of processes from higher-level preparatory and strategic processes to lower-level cue-encoding and priming processes. Task-switching costs are influenced by list-wide proportion manipulations. Note, however, that these manipulations do not center on PC but instead on proportion repeat. Task-switch costs are larger for high proportion task-repeat than for low proportion task-repeat blocks of trials (Dreisbach et al., 2002; Dreisbach and Haider, 2006; Schneider and Logan, 2006). Task-switching costs are also influenced by item-specific proportion repeat manipulations (Leboe et al., 2008). Task-switch costs are larger for items associated with a high than low proportion of repeats. Task-switching costs can also be influenced by contextual cues that are predictive of particular tasks (Mayr and Bryck, 2005, 2007; Rubin and Koch, 2006). For example, task-switching costs are reduced when tasks appear in predictive contexts (such as location) rather than in unpredictive contexts. These findings are very much in line with the idea that stimulus information can retrieve attentional control settings and apply them to adjust online performance. In this case task-sets, rather than spatial attention or word-filtering settings, are retrieved by contextual cues.

Somewhat more abstractly, task-switching costs are also influenced by context-specific proportion repeat manipulations (Leboe et al., 2008; Crump and Logan, 2010). Task-switch costs are larger in the context associated with a high than low proportion of task-repeats. In these cases, the contextual cue was not associated with a particular task, but instead associated with likelihood of switching a task. One interpretation of this latter context effect is that contextual cues can retrieve a signal that controls whether or not a recently used task-set is retrieved and applied to current performance.

CONTROL PROCESSES AND REPRESENTATIONS INVOLVED IN LWPC, ISPC, AND CSPC EFFECTS

To take stock, PC effects can take several forms (list-wide, item-, and context-specific), they are highly robust, and they have been replicated both within and across paradigm boundaries in attention. For these reasons, we view PC manipulations as a useful tool to better understand cognitive control processes in general. Earlier, we forwarded the attention and action theory (Norman and Shallice, 1986) as a tool to better classify levels of cognitive control. We distinguished between proximal and distal forms of control and endogenous and exogenous forms of control. In this section we describe how different PC effects provide insight into these forms of control.

SUPERVISORY CONTROL

Supervisory control refers collectively to those processes engaged in strategic, endogenous, anticipatory, preparatory, proactive, executive, or voluntary control. Supervisory control reflects operations of the elusive homunculus, where the intentions, plans, and strategies voluntarily adopted by a performer direct, guide, and coordinate how attention selects information in the environment. Supervisory processes comprehend task-instructions and set overarching goals for task performance like speed-accuracy tradeoffs, attention to task-relevant information, and application of task-specific rules. Supervisory control monitors ongoing performance and makes adjustments to the activation of action plans when performance runs amok. Supervisory control is endogenous and distally acts on the proximal control units that direct attention and action. Similarly, in task-specific models of attentional selection, like Stroop (Cohen et al., 1990, and in domain-general models of attention (Bundesen, 1990; Logan, 2002), the setting of weights that filter perceptual information is assumed to be under supervisory control. The setting of weights is an endogenous act of control, and the weights themselves refer to the proximal mechanisms that filter information.

Supervisory processes are often invoked to explain LWPC effects (e.g., Lowe and Mitterer, 1982). If participants become aware of the manipulation, then they have every opportunity to anticipate and prepare for upcoming congruent or incongruent trials. In the context of Stroop, this would imply that participants are capable of voluntarily adjusting the extent to which they suppress word information (Cheesman and Merikle, 1986; Balota and Faust, 2001). This is plausible, as Raz et al. (2006) demonstrated that highly suggestible subjects show reduced Stroop interference when they are told to imagine words as not word-like. Strategic control also requires effort to maintain attentional focus over the course of the task, and it is well known that PC effects vary with working memory capacity, presumably reflecting the fact that people low in working memory capacity fail to consistently maintain the attention settings required for their adopted strategy (Kane and Engle, 2003; but see Hutchison, 2011, for evidence of larger LWPC effects for individuals with low working memory capacity).

Strategic control accounts assume awareness of the PC manipulation, however participants may show list-wide effects even when unaware of the manipulation. Blais et al. (2012) found that shifts in cognitive control across mostly congruent and mostly incongruent lists largely reflected implicit knowledge of PC. Participants who were more aware of the LWPC manipulation were not more likely to show a significantly larger LWPC effect, as might be expected if participants were using a voluntary strategy based on awareness of PC. An alternative idea is that of Melara and Algom (2003) who refer to attention being differentially drawn to the irrelevant dimension in mostly congruent vs. mostly incongruent conditions depending on the irrelevant dimension's utility to responding. Note that no reference to voluntary filtering of the irrelevant dimension is made per this very viable account. Instead, the idea is that this biasing of attention occurs relatively automatically when correlations are present between the relevant and irrelevant dimension, as is the case when PC is manipulated [cf. Kinoshita et al.'s, 2011, explanation of the effects of list-wide congruency proportion on priming (with visible primes) that refers to *implicit* tracking of the prime's utility in predicting the response].

Item-specific proportion congruent and CSPC manipulations typically do not invoke voluntary control, whereby participants become aware of the associations between particular items or contexts and their likelihood of congruency, and then use this knowledge to prepare for upcoming trials. Most ISPC designs do not probe awareness of the ISPC manipulation so it is not clear whether participants have explicit knowledge of the manipulation. In an ISPC-like paradigm, Schmidt et al. (2007) found that evidence of the learning of four non-color-word contingencies was present even in participants who were explicitly unaware of the contingencies. Given that ISPC designs have many different items (usually between 4 and 8), it seems unlikely that participants would become aware of all of the item-specific associations.

It is perhaps easier to imagine that voluntary control could account for CSPC effects (cf. Heinemann et al., 2009). Many CSPC designs employ two contexts: high and low PC. It is possible that participants become aware of the CSPC manipulation and simultaneously prepare two attentional sets, one for each context. On this view, CSPC effects would reflect rapid voluntary switching of attentional set in response to contextual information. However, participants are unable to explicitly report the proportions of congruent and incongruent items in the high and low PC contexts (Crump et al., 2006). In the same set of studies, CSPC effects were observed for location cues, but not for shape cues. Following up, Crump et al. (2008) attempted to make shape cues effective by informing participants about the CSPC manipulation. Participants signed a consent form indicating they were aware of which shapes signaled high and low PC, however CSPC effects were not observed. Interestingly, at the end of the experiment participants were again probed for their knowledge of the manipulation, and at this time they failed to report the correct proportions. This leaves open the possibility that the awareness manipulation was not strong enough, however it also underscores that participants become absorbed with the task and lose knowledge of the CSPC manipulation by the end of the task.

STIMULUS-DRIVEN CONTROL

Stimulus-driven control refers to exogenous cuing or triggering of proximal representations coordinating attention and action. Stimulus-driven control covers, but is not limited to, classic automatic influences that occur in a rapid, non-voluntary, and stereotyped or inflexible manner. Well-known examples in attention include interference from distracting word or spatial information in Stroop or Flanker, attention capture by salient perceptual information (Theeuwes, 1991, 1992), or peripheral visual cuing effects (Posner and Cohen, 1984). Stimulus-driven control also guides action in the context implicit sequence learning tasks (Nissen and Bullemer, 1987), and broadly covers classical conditioning phenomena in human and animal learning (Pavlov, 1927; Rescorla and Wagner, 1972). In these examples stimuli are assumed to capture attentional resources or retrieve learned responses.

Stimulus-driven control processes are invoked to explain ISPC and CSPC effects. Here features of the item or the item's context act as environmental cues. Cues trigger associated representations controlling attention and action. In the Norman and Shallice (1986) theory stimuli retrieve action schemas. In the PC literature stimuli are assumed to retrieve associated responses or attentional settings (Jacoby et al., 2003). As we have discussed, a major debate in the PC literature aims to adjudicate between contingency learning accounts that invoke stimulus-response associations, and stimulus-driven attentional control accounts that invoke stimulus-attention associations in accounting for ISPC and CSPC effects. These representational issues are summarized in the next section. Aside from this debate, there is broad consensus that stimulus-driven processes play an important role in PC effects, and in guiding attention and action in general.

LOW AND HIGH LEVELS OF PROXIMAL CONTROL

Proximal control refers to those representations that are exogenously or endogenously brought to bear in the direct control of attention and action. The stimulus-response association is perhaps the most classic example of a representation controlling action. There is broad consensus that stimulus-response associations are a fundamental building block of performance across attention and action, and there is wide recognition that these representations mediate some, but not all PC effects.

One of the major new insights into proximal control provided by the PC literature could be termed the *stimulus-attention association.* Conventionally, stimuli are assumed to be directly associated with responses, whereas attentional settings are assumed to be controlled by supervisory processes, and not associated with or triggered by environmental cues. The stimulus-attention association allows for the possibility that environmental cues can be associated with and trigger the application of attentional filters that, like responses, have been paired together during a learning experience. This kind of representation has been invoked to explain ISPC and CSPC effects, whereby the features of the item or contextual features rapidly trigger associated attentional filters that modulate the size of congruency effects.

Stimulus-response and stimulus-attention associations point to different levels of proximal control. Stimulus-response associations can be considered as low-level proximal control. Low-level refers to the notion that stimulus-response associations are well-learned, inflexible or stereotyped responses. Once triggered they proceed with minimal internal ability to make adjustments to action. The possibility of stimulus-attention representations suggests a more flexible, higher level of proximal control. Whereas stimulus-response associations can be highly stereotyped with stimuli retrieving only those specific actions paired in the past, stimulus-attention associations can allow for flexible control by triggering generalizable attentional filters that allow selection processes to transfer across items.

The PC literature has developed precise methods to identify effects that show evidence for stimulus-attention representations, but it is not alone in providing supporting evidence. Contextual cues can direct attention toward target location in visual search (Chun and Jiang, 1998). Long-term negative priming (DeSchepper and Treisman, 1996; Grison et al., 2005), long-term inhibition of return effects (Tipper et al., 2003), long-term aftereffects in the stop-signal task (Verbruggen and Logan, 2008), and long-term priming of pop-out effects in attention capture (Thomson and Milliken, 2012a,b) show that stimuli can retrieve attentional filters applied to them from the recent and distant past. Last, prior experience with viewing natural scenes controls eye-movements and sampling of information from familiar images (Ryan et al., 2007). All of these examples demonstrate stimulus-driven control over a variety of attention processes and further support a distinction between low and high levels of proximal control.

The nature of stimulus-attention representations are currently not well understood. Crump and Milliken (2009) forwarded an episodic account whereby memory encodes the stimulus, response, contextual features, and attentional procedures or filters employed during performance. In this way, contextual information in the task environment can cue retrieval of attentional settings used in the past and apply them to control online processing in the present. This account acknowledges stimulus-driven control as a process that guides selective attention. As well, the account assumes enriched memory representations that not only code stimulus-response information, but also code the history of attentional operations that have been applied during performance. In this way, the account is similar to the event-files account Spape and Hommel (2008) forwarded to explain the selectivity of sequential modulations of the auditory Stroop effect. The modulations were limited to sequences in which the to-be-ignored word ("high" or "low") was spoken by the same voice on trial $n - 1$ and trial n. Speaker voice was apparently a (contextual) feature to which the attentional control operations associated with trial $n - 1$ were bound (along with actions, etc.). Only when the same voice was spoken on trial n were the attentional operations reactivated, leading to a performance benefit.

A critical as yet untested assumption of these accounts is that the attentional filters triggered by stimuli are themselves bound together in records of prior experience (e.g., in an episode or event file) that code stimulus, response, and attentional filtering information. With respect to the attention and action theory, this assumption is akin to saying that the contention scheduling system codes more than stimulus-response units, it also codes stimulus-attention units or perhaps stimulus-attention-response units. An alternative possibility is that attention filters are not bound in a long or short-term episodic record, but that they trigger application of attentional sets maintained outside of contention scheduling. Part of the distinction rests on whether the stimulus-triggered adjustment occurs during retrieval of bound attentional filters that are integrated into online attentional sets, or whether multiple attentional sets are maintained online and stimuli bias application of existing attentional sets. This distinction is made apparent in differing computational accounts of PC effects that are described in the next section.

COMPUTATIONAL ACCOUNTS

List-wide and item-level PC effects have been discussed in more formal computational models. The aim of these models fits well with the aims of abandoning the conventional controlled vs. automatic dichotomy, and more precisely defining the processes underlying PC effects and cognitive control in general. Perhaps the most well known of extant computational models of PC effects is the conflict-monitoring account (Botvinick et al., 2001). According to this account, the anterior cingulate acts as a conflict-monitor, a cumulative recorder of conflict on all preceding trials. When conflict is frequent, as in a mostly incongruent list, the anterior cingulate signals regions such as dorsolateral prefrontal cortex to increase its top-down influence on performance. The notion of a conflict-monitoring system also challenges the controlled vs. automatic dichotomy as the monitoring system enacts control on an automatic basis driven by conflict signals. The control is endogenous in the sense that its origin stems from the processing of conflict signals, and these signals feed global, top-down adjustments buffering against conflict in the immediate future. Botvinick et al. characterize this global influence as heightened processing of the relevant dimension (see also Egner and Hirsch, 2005), though accounts positing heightened top-down filtering of the word dimension are equally viable (Lindsay and Jacoby, 1994). Conflict-based adjustments are partly distal in the sense that the monitoring system acts on the attention processes filtering the relevant or irrelevant dimension, and they are partly proximal in the sense that there is a kind of closed processing loop between conflict detection and subsequent adjustment.

The conflict-monitoring account successfully anticipates the LWPC effect, including the finding of this effect for congruency-matched items (e.g., 50%, Bugg and Chanani, 2011; 67 or 33%, Hutchison, 2011) and neutral trials (Bugg et al., 2011b). In addition to explaining the LWPC effect via a global heightening of top-down control in mostly incongruent lists, the conflict-monitoring account might also explain the effect via a reactive heightening of control, otherwise known as conflict-adaptation. LWPC manipulations necessarily bias the number of trials that are preceded by an incongruent trial. For example, in mostly incongruent blocks, most trials are preceded by an incongruent trial. This leaves open the possibility that smaller congruency effects in mostly incongruent vs. mostly congruent blocks are driven by conflict-driven adjustments on a local, trial-to-trial basis. Such an explanation may be less likely to account for LWPC effects on congruency-matched items or neutral items, however, when such items do not share overlapping features (relevant or irrelevant) with the items that establish the bias of the list (i.e., those items that are presented as incongruent on 75% of trials in a mostly incongruent list).

Conflict-adaptation effects are more fickle under such conditions (e.g., Mayr et al., 2003; Nieuwenhuis et al., 2006).

The conflict-monitoring account has difficulty accounting for ISPC and CSPC effects. Most ISPC and CSPC designs balance congruent and incongruent items at the list-wide level. In other words, conflict occurs on 50% of trials within a list, and the degree of conflict present in a given list is equivalent for all items. The finding of differential interference effects for different items or contexts is therefore difficult to reconcile with a model that posits a global level of top-down control across the list. In addition, in the case of ISPC and CSPC manipulations, the number of trials that follow incongruent items is also balanced such that it is unlikely that sequential effects such as conflict-adaptation can account for the corresponding effects. As a case in point, Crump et al. (2006) reported a sequential analysis in their CSPC design. They found significant sequential effects – the Stroop effect was smaller for trials following incongruent than congruent trials – however, these did not interact with the CSPC effect (see also Vietze and Wendt, 2009, for a similar pattern in their study on color-based CSPC effects in the flanker task).

Several researchers have recently forwarded computational models that include item-specific mechanisms in an attempt to provide alternatives to the conflict-monitoring model. For instance, Blais et al. (2007) proposed an item-specific conflict-monitoring account of ISPC effects. The model assumes that conflict signals are item-specific rather than general. The model also assumes a role for online maintenance of multiple, item-specific attentional sets, and thus does not assume that online-adjustments are driven by memory retrieval processes. When conflict is frequently experienced for a given item (e.g., GREEN is shown frequently in white), control adjustments are made only to the relevant pathway for the specific item generating the conflict (e.g., attention to the color white is boosted for the word GREEN). Mostly incongruent items frequently cause conflict, which in turn triggers conflict-induced adjustments for those items, leading to smaller interference effects relative to mostly congruent items. Recent findings showing that item-specific control, when dissociated from contingency learning mechanisms, has a selective or more pronounced influence on incongruent trials (Bugg et al., 2011a; Bugg and Hutchison, 2012) lend support to the idea that conflict plays a role in triggering item-specific adjustments. Interestingly, this pattern is not consistently observed in CSPC paradigms that index context-specific control (e.g., Crump et al., 2006), which raises the question of whether context-specific control adjustments are triggered by the occurrence of conflict. It is possible that access to the contextual cues, perhaps especially in the case of location, occurs sufficiently rapidly such that control adjustments are triggered prior to the detection of conflict. This remains to be explored in future modeling efforts.

Another open question is how the item-specific conflict-monitoring model accounts for transfer in ISPC and possibly CSPC paradigms. If the model boosts control only for the specific item (word-color compound) producing conflict, then transfer would seem unlikely given that transfer trials typically include an old word paired with a new color (or picture) or an old color (or picture) paired with a new word (e.g., Crump and Milliken, 2009; Bugg et al., 2011a; Bugg and Hutchison, 2012).

An appeal of the model of Blais et al. (2007) is that it successfully models not only ISPC but also LWPC effects. However, with regard to the latter, it is important to note that the LWPC effects that were modeled were confounded with ISPC effects. Thus, it is uncertain whether the item-specific conflict-monitoring model can account for more "pure" indicators of list-level control such as LWPC effects for congruency-matched items (Bugg and Chanani, 2011; Hutchison, 2011). It seems rather unlikely, given that the congruency-matched items are identical in the mostly congruent and mostly incongruent lists (Bugg and Chanani; Hutchison). Because conflict is identical for these items across lists, similar levels of interference should be observed according to the model. Similarly, it is unlikely that conflict-adaptation effects resulting from item-specific conflict can account for the LWPC effect observed for congruency-matched items. In the study of Bugg and Chanani, congruency-matched items shared no features (relevant or irrelevant) with the items that established the bias of the list. According to the item-specific conflict-monitoring model, such adjustments would influence performance on the subsequent trial only when the word repeats (Blais et al., 2007), and the word never repeated when transitioning from a high conflict (biased) to a 50% congruent item (see also Hutchison, 2011). By contrast, another model that includes an item-specific component, here one that reflects a conflict modulated Hebbian learning rule, accommodates both item-specific and item-non-specific adaptation (e.g., sequential effects on non-repeating items; Verguts and Notebaert, 2008). As such, it accounts for the ISPC effect. Additionally, it might accommodate the LWPC effect for congruency-matched and neutral items if the sequential effects on the non-repeating trials in the paradigms that have revealed these LWPC effects are sufficiently robust. The model does, however, require an additional assumption to account for such effects, namely that there is a little bit of carry-over of the top-down control settings from trial $n - 1$ to trial n.

OPEN ISSUES FOR FUTURE RESEARCH

The already rich PC literature might be expanded in several theoretically important ways that would enhance our understanding of the many faces of cognitive control. For instance, PC designs involve learning about stimulus-response and stimulus-attention associations. It remains unclear whether these associations are formed by the same learning processes, and are established at the same rate. There is some evidence to suggest they may not be. For example, Jacoby et al. (2003) found that the contingency-confounded ISPC effect was present within just 16 trials. By contrast, Crump et al. (2006) and Bugg et al. (2008) found that their CSPC effects, believed to reflect stimulus-attention associations, developed more slowly, and in some cases these effects interacted with block (i.e., were not observed in initial blocks but grew stronger with time).

Second, PC effects have been studied in single session designs where the learning occurs inside the experimental session. It remains unclear whether these learning experiences establish long-term associations that would continue to influence performance a day or a month in the future, similar to the types of effects that have been observed in other attention paradigms (e.g., see DeSchepper and Treisman, 1996, for evidence of negative priming at a 1 month interval). Similarly, stimulus-attention associations may be useful for particular items and contexts in specific situations, and new stimulus-attention associations may be required when task demands change at the item or context level. The time-course with which old stimulus-attention associations interfere with learning of new stimulus-attention associations remains an open question for future research.

A third open issue concerns the experience of conflict and the formation of associations. There is some evidence that learning about PC not only depends on item-frequency, but also depends on experiencing conflict during the learning experience. For example, Crump et al. (2008) showed a location-based CSPC effect in a Stroop color-naming task with word primes and color patch probes, but CSPC effects were not observed when the task was reversed (word naming with color patch primes and word probes). The word-naming version of the Stroop task reduces the experience of conflict, and suggests that conflict may play an important role in the learning of stimulus-attention associations.

Broadly speaking, we have endorsed the view that ISPC and CSPC effects reflect stimulus-driven control whereby item or context-level cues trigger attentional adjustments. This shows that stimulus-attention associations can be tailored for specific items, and for classes of items that appear in similar contexts. The principles guiding reliance on item- or context-specific stimulus-attention associations remain unclear, however. For example, as noted above, it is not certain whether item and context-level control adjustments are both conflict triggered, nor is it clear when item-level associations would dominate over context associations or vice versa. Bugg et al. (2008) found that when words (an item-level signal) and font (a contextual signal) were correlated with PC, the PC effect was no larger than when words independently signaled PC, which might be interpreted as preliminary support for the dominance of item-specific signals of control.

Another issue is that PC designs, especially those examining manipulations other than LWPC, have largely been carried out in behavioral paradigms (but see Blais and Bunge, 2010; King et al., 2012). The neural substrates coding stimulus-attention associations remain unspecified. Imaging studies may be particularly advantageous for examining whether control is indeed anticipatory/preparatory in the context of LWPC manipulations vs. stimulus-driven in the context of ISPC and CSPC manipulations. Methodologies are available that permit the examination of sustained activation patterns, which one would expect to accompany list-level control, and transient activation patterns, which should characterize the item- and context-level adjustments. Moreover, it would be advantageous to examine the time-course of LWPC, ISPC, and CSPC effects. Cleary, the latter two reflect very rapid control adjustments, but it is not certain whether the time courses differ for ISPC and CSPC effects, or whether the trigger for such adjustments is a perceptual feature (e.g., shape differences such as font, Bugg et al., 2008) as opposed to the conflict associated with processing of the irrelevant dimension. These questions could be addressed using event-related potentials.

Finally, there is much room for application of the range of PC manipulations to other tasks in which it would be theoretically advantageous to isolate voluntary control from stimulus-driven mechanisms, and vice versa. There is evidence of LWPC effects in the Simon task (e.g., Hommel, 1994; Toth et al., 1995), which shares some features with the Stroop and flanker tasks to which we devoted much attention. Determining whether PC manipulations are useful for examining cognitive control in tasks that are quite different from Stroop and flanker (e.g., task-switching; Go No-Go) is a necessary next step in evaluating whether the levels of control concepts presented herein might have broader appeal.

FINAL THOUGHTS

People learn to optimize their performance in a complex and unruly world. It is increasingly clear that multiple levels of control guide performance during and after learning. We have distinguished broadly between endogenous vs. exogenous and proximal vs. distal aspects of control. Endogenous control highlights opportunity for volition to guide performance, and exogenous control highlights opportunity to offload control to the environment. Both endogenous and exogenous control act distally on the many kinds of proximal control representations that allow attention to filter relevant from irrelevant information and the motor system to guide action. The relationships between levels of control and the extent to which different control processes contribute to performance are not well understood.

On the one hand, investigating multiple levels of control calls for researchers to study each level independently. This involves a terminology of control that permits fine distinctions between levels and experimental rigor to create process pure measurements of each process. The PC literature is an illustration of the difficulty in achieving this rigor. Nevertheless, such an approach could potentially answer quantitative questions about control-based adjustments. For example, do voluntary, conflict-based, or stimulus-driven forms of adjusting attention all have the same power to change attentional settings? The approach might also encourage researchers to experimentally disentangle contributions of stimulus-response, stimulus-attention, conflict-based, and voluntary processes potentially mediating effects of interest.

On the other hand, it would be unfortunate if by dividing-and-conquering the territory of cognitive control it would be left without some theme of unification. Investigating multiple levels of control also calls for researchers to clarify relationships between levels. Do different forms of control act on the same proximal mechanisms of control? Do these forms of control operate exclusively from each other or are they interrelated over the course of learning? For example, are stimulus-attention associations formed strictly by incidental or implicit learning processes, or do they reflect a learning process that associates voluntary control of attention together with particular cues eventually mediating stimulus-driven control? We encourage both approaches, investigating the nuances of different forms of control and how they act in concert as paths toward understanding the manifold ways that cognitive control coordinates performance.

REFERENCES

Atalay, N. B., and Misirlisoy, M. (2012). Can contingency learning alone account for item-specific control? Evidence from within- and between-language ISPC effects. *J. Exp. Psychol. Learn. Mem. Cogn.* [Epub ahead of print].

Balota, D. A., and Faust, M. E. (2001). "Attention in dementia of the Alzheimer's type," in *Handbook of Neuropsychology*, 2nd Edn, eds F. Boller and S. Cappa (New York, NY: Elsevier Science), 51–80.

Blais, C., and Bunge, S. (2010). Behavioral and neural evidence for item-specific performance monitoring. *J. Cogn. Neurosci.* 22, 2758–2767.

Blais, C., Harris, M. B., Guerrero, J. V., and Bunge, S. A. (2012). Rethinking the role of automaticity in cognitive control. *Q. J. Exp. Psychol.* 65, 268–276.

Blais, C., Robidoux, S., Risko, E. F., and Besner, D. (2007). Item-specific adaptation and the conflict monitoring hypothesis: a computational model. *Psychol. Rev.* 114, 1076–1086.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., and Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychol. Rev.* 114, 1076–1086.

Braver, T. S., Gray, J. R., and Burgess, G. C. (2007). "Explaining the many varieties of working memory variation: dual mechanisms of cognitive control," in *Variation in Working Memory*, eds A. R. A. Conway, C. Jarrold, M. J. Kane, A. Miyake, and J. N. Towse (Oxford: Oxford University Press), 76–106.

Bugg, J. M., and Chanani, S. (2011). List-wide control is not entirely elusive: evidence from picture-word Stroop. *Psychon. Bull. Rev.* 18, 930–936.

Bugg, J. M., and Hutchison, K. A. (2012). Converging evidence for control of color-word Stroop interference at the item level. *J. Exp. Psychol. Hum. Percept. Perform.* doi:10.1037/a0029145

Bugg, J. M., Jacoby, L. L., and Chanani, S. (2011a). Why it is too early to lose control in accounts of item-specific proportion congruency effects. *J. Exp. Psychol. Hum. Percept. Perform.* 37, 844–859.

Bugg, J. M., McDaniel, M. A., Scullin, M. K., and Braver, T. S. (2011b). Revealing list-level control in the Stroop task by uncovering its benefits and a cost. *J. Exp. Psychol. Hum. Percept. Perform.* 37, 1595–1606.

Bugg, J. M., Jacoby, L. L., and Toth, J. P. (2008). Multiple levels of control in the Stroop task. *Mem. Cognit.* 36, 1484–1494.

Bundesen, C. (1990). A theory of visual attention. *Psychol. Rev.* 97, 523–547.

Cheesman, J., and Merikle, P. M. (1986). Distinguishing conscious from unconscious perceptual processes. *Can. J. Psychol.* 40, 343–367.

Chun, M. M., and Jiang, Y. (1998). Contextual cueing: implicit learning and memory of visual context guides spatial attention. *Cogn. Psychol.* 36, 28–71.

Cohen, A., Fuchs, A., Bar-Sela, A., Brumberg, Y., and Magen, H. (1999). Correlational cuing as a function of target complexity and target-flanker similarity. *Percept. Psychophys.* 61, 275–290.

Cohen, J. D., Dunbar, K., and McClelland, J. L. (1990). On the control of automatic processes: a parallel distributed processing model of the Stroop effect. *Psychol. Rev.* 97, 332–361.

Cooper, R., and Shallice, T. (2000). Contention scheduling and the control of routine activities. *Cogn. Neuropsychol.* 17, 297–338.

Corballis, P. M., and Gratton, G. (2003). Independent control of processing strategies for different locations in the visual field. *Biol. Psychol.* 64, 191–209.

Crump, M. J. C., Gong, Z., and Milliken, B. (2006). The context-specific proportion congruent effect: location as a contextual cue. *Psychon. Bull. Rev.* 13, 316–321.

Crump, M. J. C., and Logan, G. D. (2010). Contextual control over task-set retrieval. *Atten. Percept. Psychophys.* 72, 2047–2053.

Crump, M. J. C., and Milliken, B. (2009). The flexibility of context-specific control: evidence for context-driven generalization of item-specific control settings. *Q. J. Exp. Psychol. (Hove)* 62, 1523–1532.

Crump, M. J. C., Vaquero, J. M. M., and Milliken, B. (2008). Context-specific learning and control: the role of awareness, task-relevance, and relative salience. *Conscious. Cogn.* 17, 22–36.

Dell'Acqua, R., Job, R., Peressotti, F., and Pascali, A. (2007). The picture–word interference effect is not a Stroop effect. *Psychon. Bull. Rev.* 14, 717–722.

DeSchepper, B. G., and Treisman, A. M. (1996). Visual memory for novel shapes: implicit coding without attention. *J. Exp. Psychol. Learn. Mem. Cogn.* 22, 27–47.

Dishon-Berkovits, M., and Algom, D. (2000). The Stroop effect: it is not the robust phenomenon that you have thought it to be. *Mem. Cognit.* 28, 1437–1449.

Dreisbach, G., and Haider, H. (2006). Preparatory adjustment of cognitive control in the task switching paradigm. *Psychon. Bull. Rev.* 13, 334–338.

Dreisbach, G., Haider, H., and Kluwe, R. H. (2002). Preparatory processes in the task-switching paradigm: evidence from the use of probability cues. *J. Exp. Psychol. Learn. Mem. Cogn.* 28, 468–483.

Egner, T. (2008). Multiple conflict-driven control mechanism in the human brain. *Trends Cogn. Sci. (Regul. Ed.)* 12, 374–380.

Egner, T., and Hirsch, J. (2005). Cognitive control mechanisms resolve conflict through cortical amplification of task-relevant information. *Nat. Neurosci.* 8, 1784–1790.

Eriksen, B. A., and Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a non-search task. *Percept. Psychophys.* 16, 143–149.

Gratton, G., Coles, M. G. H., and Donchin, E. (1992). Optimizing the use of information: strategic control of activation and responses. *J. Exp. Psychol. Gen.* 121, 480–506.

Grison, S., Tipper, S. P., and Hewitt, O. (2005). Long-term negative priming: support for retrieval of prior attentional processes. *Q. J. Exp. Psychol. A* 58, 1199–1224.

Heinemann, A., Kunde, W., and Kiesel, A. (2009). Context-specific prime-congruency effects: on the role of conscious stimulus representations for cognitive control. *Conscious. Cogn.* 18, 966–976.

Hommel, B. (1994). Spontaneous decay of response-code activation. *Psychol. Res.* 56, 261–268.

Hommel, B. (2007). Consciousness and control: not identical twins. *J. Conscious. Stud.* 12, 155–176.

Hutchison, K. A. (2011). The interactive effects of list-based control, item-based control, and working memory capacity on Stroop performance. *J. Exp. Psychol. Learn. Mem. Cogn.* 37, 851–860.

Jacoby, L. L., Lindsay, D. S., and Hessels, S. (2003). Item-specific control of automatic processes: Stroop process dissociations. *Psychon. Bull. Rev.* 10, 638–644.

Jacoby, L. L., McElree, B., and Trainham, T. N. (1999). "Automatic influences as accessibility bias in memory and Stroop-like tasks: toward a formal model," in *Attention and Performance XVII: Cognitive Regulation of Performance. Interaction of Theory and Application*, eds D. Gopher and A. Koriat (Cambridge, MA: MIT Press), 461–486.

Jordan, M. I., and Rumelhart, D. E. (1992). Forward models: supervised learning with a distal teacher. *Cogn. Sci.* 16, 307–354.

Kane, M. J., and Engle, R. W. (2003). Working memory capacity and the control of attention: the contributions of goal neglect, response competition, and task to Stroop interference. *J. Exp. Psychol. Gen.* 132, 47–70.

King, J. A., Korb, F. M., and Egner, T. (2012). Priming of control: implicit contextual cuing of top-down attentional set. *J. Neurosci.* 32, 8192–8200.

Kinoshita, S., Mozer, M. C., and Forster, K. I. (2011). Dynamic adaptation to history of trial difficulty explains the effect of congruency proportion on masked priming. *J. Exp. Psychol. Gen.* 140, 622–636.

Leboe, J. P., Wong, J., Crump, M. J. C., and Stobbe, K. (2008). Probe-specific proportion task repetition effects on switching costs. *Percept. Psychophys.* 70, 935–945.

Lehle, C., and Hübner, R. (2008). On-the-fly adaptation of selectivity in the flanker task. *Psychon. Bull. Rev.* 15, 814–818.

Lindsay, D. S., and Jacoby, L. L. (1994). Stroop process dissociations: the relationship between facilitation and interference. *J. Exp. Psychol. Hum. Percept. Perform.* 20, 219–234.

Logan, G. D. (1980). Attention and automaticity in Stroop and priming tasks: theory and data. *Cogn. Psychol.* 12, 523–553.

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychol. Rev.* 95, 492–527.

Logan, G. D. (1998). What is learned during automatization? II. Obligatory encoding of spatial location. *Hum. Percept. Perform.* 24, 1720–1736.

Logan, G. D. (2002). An instance theory of attention and memory. *Psychol. Rev.* 109, 376–400.

Logan, G. D., and Zbrodoff, N. J. (1979). When it helps to be misled: facilitative effects of increasing the frequency of conflicting stimuli in a Stroop-like task. *Mem. Cognit.* 7, 166–174.

Logan, G. D., Zbrodoff, N. J., and Williamson, J. (1984). Strategies in the color-word Stroop task. *Bull. Psychon. Soc.* 22, 135–138.

Lowe, D., and Mitterer, J. O. (1982). Selective and divided attention in a Stroop task. *Can. J. Psychol.* 36, 684–700.

Mayr, U., Awh, E., and Laurey, P. (2003). Conflict adaptation effects in the absence of executive control. *Nat. Neurosci.* 6, 450–452.

Mayr, U., and Bryck, R. L. (2005). Sticky rules: integration between abstract rules and specific actions. *J. Exp. Psychol. Learn. Mem. Cogn.* 31, 337–350.

Mayr, U., and Bryck, R. L. (2007). Outsourcing control to the environment: effects of stimulus/response locations on task selection. *Psychol. Res.* 71, 107–116.

Melara, R. D., and Algom, D. (2003). Driven by information: a tectonic theory of Stroop effects. *Psychol. Rev.* 110, 422–471.

Miall, R. C., and Wolpert, D. M. (1996). Forward models for physiological motor control. *Neural Netw.* 9, 1265–1279.

Miller, J. (1987). Priming is not necessary for selective-attention failures: semantic effect of unattended, unprimed letters. *Percept. Psychophys.* 41, 419–434.

Musen, G., and Squire, L. R. (1993). Implicit learning of color-word associations using a Stroop paradigm. *J. Exp. Psychol.* 19, 789–798.

Nieuwenhuis, S., Stins, J. F., Posthuma, D., Polderman, T. J. C., Boomsma, D. I., and de Geus, E. J. (2006). Accounting for sequential trial effects in the flanker task: conflict adaptation or associative priming? *Mem. Cognit.* 34, 1260–1272.

Nissen, M. J., and Bullemer, P. (1987). Attentional requirements of learning: evidence from performance measures. *Cogn. Psychol.* 19, 1–32.

Norman, D. A., and Shallice, T. (1986). "Attention to action: willed and automatic control of behaviour," in *Consciousness and Self-Regulation*, eds R. J. Davidson, G. E. Schwartz, and D. Shapiro (New York: Plenum Press).

Pavlov, I. P. (1927). *Conditional Reflexes*. New York: Dover Publications.

Posner, M. I., and Cohen, Y. (1984). "Components of visual orienting," in *Attention and Performance*, Vol. X, eds H. Bouma and D. Bouwhuis (Hillsdale: Erlbaum), 51–556.

Posner, M. I., and Snyder, C. R. R. (1975). "Attentoin and cognitive control," in *Information Processing and Cognition* ed. R. L. Solso (Hillsdale: Erlbaum), 55–85.

Raz, A., Kirsch, I., Pollard, J., and Nitkin-Kaner, Y. (2006). Suggestion reduces the Stroop effect. *Psychol. Sci.* 17, 91–95.

Rescorla, R. A., and Wagner, A. R. (1972). "A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement," in *Classical Conditioning: Current Research and Theory*, eds A. H. Black and W. F. Prokasy (New York: Appleton-Century-Crofts), 64–99.

Rubin, O., and Koch, I. (2006). Exogenous influences on task set activation in task switching. *Q. J. Exp. Psychol. (Hove)* 59, 1033–1046.

Ryan, J. D., Hannula, D. E., and Cohen, N. J. (2007). The obligatory effects of memory on eye movements. *Memory* 15, 508–525.

Schmidt, J., Crump, M. J. C., Cheesman, J., and Besner, D. (2007). Contingency learning without awareness: evidence for implicit control. *Conscious. Cogn.* 16, 421–435.

Schmidt, J. R., and Besner, D. (2008). The Stroop effect: why proportion congruent has nothing to do with congruency and everything to do with contingency. *J. Exp. Psychol. Learn. Mem. Cogn.* 34, 514–523.

Schneider, D. W., and Logan, G. D. (2006). Priming cue encoding by manipulating transition frequency in explicitly cued task switching. *Psychon. Bull. Rev.* 13, 145–151.

Shiffrin, R. M., and Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychol. Rev.* 84, 127–190.

Shor, R. E. (1975). An auditory analog of the Stroop test. *J. Gen. Psychol.* 93, 281–288.

Spape, M. M., and Hommel, B. (2008). He said, she said: episodic retrieval induces conflict adaptation in an auditory Stroop task. *Psychon. Bull. Rev.* 15, 1117–1121.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *J. Exp. Psychol.* 18, 643–662.

Taylor, D. A. (1977). Time course of context effects. *J. Exp. Psychol. Gen.* 106, 404–426.

Theeuwes, J. (1991). Cross-dimensional perceptual selectivity. *Percept. Psychophys.* 18, 643–662.

Theeuwes, J. (1992). Perceptual selectivity for color and form. *Percept. Psychophys.* 51, 599–606.

Thomson, D. R., and Milliken, B. (2012a). Perceptual distinctiveness produces long-lasting priming of pop-out. *Psychon. Bull. Rev.* 19, 170–176.

Thomson, D. R., and Milliken, B. (2012b). Contextual distinctiveness produces long-lasting priming of pop-out. *J. Exp. Psychol. Hum. Percept. Perform.* [Epub ahead of print].

Tipper, S. P., Grison, S., and Kesler, K. (2003). Long-term inhibition of return attention. *Psychol. Sci.* 14, 19–25.

Toth, J. P., Levine, B., Stuss, D. T., Oh, A., Winocur, G., and Meiran, N. (1995). Dissociation of processes underlying spatial S-R compatibility: evidence for the independent influence of what and where. *Conscious. Cogn.* 4, 483–501.

Trainham, T. N., Lindsay, D. S., and Jacoby, L. L. (1997). Stroop process dissociations: reply to Hillstrom and Logan (1997). *J. Exp. Psychol. Hum. Percept. Perform.* 23, 1579–1587.

van Maanen, L., van Rijn, H., and Borst, J. P. (2009). Stroop and picture–word interference are two sides of the same coin. *Psychon. Bull. Rev.* 16, 987–999.

Verbruggen, F., and Logan, G. D. (2008). Long-term aftereffects of response inhibition: memory retrieval, task goals, and cognitive control. *J. Exp. Psychol. Hum. Percept. Perform.* 34, 1229–1235.

Verguts, T., and Notebaert, W. (2008). Hebbian learning of cognitive control: dealing with specific and nonspecific adaptation. *Psychol. Rev.* 115, 518–525.

Vietze, I., and Wendt, M. (2009). Context specificity of conflict frequency-dependent control. *Q. J. Exp. Psychol. (Hove)* 62, 1391–1400.

Wendt, M., and Kiesel, A. (2011). Conflict adaptation in time: foreperiods as contextual cues for attentional adjustment. *Psychon. Bull. Rev.* 18, 910–916.

Wendt, M., Kluwe, R. H., and Vietze, I. (2008). Location-specific versus hemisphere-specific adaptation of processing selectivity. *Psychon. Bull. Rev.* 15, 135–140.

Wendt, M., and Luna-Rodriguez, A. (2009). Conflict-frequency affects flanker-interference. *Exp. Psychol.* 56, 206–217.

West, R., and Baylis, G. C. (1998). Effect of increased response dominance and contextual disintegration on the Stroop interference effect in older adults. *Psychol. Aging* 13, 206–217.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 15 June 2012; accepted: 07 September 2012; published online: 27 September 2012.

*Citation: Bugg JM and Crump MJC (2012) In support of a distinction between voluntary and stimulus-driven control: a review of the literature on proportion congruent effects. Front. Psychology **3**:367. doi: 10.3389/fpsyg.2012.00367*

This article was submitted to Frontiers in Cognition, a specialty of Frontiers in Psychology.

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 12 October 2012
doi: 10.3389/fpsyg.2012.00397

A cue from the unconscious – masked symbols prompt spatial anticipation

Heiko Reuss[1]*, Andrea Kiesel[1], Wilfried Kunde[1] and Peter Wühr[2]

[1] Department of Psychology III, Julius-Maximilians-Universität Würzburg, Würzburg, Germany
[2] Department of Psychology, Technische Universität Dortmund, Dortmund, Germany

Edited by:
Rico Fischer, Technische Universität Dresden, Germany

Reviewed by:
Mike Le Pelley, University of New South Wales, Australia
Friederike Schlaghecken, University of Warwick, UK

***Correspondence:**
Heiko Reuss, Department of Psychology, University of Würzburg, Röntgenring 11, 97070 Würzburg, Germany.
e-mail: reuss@psychologie.uni-wuerzburg.de

Anticipating where an event will occur enables us to instantaneously respond to events that occur at the expected location. Here we investigated if such spatial anticipations can be triggered by symbolic information that participants cannot consciously see. In two experiments involving a Posner cueing task and a visual search task, a central cue informed participants about the likely location of the next target stimulus. In half of the trials, this cue was rendered invisible by pattern masking. In both experiments, visible cues led to cueing effects, that is, faster responses after valid compared to invalid cues. Importantly, even masked cues caused cueing effects, though to a lesser extent. Additionally, we analyzed effects on attention that persist from one trial to the subsequent trial. We found that spatial anticipations are able to interfere with newly formed spatial anticipations and influence orienting of attention in the subsequent trial. When the preceding cue was visible, the corresponding spatial anticipation persisted to an extent that prevented a noticeable effect of masked cues. The effects of visible cues were likewise modulated by previous spatial anticipations, but were strong enough to also exert an impact on attention themselves. Altogether, the results suggest that spatial anticipations can be formed on the basis of unconscious stimuli, but that interfering influences like still active spatial anticipations can suppress this effect.

Keywords: endogenous shifts of attention, anticipation, unconscious processing, spatial cueing, masked priming

INTRODUCTION

Humans can give priority to spatial locations where behaviorally relevant stimuli occur, a process referred to as spatial attention. Such orienting of attention can happen in two different ways, either *exogenously* driven or *endogenously* controlled (e.g., Posner, 1980; Jonides, 1981; Posner and Cohen, 1984; Müller and Rabbitt, 1989; Yantis and Johnson, 1990; Yantis and Jonides, 1990; Theeuwes, 1991; Folk et al., 1992). On the one hand, exogenous orienting of attention is induced by particular events in the environment. Here, anticipation plays a role insofar as only events that are behaviorally relevant are able to capture attention. If, for example, participants search for targets that abruptly onset on a screen, cues that abruptly onset automatically capture attention (Folk et al., 1992). Interestingly, if participants anticipate particular behaviorally relevant features, cues that owe these features grab attention automatically, even if they are overall not predictive for the target location.

On the other hand, humans can deliberately orient attention to certain locations in space, or in Helmholtz's words, "it is possible, simply by a conscious and voluntary effort, to focus the attention on some definite spot in an absolutely dark and featureless field" (von Helmholtz, 1866, cited after Yantis, 1998, p. 225). Typically, such conscious efforts are suggested to the participant by some symbolic cue presented in the center of vision (Posner, 1980; Posner et al., 1980). Again, anticipation is a necessary process for such cues to work, but at a different point in time, namely after rather than before cue presentation. Only if subjects anticipate targets at the cued location will central cues leave a trace in performance. Endogenously controlled shifts of attention are only executed when the target in fact appears at the cued location more often than not and is thus anticipated there. Perhaps anticipation of the target location is what we typically describe as cueing effects or validity effects: faster response times (RT) to targets at validly cued compared to invalidly cued locations.

The distinction between these two forms of orienting of attention and also their dependency on awareness is nicely illustrated in a study by McCormick (1997). The cues in this experiment were peripheral bars that were either visible or masked. Critically, the target appeared at the opposite location of the cue in 85% of the trials, so that participants would anticipate the target at the non-cued location. McCormick reasoned that when a cue appeared, it would at first capture attention exogenously. However, this exogenous cue could then be used strategically by the participants to endogenously shift their attention to the opposite location, where they anticipated the target. McCormick found that with visible cues, participants were indeed reorienting their exogenously captured attention in anticipation of the target. When the cues were masked, however, performance was better when the target appeared at the location of the cue. This indicates that the masked cues were able to exogenously capture attention, but that the participants were not able to voluntarily reorient their attention when the cue was masked. Subsequent work has confirmed many times that masked cues trigger exogenous shifts of attention (McCormick, 1997; Lambert et al., 1999; Scharlau, 2002; Ivanoff and Klein, 2003; Scharlau and Ansorge, 2003; Scharlau and Neumann, 2003; Ansorge and Neumann, 2005; Ansorge and Heumann, 2006; Mulckhuyse et al., 2007; for a review, see Mulckhuyse and Theeuwes, 2010).

While the possibility of exogenous cues to work outside of awareness is in line with classical theories of automaticity and control (Atkinson and Shiffrin, 1968; Posner and Snyder, 1975), the more intriguing question is the relation of endogenously controlled orienting of attention and consciousness. In recent years, a steadily growing field of research is concerned with this relation of consciousness and cognitive control processes. For example, it was shown that the activation of task sets, a typical instance of cognitive control, can be triggered unconsciously by masked task cues (Mattler, 2006; Lau and Passingham, 2007; Reuss et al., 2011a). Also, there are findings that inhibition, a cognitive control process that is oftentimes conceptualized as the functional opposite of attention, can be induced unconsciously. When participants were presented with masked nogo-signals or masked stop signals, they tended to respond slower than without such a signal or they even inhibited their response altogether (van Gaal et al., 2008, 2009; Hughes et al., 2009). These findings suggest that the link between consciousness and cognitive control may not be as obligatory as traditional views of consciousness and control propose. As the focusing of attention on relevant information is regarded as one of the most elementary executive functions (Smith and Jonides, 1999), insights into the role of cue awareness in this process are essential for an understanding of the functional role of consciousness and different aspects of cognitive control.

Interestingly, however, findings regarding the role of awareness and endogenously controlled shifts of attention are scarce. As noted, McCormick (1997) found that cue awareness is necessary to perform shifts of attention in direction opposite to that indicated by a peripheral cue. Note, however, that subjects in that study had to first countermand the impact of a peripheral cue before subsequently attending to a new location. It remains therefore an open question whether masked central cues would have the power to induce shifts of attention when such countermanding is not needed. In fact Reuss et al. (2011b) found preliminary evidence for the orienting of attention by masked central cues provided attention has not already been grabbed by another event. However, this finding is preliminary due to the specific type of cues used, namely arrows (for a similar study with eye gaze cues, see Al-Janabi and Finkbeiner, 2012). Arrow cues and other stimuli such as eye gaze and hand gestures carry an over learned spatial meaning. Most crucially, they were found to successfully capture attention even when they were not informative regarding the target location, which is in fact a hallmark of a reflexive rather than voluntary orienting of attention (Eimer, 1997; Hommel et al., 2001; Tipples, 2002; Friesen et al., 2004; Gibson and Bryant, 2005; Stevens et al., 2008; Pratt et al., 2010).

Given these limitations of previous research, the present study explored if symbolic cues that carry no inherent spatial meaning have the power to bias attention without cue awareness. To study this, we presented letters that indicated the locations of the target stimuli. These cue letters were presented masked or unmasked with the presence or absence of masks changing randomly from trial to trial. The primary question was if central cues were able to impact on attention at all when they are presented unconsciously.

The experimental protocol allowed us to study another debated question regarding the effects of masked stimuli, namely carry-over effects from one trial to the next trial. A well-known sequential effect is the so-called Gratton-effect, which deals with the influence of the congruency of the previous trial on the congruency effect in the current trial. Typically, congruency effects are smaller after trials with incongruent primes than after trials with congruent primes. Several studies found such carry-over effects when primes in the preceding trial were visible but not when they were invisible (Greenwald et al., 1996; Kunde, 2003; Frings and Wentura, 2008; Ansorge et al., 2011), though under certain circumstances even invisible primes might prompt such carry-over effects (van Gaal et al., 2010). Here, we will investigate if a cue is able to impact on the next trial depending on its visibility and the visibility of the next cue. To this end, we will analyze if the size of the validity effects is modulated by these two factors. There are reasons to expect this, though the type of impact is admittedly less clearly predictable. On the one hand, one may argue that strategies from the processing of visible cues are simply transferred to masked trials (cf. Klapp and Haas, 2005). Consequently, the impact of masked cues should increase the more recent (ideally in the last trial) a visible cue had been encountered. On the other hand, one may assume that attentional orienting by visible cues is much stronger than that by masked cues (which is in fact what we found). Perhaps masked cues have a better chance to impact on performance the less attention is still influenced by a preceding visible cue. To specifically investigate if spatial anticipations are still active in the next trial, we will compare validity effects when the cued location repeats in contrast to when the cued location switches. Persisting spatial anticipations would result in larger validity effects when the cued location repeats compared to when the cued location switches.

MATERIALS AND METHODS

EXPERIMENT 1

This experiment is based on the spatial cueing paradigm by Posner (1980), with centrally presented cues that indicate a lateral location, and a target display that either includes a target or consists solely of distractors. To make sure that any shifts of attention are truly endogenous, we used letters as cues which are normally not in any way pre-experimentally associated with a direction or location. The crucial manipulation was the visibility of the cues, which were presented either visibly or backward masked. Furthermore, to investigate the temporal dynamics of the cues' possible effects on attention, the cue target stimulus-onset-asynchrony (SOA) varied from 100 to 600 ms. As it has been shown that the magnitude of validity effects can depend strongly on the cue target SOA (e.g., Shulman et al., 1979; Posner, 1980), and masked priming effects are likewise susceptible to the interval between masked stimuli and target (e.g., Vorberg et al., 2003), this relatively broad range of SOAs was applied, especially regarding the novelty of this research. Finally, we analyzed sequential interactions of visible and masked cues, specifically whether masked cues are able to impact on attention in the context of a previously presented visible or masked cue.

Participants

Twenty-six students (five males) of the University of Würzburg with an average age of 22 years participated in the experiment in fulfillment of course requirements or payment (18 Euro). All reported having normal or corrected-to-normal vision, and were not familiar with the purpose of the experiment. The experiment was completed in three sessions that were run on separate days. Each session lasted approximately 1 h.

Apparatus and stimuli

The experiment took place in a dimly lit room. An IBM compatible computer with a 17″ VGA-Display and the software package E-Prime™ (Schneider et al., 2002) were used for stimulus presentation and response sampling. Stimulus presentation was synchronized with the vertical retraces of a 100-Hz monitor. Responses were executed with the index fingers of both hands and collected with external response keys. All stimuli were presented in white on a black background. The letter *V* or *H* functioned as central cues, presented in Arial font with a size of 30 pixels. Diamonds and squares with an edge length of 2 cm served as targets and distractors, respectively. In each target display, either one target and one distractor, or two distractors were presented on the left and on the right side, with a distance of 5 cm to the center. In trials with masked cues, the forward mask and backward mask consisted of a random string of four symbols (chosen from: #, &, $, and %), presented in Arial font with a size of 40 pixels.

Procedure and design

The sequence of events in a trial is depicted in **Figure 1**. Each trial started with a central fixation cross extending 0.7 cm × 0.7 cm that was presented for 600 ms. Following the fixation cross, a forward mask was presented for 70 ms. In trials with masked cues, the cue was presented for 30 ms, followed by a backward mask that was presented for 70 ms. In trials with visible cues, the cue was presented for 100 ms, and the backward mask was omitted. The target display appeared either immediately or after an interval of 100, 200, 300, 400, 500, or 600 ms and remained for 500 ms. Participants could respond within a time window of 2000 ms after target onset. After response execution a fixed time interval of 1000 ms elapsed before the next trial started.

Participants had to perform a single choice RT task. They were instructed to respond as fast as possible by pressing the spacebar when a target was present on either the left location or the right location, and not to respond when no target was present. Errors were indicated by the German word for wrong ("Falsch!") presented in red in the lower part of the monitor. RT were recorded from the onset of the target stimulus until a response was given.

Each block of 144 trials featured 24 catch trials in which no target was present and the participants were instructed not to respond. When a target was present, the cue indicated the location of the target correctly in 96 of these trials, i.e., with a validity of 80%. Considering all trials including catch trials, this results in an overall cue validity of 67%. During each block, each possible combination of the factors visibility of the cue, identity of the cue, location of the target, and cue target-interval was presented once in the case of an invalid trial (24 trials), and four times in the case of a valid trial (96 trials), with the sequence of trials being randomly determined. The experiment consisted of three sessions (two sessions for the main experiment, one session for assessment of cue visibility) that took approximately 1 h each. Participants performed one practice block and five experimental blocks in the first two sessions.

Participants were informed that a visible cue, the letter *V* or *H*, will be presented in 50% of the trials, and that the

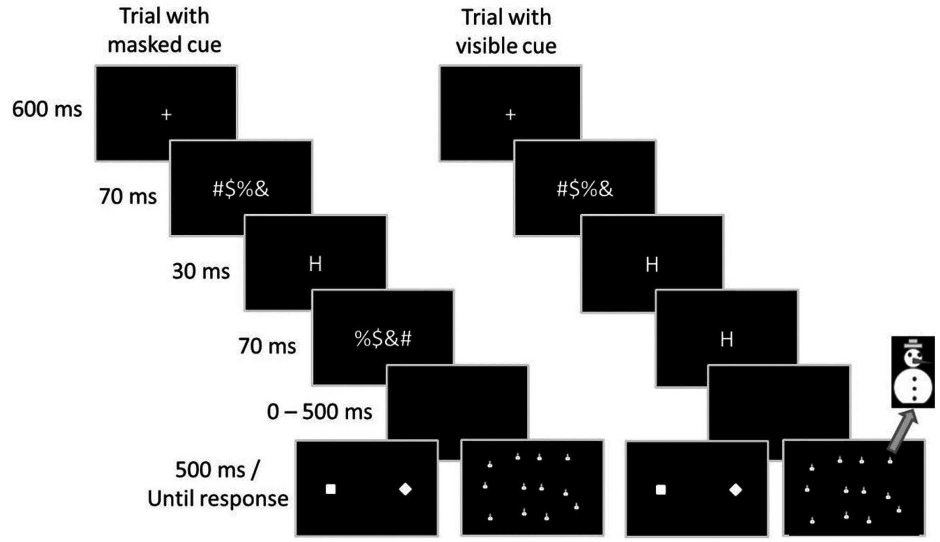

FIGURE 1 | Sequence of stimuli in Experiments 1 and 2. On the left side, a trial with short cue duration and a backward mask is depicted. On the right side, a trial with longer cue duration and no backward mask is depicted. The left target display shows the target display of Experiment 1, with target present. The right target display depicts the search display of Experiment 2.

cue predicted the correct location of the target in most trials. They were told not to move their eyes away from fixation when they shifted their attention. Eye movements were, however, not measured during the experiment, so that we cannot exclude the possibility of eye movements. Participants were not informed about the masked cues. The mapping of each cue to the left or to the right location was counterbalanced across participants.

Assessment of cue visibility

A visibility test consisting of 10 blocks of 72 trials each featuring both non-masked cues and masked cues constituted the third session of the experiment. Participants were fully informed about the structure of a trial and the presence of masked cues. They had to perform a forced choice discrimination task. For this task, the sequence of stimuli was exactly the same as in the main experiment. However, there was no time limit after target onset, and the overall cue validity was lowered to 50%, so that the participants could not infer from the location of the target which cue was more likely. Participants were asked to discriminate whether a *V* or an *H* was presented, and had to press one of two response keys accordingly. Participants were instructed to take their time, to try to be as accurate as possible, and if they had not seen anything to guess, bearing in mind the probability for either cue was equal.

RESULTS

EXPERIMENT 1

Trials with RTs deviating more than 2.5 standard deviations (SDs) from the mean RT of each participant and each condition were excluded (1.3% of all trials). Mean RTs for correct responses were submitted to a repeated measures analysis of variance (ANOVA) with the within-subject factors cue visibility (visible cue vs. masked cue), validity (valid cue vs. invalid cue), cue target SOA (100, 200, 300, 400, 500, and 600 ms), and previous cue visibility (visible cue vs. masked cue in trial n-1). The results are depicted in **Figure 2**.

We found significant main effects for the factors cue validity, $F(1, 25) = 12.5$, $p = 0.002$), cue visibility, $F(1, 25) = 6.05$, $p = 0.021$, and cue target SOA, $F(1, 25) = 122.7$, $p < 0.001$. These main effects indicate faster responses after valid cues than after invalid cues (384 vs. 399 ms), faster responses after visible cues than after masked cues (388 vs. 395 ms), and faster responses with longer SOAs. The interaction of cue validity and cue visibility was significant, $F(1, 25) = 5.51$, $p = 0.027$, as well as the interaction of cue validity and cue target SOA, $F(1, 25) = 3.58$, $p = 0.005$. Additionally, the three-way interaction of cue validity, cue visibility, and previous cue visibility reached significance, $F(1, 25) = 5.11$, $p = 0.033$. To further analyze these interactions, we conducted two separate ANOVAs for trials with visible cues and trials with masked cues.

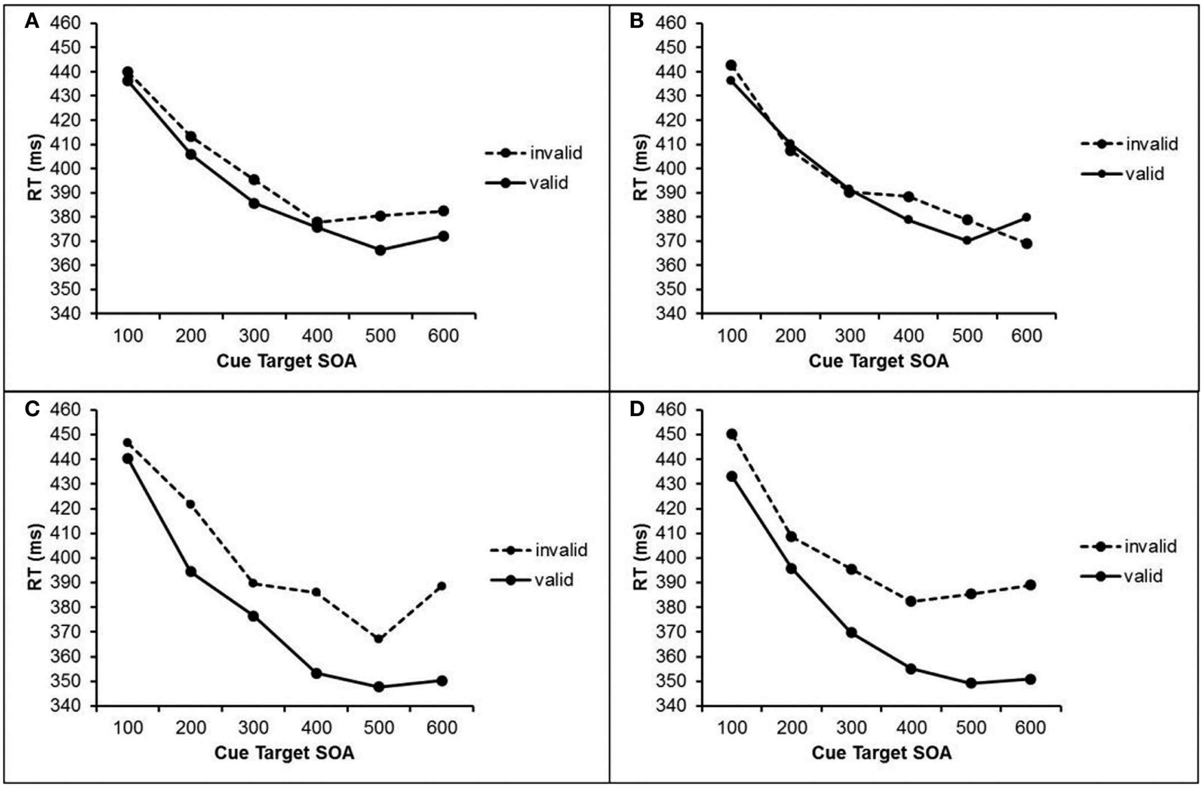

FIGURE 2 | RTs in Experiment 1 after masked cues (upper half) and visible cues (lower half) as a function of cue validity, cue target SOA, and visibility of the previous cue. (A) RTs after masked cue when the previous cue was masked. **(B)** RTs after masked cues when the previous cue was visible. **(C)** RTs after visible cues when the previous cue was masked. **(D)** RTs after visible cues when the previous cue was visible.

With visible cues, we found a 25 ms effect of cue validity, $F(1, 25) = 9.28$, $p = 0.005$, as well as an interaction of cue validity and cue target SOA, $F(5, 125) = 5.05$, $p < 0.001$, which reflects larger validity effects with longer SOAs. There was no interaction of cue validity and previous visibility, $F < 1$.

With masked cues, there was also a significant effect of cue validity, $F(1, 25) = 4.61$, $p = 0.042$. Participants responded 5 ms faster after valid cues than after invalid cues. Additionally, we found a marginally significant interaction of cue validity and previous cue visibility, $F(1, 25) = 3.13$, $p = 0.089$. Single comparisons revealed that validity effects of masked cues were present only after trials with masked cues, $t(25) = 2.43$, $p = 0.023$. Here, participants responded 8 ms faster after valid cue than after invalid cues. After trials with visible cues, however, no such validity effect was found $t(25) = 0.71$, $p = 0.472$.

To shed light on possible underlying mechanisms of this observation, we analyzed RTs regarding cue validity (valid vs. invalid) and cued location repetition (repetition or change compared to previous trial) separately for visible and masked cues. For masked cues that follow a visible cue, we found an interaction of cue validity and previously cued location $F(1, 25) = 10.12$, $p = 0.004$. When the cued location was the same as in the previous trial, a regular validity effect of 9 ms was present. When the cued locations changed, however, the validity effect was reversed, with RTs of 392 ms after invalid cues and RTs of 397 ms after valid cues. In other words, in both cases, RTs were shorter at the previously cued location compared to the previously non-cued location. In masked trials after masked cues, a similar interaction was present, with a larger cuing effect when the cued location remained the same (9 ms) than when it changed (3 ms), but this interaction was not significant, $F(1, 25) = 1.47$, $p = 0.24$.

For visible cues that follow a visible cue, this interaction was also significant, $F(1, 25) = 8.61$, $p = 0.007$, with a larger validity effect (33 ms) when the cued location remained the same than when it changed (18 ms). For visible cues that follow a masked cue the cuing effect was also a larger when the cued location remained the same (26 ms) than when it changed (20 ms), but this interaction missed significance, $F(1, 25) = 1.22$, $p = 0.279$.

Cue visibility was assessed by computing the signal detection measure d', treating the cue V as signal, and the cue H as noise. Participants' discrimination performance for the masked cues was $d' = 0.54$, with a mean hit rate of 55.7% and a mean false alarm rate of 37.8%. This value deviated from zero $t(24) = 4.64$, $p < 0.001$. To test whether any validity effects with masked cues can be ascribed to cue visibility, we assessed the relationship between each participants' individual d' score and the effect of valid and invalid masked cues on RT. We adopted a procedure suggested by Greenwald et al. (1995), see also Greenwald et al. (1996), Draine and Greenwald (1998) and regressed the validity effect of each participant onto individual d' scores. This analysis showed that d' scores and the effects of masked cues are not significantly correlated ($r = 0.283$, $p = 0.16$), which implies that while it cannot be definitely ruled out that some masked cues were consciously perceived, the observed effects are mostly independent of individual cue visibility and are by and large not due to conscious perception of some of the cues.

DISCUSSION

EXPERIMENT 1

In Experiment 1, we investigated whether centrally presented cues lead to spatial anticipations and accordant shifts of attention, and how this effect depends on the visibility of the cues, the cue target SOA, and the visibility of the previous cue. The results show that participants did form spatial anticipations based on the cues' information and shifted their attention accordingly. With visible cues, the cueing effect increased with SOA. As the interpretation of the cue and voluntarily shifting of attention takes effort and time, the benefits of correct anticipations are more pronounced when they happen before target onset.

Remarkably, masked cues also lead to the formation of spatial anticipations and accordant shifts of attention. These anticipations based on masked cues were, however, found to be more susceptible to external modulation. Specifically, masked cues were only able to impact on attention when there were no current spatial anticipations that were induced by visible information in the previous trial, i.e., masked cues were effective only when the previously presented cue was also masked, but not when it was visible.

Previously formed spatial anticipations generally had an impact on orienting of attention. Responding was faster when the target appeared at the previously cued location than when it did not. As noted above this impact of the previously cued location was stronger when the cue in the preceding trial was visible than when it mas masked. In fact, when the previous cue was visible and the current cue was masked, the impact of the previous cuing seemed strong enough to override the cuing effect by a masked cue. We found that the validity effect was actually reversed when the previous cue indicated another location than the current masked cue. In this case, responses were faster with invalid cues, as this location was the one that was previously cued, and slower with valid cues, as the target then appeared at the previously non-cued location. In other words, orienting of attention was influenced stronger by the spatial anticipation formed in the previous trial than by the current masked cue. It is not entirely clear, however, whether the persisting spatial anticipation is in fact solely due to the previous cue's information, or whether it is also influenced by the previous target location. Possibly, participants tended to orient their attention toward the previous target location. Effects of the previous target location might be disentangled from effects of the previous cued location by additionally analyzing the cue validity in the previous trial. Unfortunately, the experimental design at hand does not allow for a statistically sound analysis with this additional factor, as particular factor combinations yield too few cases for each participants to perform a meaningful analysis. Thus, the distinct role of the previous target location in the sequential modulation cannot be clarified with the data at hand.

The fragility of the masked cueing effect could explain previously unsuccessful efforts to find this effect. In McCormick's (1997) study, the exogenous shift of attention triggered by the peripherally presented cue might have suppressed an effect of spatial anticipation (which would be directed on the opposite side of the screen), as an anticipation that is based on a masked cue can be influenced and possibly suppressed by other spatial information currently present.

MATERIAL AND METHODS

EXPERIMENT 2

Experiment 1 showed that anticipative shifts of attention can be triggered by masked symbolic cues. To replicate and further elaborate this finding, a more demanding visual search task with a two forced choice RT task was implemented in Experiment 2 instead of the elementary target detection task of Experiment 1. To this end, Experiment 2 featured a visual search display with 11 distractors and 1 target. Here, participants had to search for one of two possible targets amongst several stimuli. The difficulty to find the target and to identify it was therefore far higher than in Experiment 1, and the benefits and costs after valid and invalid cues, respectively, were thus supposedly much larger. Like in Experiment 1, we varied the visibility of the cue and the cue target SOA. Also, we again analyze sequential effects of cue visibility.

Participants

Twenty-one students (13 males) of the University of Würzburg with an average age of 24 years participated in the experiment in fulfillment of course requirements or for payment (18 Euro). Informed consent was obtained from all participants. All participants reported having normal or corrected-to-normal vision, and were not familiar with the purpose of the experiment. The experiment was completed in three sessions that were run on separate days. Each session lasted approximately 1 h.

Apparatus and stimuli

The experiment took place in a dimly lit room. An IBM compatible computer with a 17″ VGA-Display and the software package E-Prime™ were used for stimulus presentation and response sampling. Stimulus presentation was synchronized with the vertical retraces of a 100-Hz monitor. Responses were executed with the index fingers of both hands and collected with external response keys. All stimuli were presented on a black background. Cue stimuli and masking stimuli were identical to those in Experiment 1. The target display consisted of 12 snowmen (extending 8 mm × 18 mm) wearing colored hats (extending 7 mm × 3 mm; see lower right screen of **Figure 1**). The snowmen were quasi-randomly distributed over the screen. For this, the screen was subdivided in 4 × 3 grids (invisible to the participants), and in each grid a snowman was presented at a random location, so that six snowmen were presented on locations on the left half of the screen, and six snowmen were presented on locations on the right side of the screen. There was always exactly one target snowman present, which was denoted by wearing either a blue hat or a gray hat. The other 11 distractor snowmen wore red, violet, orange, yellow, and green hats.

Procedure and design

The sequence of events in a trial until the target display was identical to that of Experiment 1 (see **Figure 1**). The target search display was presented until a response was given, with no limitation by a response window. After response execution a fixed time interval of 1000 ms elapsed before the next trial started.

Participants had to perform a two forced choice RT task. They were instructed to respond as fast as possible to the color of the target snowman's hat. The mapping of left and right responses to gray vs. blue hats was counterbalanced across participants. Errors were indicated by the German word for wrong ("Falsch!") presented in red in the lower part of the monitor. RT were recorded from the onset of the target stimulus until a response was given.

Participants were informed that a visible cue is presented in 50% of the trials, and that the cue indicates the correct side of the screen (left vs. right) where the target appears in most trials. They were not informed about the masked cues. The mapping of each cue to the left or to the right side of the screen was counterbalanced across participants.

Each of two experimental sessions consisted of one short training block (20 trials) and nine experimental blocks of 80 trials each. All 96 possible combinations of cues, visibility of the cue, target location (left or right), cue target SOA, and target identity (either blue or gray hat) were presented within the span of three blocks. All combinations with invalid cues were presented once within three blocks, and all combinations with valid cues were presented four times within three blocks. The overall cue validity was thus 80%. The experiment consisted of three sessions that took approximately 1 h each.

Assessment of cue visibility

A visibility test consisting of six blocks of 96 trials featuring both non-masked cues and masked cues was applied in the third experimental session. Participants were fully informed about the structure of a trial and the presence of masked cues. They had to perform a forced choice discrimination task. For this task, the sequence of stimuli was exactly the same as in the main experiment. However, the overall cue validity was lowered to 50%, so that the participants could not infer from the location of the target which cue was more likely. Participants were asked to discriminate whether a *V* or an *H* was presented, and had to press one of two response keys accordingly. Participants were instructed to take their time, to try to be as accurate as possible, and if they had not seen anything to guess, bearing in mind the probability for either cue was equal.

RESULTS

EXPERIMENT 2

Trials with RTs deviating more than 2.5 SDs from the mean RT of each participant and each condition were excluded (1.7% of all trials). RT data were submitted to a repeated measures ANOVA with the within-subject factors cue visibility (visible cue vs. masked cue), validity (valid cue vs. invalid cue), cue target SOA (100–600 ms), and previous cue visibility (visible cue vs. masked cue in trial n-1). The results are depicted in **Figure 3**.

This analysis revealed a main effect of validity, $F(1, 20) = 24.2$, $p < 0.001$. Participants responded faster after valid cues (1185 ms) than after invalid cues (1327 ms) The main effect of cue visibility was also significant, $F(1, 23) = 15.8$, $p = 0.001$. Participants responded faster (1221 ms) after masked cues than after visible cues (1291 ms). The interaction of cue visibility and validity was significant, $F(1, 20) = 17.0$, $p = 0.001$, as was the three-way interaction of cue visibility, validity, and previous cue visibility, $F(1, 20) = 13.0$, $p = 0.002$. No other main effects or interactions were significant ($ps > 0.166$). To further investigate these interactions, we conducted two separate ANOVAs for visible and masked cues.

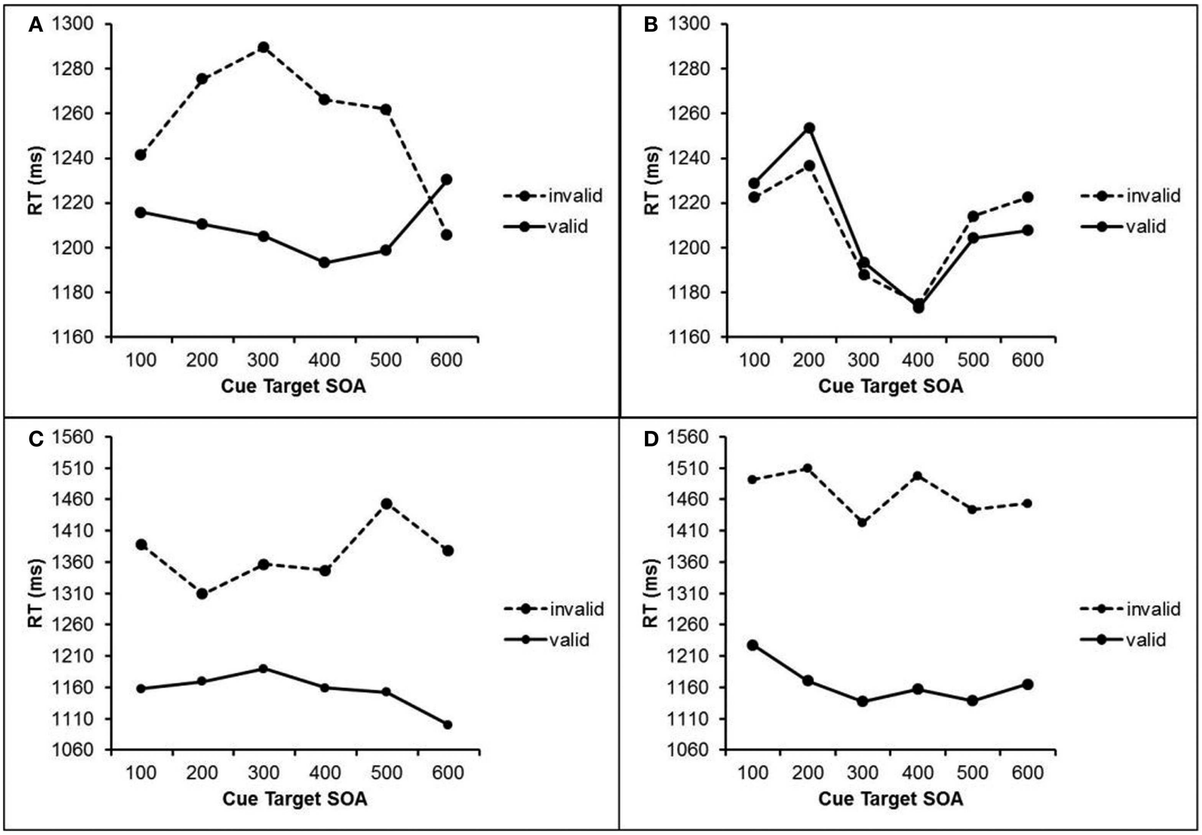

FIGURE 3 | RTs in Experiment 2 after masked cues (upper half) and visible cues (lower half) as a function of cue validity, cue target SOA, and visibility of the previous cue. (A) RTs after masked cue when the previous cue was masked. **(B)** RTs after masked cues when the previous cue was visible. **(C)** RTs after visible cues when the previous cue was masked. **(D)** RTs after visible cues when the previous cue was visible.

With visible cues, a main effect of validity revealed faster responses after valid (1164 ms) than after invalid (1440 ms) cues, $F(1, 20) = 21.2$, $p < 0.001$. The interaction of validity and previous cue visibility was significant, $F(1, 20) = 11.3$, $p = 0.003$. Here, validity effects were larger after trials with visible cues (303 ms) than after trials with masked cues (217 ms).

With masked cues, we also found a significant main effect of validity, $F(1, 20) = 4.84$, $p = 0.04$, with responses that were 24 ms faster after valid than after invalid masked cues. The interaction of validity and previous visibility just failed to reached marginal significance, $F(1, 20) = 2.84$, $p = 0.11$. In contrast to visible cues, masked cues only impacted on attention when the previous cue was also masked, $t(20) = 2.38$, $p = 0.027$ which is reflected in a validity effect of 48 ms. When following a visible cue, masked cues were not able to impact on attention at all, evident by an absent validity effect (0 ms).

To further understand these sequential effects we analyzed RTs regarding cue validity (valid vs. invalid) and cued location repetition (repetition or change compared to previous trial) separately for visible and masked cues. For masked cues that follow a visible cue, we found an interaction of cue validity and repetition of cued location, $F(1, 20) = 7.35$, $p = 0.013$. Similar to Experiment 1, there was a regular cuing effect of 72 ms when the cued locations repeated, which was reversed to a negative cuing effect (−58 ms) when the cued location changed. For masked cues that follow a masked cue, the interaction was marginally significant, $F(1, 20) = 3.27$, $p = 0.086$, and also reflected a regular cuing effect when the cued location repeated (77 ms) and a reversed cueing effect when the location switched (−3 ms).

For visible cues that follow a masked cue, the interaction was also significant, $F(1, 20) = 14.48$, $p = 0.001$, with a larger cuing effect (278 ms) when the cued location repeated than when it changed (150 ms). However, no significant interaction was found for visible cues that follow a visible cue, $F < 1$.

Cue visibility was assessed by computing the signal detection measure d', treating the cue V as signal and the cue H as noise. Participants' discrimination performance for the masked cues was $d' = 0.186$, with a mean hit rate of 54.5% and a mean false alarm rate of 47.4%. This value deviated from zero $t(20) = 2.48$, $p = 0.023$. To test whether any validity effects of masked cues can be ascribed to cue visibility, we assessed the relationship between each participants' individual d' score and the effect of valid and invalid masked cues on RT. Following a procedure suggested by Greenwald et al. (1995), see also Draine and Greenwald (1998), Greenwald et al. (1996) and regressed the validity effect of each participant (RT invalid trials – RT valid trials) onto individual d'

scores. This analysis showed that d' scores and the effects of masked cues were not significantly correlated ($r = 0.126, p = 0.596$), which implies that the observed effects are mostly independent of individual cue visibility and are by and large not due to conscious perception of some of the cues.

DISCUSSION

EXPERIMENT 2

The results of Experiments 2 confirmed the findings of Experiment 1 in a visual search context: participants are able to form spatial anticipations and shift their attention accordingly on the basis of both visible and masked centrally presented cues. Participants shifted their attention to the side where they anticipated the target, which resulted in shorter RTs when the target was in fact amongst the stimuli on this side of the screen, and in longer RTs when the target was actually on the other side of the screen. With visible cues, this resulted in responses that were 276 ms faster after valid than after invalid cues. With masked cues, this effect was much smaller (24 ms) but still present. This shows that even cues that we are not aware of are able to induce spatial anticipations that lead to according shifts of attention. However, an effect of masked cues was found only when the previous trial did not contain a visible cue. This observation suggests that information provided by masked stimuli takes effect only when no stronger spatial information, i.e., that of visible stimuli, is in a still active state. The cuing effect depended on the previously cued location. It was stronger when the cued locations repeated from previous to current trials than when they switched. As in Experiment 1, the impact of a previous visible cue was strong enough to invert the regular cuing effect from a current masked cue. Yet, even previous masked cues were able to modify cuing effects in the current trial to some degree. Within this regard, it again remains unclear whether the location of the previous target additionally influenced orienting of attention in the current trial. One exception from this overall pattern, which otherwise emerged quite consistently in both experiments, was the lack of sequence effects with two subsequent visible cues in Experiment 2. At present we have no obvious explanation for this.

The cue target SOA had less of an influence than in Experiment 1, probably because of the different time frame of the tasks. Conceivably, the information provided by the cue was not effectively used with very short cue target SOAs in Experiment 1 due to RTs that were shorter than the time needed to interpret the cue and shift one's attention. When the target display appeared shortly after the cue and probably before the shift of attention was initiated, the simple task was carried out before the accordant shift of attention was performed. With the visual search task in Experiment 2, the target display could appear before the shift of attention was initiated, but the information provided by the cue could still be effectively used because of the rather long search RTs to find the target.

GENERAL DISCUSSION

The ability to shift our attention in anticipation of future events is an elementary process of cognitive control. Here, we provided evidence that such shifts of attention can be elicited by masked cues.

In two experiments, centrally presented letter cues informed the participants about the likely location of the upcoming target. Participants responded faster when the target appeared at the anticipated location (i.e., after a valid cue) rather than at another location (i.e., after an invalid cue). This indicated that participants formed spatial anticipations regarding the location of the target and shifted their attention accordingly. Strikingly, this was true for visible as well as for masked cues. This is especially remarkable as the cues were deliberately chosen to be spatially arbitrary. Unlike arrows, letters possess no inherent spatial meaning. Thus, letters have to be interpreted regarding their spatial meaning to form spatial anticipations. The observed effects of the cues therefore cannot be attributed to automatically induced shifts of attention that are based on over learned spatial relations like in previous studies (Reuss et al., 2011b; Al-Janabi and Finkbeiner, 2012), but must be attributed to anticipatory shifts of attention generated endogenously.

In Experiment 1, participants had to recognize whether the target display contained one distractor and one target, or two distractors. With this single choice task, RT were very short. Consequently, the cue target SOA modulated the validity effect. Only with longer SOAs, the spatial information provided by the cue could be used effectively to orient attention before the target occurred and the response was given. This interaction was more pronounced in the visible cue condition. Most importantly, however, the validity effects were found both for visible and masked cues. The latter, however, were only able to impact on attention when the previous cue was also masked. This indicates that visible cues lead to strong spatial anticipations that are able to persist at least until the next trial and interfere with forming new spatial anticipations, especially those based on masked cues.

In Experiment 2, participants had to actually search for the target among eleven distractors. Thus, target detection was harder and RT were longer than in Experiment 1. The increased task difficulty worked as an incentive to use the cues, as the information provided by the cues is potentially more beneficial the harder the target is to detect, which resulted in large effects of cue validity. Also, the influence of the cue target SOA was reduced in Experiment 2 compared to Experiment 1, so that effects of cue validity were also present with very short SOAs. Besides that, the overall pattern of results was very similar to Experiment 1. Again, validity effects were found both for masked cues and visible cues, and the effects of masked cues strongly depended on the visibility of the previous cue: when the previous cue was visible, no effects of a masked cue could be observed at all. Masked cues were effective only after trials with masked cues.

To further investigate the underlying mechanism of these sequential modulations, we analyzed whether the cue information of the previous trial interacts with the current cue information depending on their respective visibility. These analyses revealed for both experiments that when the current cue was masked, participants oriented their attention toward the location that was previously cued if this previous cue was visible (and to a lesser extent also when the previous cue was masked). When the same location as in the previous trial was cued, participants responded faster when the target appeared at the cued than at the non-cued location. However, when the currently cued location differed from the cued location in the previous trial, participants responded actually faster when the target appeared at the currently non-cued location than at the currently cued location. In other words, responses were faster when the target appeared at the location indicated by the previous visible cue compared to when the target appeared at the location that was not indicated by the previous visible cue, whereas the current masked cue had no substantial impact. This strongly suggests that spatial anticipations persist until the next trial and still influence the orienting of attention to an extent that nullifies effects of masked cues (in the case of a previous visible cue) or at least modulates their effect (in the case of a previous masked cue). It is also plausible, however, that not only the information of the previous cue, but also the actual target location in the previous trial influenced the spatial anticipation that carried over to the next trial. As the cued location is identical to the target location in the majority of trials, the observed effects can be due to either of these factors. While we analyzed whether the visibility of the previous cue influenced the observed validity effects, previous cue information is confounded with previous target location because cues were valid in most trials. Unfortunately, the present data set does not allow us to soundly disentangle the effects of both the cue information and the target location on orienting of attention in the subsequent trial, as particular factor combinations in the necessary analysis occur too infrequently to enable a meaningful analysis.

The persisting effect of already active spatial anticipations is strong enough to still impact on attention even when a visible cue is presented. However, spatial anticipations that are induced by visible cues are more resilient to such influences, so that their effect is merely modulated by already active spatial anticipations, but not completely overridden. Interestingly, this modulation took place even when the previous cue was masked.

The observation that even masked cues lead to anticipatory shifts of attention is remarkable because endogenous orienting of attention is regarded as one of the most elementary processes of cognitive control, and cognitive control processes are traditionally associated with consciousness (e.g., Atkinson and Shiffrin, 1968; Norman and Shallice, 1986), and "authors speak of "conscious control" as if there could be no alternative" (Hommel, 2007, p. 161). An effect of masked cues thus casts doubt on this proposed correlation. As outrageous as this devaluation of the functionality of our consciousness may seem, given its antagonism to our introspective impression that our conscious will controls our actions (Wegner, 2002), it is in line with recent findings concerning the relation of consciousness and other cognitive control processes like inhibitory processes (van Gaal et al., 2008, 2009; Hughes et al., 2009) and task set activation (Mattler, 2006; Lau and Passingham, 2007; Reuss et al., 2011a). These cognitive control processes were shown to be able to work outside of awareness.

However, one should not dismiss the differences that exist between the effects of visible cues and masked cues. First, there are quantitative differences when looking at the benefits and costs of valid cues and invalid cues depending on their visibility. The effects of visible cues are distinctively larger than those of masked cues. This indicates a stronger and more reliable impact on cognitive control processes than the one provided by masked cues. Such a quantitative difference was, for example, also found by Reuss et al. (2011a) regarding the activation of task sets by masked cues. Second, the effects of visible cues are less prone to potential interference than the effects of masked cues. In the two experiments presented here, this is illustrated by the impact of cues on attention in the next trial. In both experiments, spatial anticipations induced by visible cues were still active in the subsequent trial. In trials with masked cues, this persisting spatial anticipation was able to strongly influence the orienting of attention, sometimes to an extent that the current cue had no noticeable effect on attention. Visible cues were in contrast more robust against such a persisting influence. Persisting spatial anticipations were able to impact on attention in trials with visible cues as well, but the effect of visible cues was strong enough to also significantly impact on attention.

To conclude, we showed in two experiments that spatial anticipations and corresponding shifts of attention are able to be induced both by visible cues and by masked cues. This observation challenges the notion of a strong link between orienting of attention as a prototypical control process and consciousness. However, awareness of the cue still played a role regarding the reliability and robustness of the control process.

ACKNOWLEDGMENT

This research was supported by a grant from the German Science Foundation (DFG, Grant Ku 1964/7-1).

REFERENCES

Al-Janabi, S., and Finkbeiner, M. (2012). Effective processing of masked eye gaze requires volitional control. *Exp. Brain Res.* 216, 433–443.

Ansorge, U., Fuchs, I., Khalid, S., and Kunde, W. (2011). No conflict control in the absence of awareness. *Psychol. Res.* 351–365.

Ansorge, U., and Heumann, M. (2006). Shifts of visuospatial attention to invisible (metacontrast-masked) singletons: clues from reaction times and event-related potentials. *Adv. Cogn. Psychol.* 2, 61–76.

Ansorge, U., and Neumann, O. (2005). Intentions determine the effect of invisible metacontrast-masked primes: evidence for top-down contingencies in a peripheral cuing task. *J. Exp. Psychol. Hum. Percept. Perform.* 31, 762–777.

Atkinson, R. C., and Shiffrin, R. M. (1968). "Human memory: a proposed system and its control processes," in *The Psychology of Learning and Motivation*, Vol. 2, eds K. W. Spence and J. T. Spence (New York: Academic Press), 89–195.

Draine, S. C., and Greenwald, A. G. (1998). Replicable unconscious semantic priming. *J. Exp. Psychol. Gen.* 127, 286–303.

Eimer, M. (1997). Uninformative symbolic cues may bias visual-spatial attention: behavioral and electrophysiological evidence. *Biol. Psychol.* 46, 67–71.

Folk, C. L., Remington, R. W., and Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *J. Exp. Psychol. Hum. Percept. Perform.* 18, 1030–1044.

Friesen, C. K., Ristic, J., and Kingstone, A. (2004). Attentional effects of counter predictive gaze and arrow cues. *J. Exp. Psychol. Hum. Percept. Perform.* 30, 319–329.

Frings, C., and Wentura, D. (2008). Trial-by-trial effects in the affective priming paradigm. *Acta Psychol. (Amst.)* 128, 318–323.

Gibson, B. S., and Bryant, T. A. (2005). Variation in cue duration reveals top-down modulation of involuntary orienting to uninformative symbolic cues. *Percept. Psychophys.* 67, 749–758.

Greenwald, A. G., Draine, S. C., and Abrams, R. L. (1996). Three cognitive markers of unconscious semantic activation. *Science* 273, 1699–1702.

Greenwald, A. G., Klinger, M. R., and Schuh, E. S. (1995). Activation by marginally perceptible ("subliminal") stimuli: dissociation of unconscious from conscious cognition. *J. Exp. Psychol. Gen.* 124, 22–42.

Hommel, B. (2007). Consciousness and control: not identical twins. *J. Conscious. Stud.* 14, 155–176.

Hommel, B., Pratt, J., Colzato, L., and Godijn, R. (2001). Symbolic control of visual attention. *Psychol. Sci.* 12, 360–365.

Hughes, G., Velmans, M., and De Fockert, J. D. (2009). Unconscious priming of a no-go response. *Psychophysiology* 46, 1258–1269.

Ivanoff, J., and Klein, R. M. (2003). Orienting of attention without awareness is affected by measurement-induced attentional control settings. *J. Vis.* 3, 32–40.

Jonides, J. (1981). "Voluntary versus automatic control over the mind's eye's movement," in *Attention & performance IX*, eds J. B. Longand and A. D. Baddeley (Hillsdale, NJ: Erlbaum), 187–203.

Klapp, S. T., and Haas, B. W. (2005). Nonconscious influence of masked stimuli on response selection is limited to concrete stimulus-response associations. *J. Exp. Psychol. Hum. Percept. Perform.* 31, 193–209.

Kunde, W. (2003). Sequential modulations of stimulus-response correspondence effects depend on awareness of response conflict. *Psychon. Bull. Rev.* 10, 198–205.

Lambert, A., Naikar, N., McLachlan, K., and Aitken, V. (1999). A new component of visual orienting: Implicit effects of peripheral information and subthreshold cues on covert attention. *J. Exp. Psychol. Hum. Percept. Perform.* 25, 321–340.

Lau, H. C., and Passingham, R. E. (2007). Unconscious activation of the cognitive control system in the human prefrontal cortex. *J. Neurosci.* 27, 5805–5811.

Mattler, U. (2006). On the locus of priming and inverse priming effects. *Percept. Psychophys.* 68, 975–991.

McCormick, P. A. (1997). Orienting attention without awareness. *J. Exp. Psychol. Hum. Percept. Perform.* 23, 168–180.

Mulckhuyse, M., Talsma, D., and Theeuwes, J. (2007). Grabbing attention without knowing: automatic capture of attention by subliminal spatial cues. *Vis. Cogn.* 15, 779–788.

Mulckhuyse, M., and Theeuwes, J. (2010). Unconscious attentional orienting to exogenous cues: a review of the literature. *Acta Psychol. (Amst.)* 134, 299–309.

Müller, H. J., and Rabbitt, P. M. A. (1989). Reflexive and voluntary orienting of visual attention: time course of activation and resistance to interruption. *J. Exp. Psychol. Hum. Percept. Perform.* 15, 315–330.

Norman, D. A., and Shallice, T. (1986). "Attention to action: willed and automatic control of behavior," in *Consciousness and Self-Regulation*, Vol. 4, eds R. J. Davidson, G. E. Schwartz, and D. Shapiro (New York: Plenum Press), 1–18.

Posner, M. I. (1980). Orienting of attention. *Q. J. Exp. Psychol.* 32, 3–25.

Posner, M. I., and Cohen, Y. (1984). "Components of visual orienting," in *Attention and Performance X*, eds H. Bouma and D. G. Bowhui (Hillsdale, NJ: Erlbaum), 531–556.

Posner, M. I., and Snyder, C. R. R. (1975). "Attention and cognitive control," in *Information Processing and Cognition: The Loyola Symposium*, ed. R. L. Solso (Hillsdale, NJ: Erlbaum), 55–85.

Posner, M. I., Snyder, C. R. R., and Davidson, B. J. (1980). Attention and the detection of signals. *J. Exp. Psychol. Gen.* 109, 160–174.

Pratt, J., Radulescu, P., Guo, R. M., and Hommel, B. (2010). Visuospatial attention is guided by both the symbolic value and the spatial proximity of selected arrows. *J. Exp. Psychol. Hum. Percept. Perform.* 36, 1321–1324.

Reuss, H., Kiesel, A., Kunde, W., and Hommel, B. (2011a). Unconscious activation of task sets. *Conscious. Cogn.* 20, 556–567.

Reuss, H., Pohl, C., Kiesel, A., and Kunde, W. (2011b). Follow the sign! Top-down contingent attentional capture of masked arrow cues. *Adv. Cogn. Psychol.* 7, 82–91.

Scharlau, I. (2002). Leading, but not trailing, primes influence temporal order perception: further evidence for an attentional account of perceptual latency priming. *Percept. Psychophys.* 64, 1346–1360.

Scharlau, I., and Ansorge, U. (2003). Direct parameter specification of an attention shift: evidence from perceptual latency priming. *Vision Res.* 43, 1351–1363.

Scharlau, I., and Neumann, O. (2003). Perceptual latency priming by masked and unmasked stimuli: evidence for an attentional interpretation. *Psychol. Res.* 67, 184–196.

Schneider, W., Eschman, A., and Zuccolotto, A. (2002). *E-Prime Users' Guide*. Pittsburgh: Psychology Software Tools Inc.

Shulman, G. L., Remington, R. W., and McLean, J. P. (1979). Moving attention through visual space. *J. Exp. Psychol. Hum. Percept. Perform.* 5, 522–526.

Smith, E. E., and Jonides, J. (1999). Storage and executive processes in the frontal lobes. *Science* 283, 1657–1661.

Stevens, S. A., West, G. L., Al-Aidroos, N., Weger, U. W., and Pratt, J. (2008). Testing whether gaze cues and arrow cues produce reflexive or volitional shifts of attention. *Psychon. Bull. Rev.* 15, 1148–1153.

Theeuwes, J. (1991). Exogenous and endogenous control of attention: the effect of visual onsets and offsets. *Percept. Psychophys.* 49, 83–90.

Tipples, J. (2002). Eye gaze is not unique: automatic orienting in response to uninformative arrows. *Psychon. Bull. Rev.* 9, 314–318.

van Gaal, S., Lamme, V. A. F., and Ridderinkhof, K. R. (2010). Unconsciously triggered conflict adaptation. *PLoS ONE* 5, 11508. doi:10.1371/journal.pone.0011508

van Gaal, S., Ridderinkhof, K. R., Fahrenfort, J. J., Scholte, H. S., and Lamme, V. A. F. (2008). Frontal cortex mediates unconsciously triggered inhibitory control. *J. Neurosci.* 28, 8053–8062.

van Gaal, S., Ridderinkhof, K. R., van den Wildenberg, W. P. M., and Lamme, V. A. F. (2009). Dissociating consciousness from inhibitory control: evidence for unconsciously triggered response inhibition in the stop-signal task. *J. Exp. Psychol. Hum. Percept. Perform.* 35, 1129–1139.

von Helmholtz, H. (1866). *Treatise on Physiological Optics*. Washington: The Optical Society of America.

Vorberg, D., Mattler, U., Heinecke, A., Schmidt, T., and Schwarzbach, J. (2003). Different time courses for visual perception and action priming. *Proc. Natl. Acad. Sci. U.S.A.* 100, 6275–6280.

Wegner, D. M. (2002). *The Illusion of Conscious Will*. Cambridge, MA: MIT Press.

Yantis, S. (1998). "Control of visual attention," in *Attention*, ed. H. Pashler (Hove: Psychology Press), 223–256.

Yantis, S., and Johnson, D. N. (1990). Mechanisms of attentional priority. *J. Exp. Psychol. Hum. Percept. Perform.* 16, 812–825.

Yantis, S., and Jonides, J. (1990). Abrupt visual onsets and selective attention: voluntary versus automatic allocation. *J. Exp. Psychol. Hum. Percept. Perform.* 16, 121–134.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 30 June 2012; accepted: 22 September 2012; published online: 12 October 2012.

*Citation: Reuss H, Kiesel A, Kunde W and Wühr P (2012) A cue from the unconscious – masked symbols prompt spatial anticipation. Front. Psychology **3**:397. doi: 10.3389/fpsyg.2012.00397*

This article was submitted to Frontiers in Cognition, a specialty of Frontiers in Psychology.

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 08 August 2012
doi: 10.3389/fpsyg.2012.00282

When predictions take control: the effect of task predictions on task switching performance

Wout Duthoo[1]*, Wouter De Baene[1], Peter Wühr[2] and Wim Notebaert[1]*

[1] Department of Experimental Psychology, Ghent University, Ghent, Belgium
[2] Institut für Psychologie, Technische Universität Dortmund, Dortmund, Germany

Edited by:
Rico Fischer, Technische Universität Dresden, Germany

Reviewed by:
Matthew J. C. Crump, Vanderbilt University, USA
Valentin J. Umbach, Humboldt-Universität zu Berlin, Germany

***Correspondence:**
Wout Duthoo and Wim Notebaert, Department of Experimental Psychology, Ghent University, Henri Dunantlaan 2, B-9000 Ghent, Belgium.
e-mail: wout.duthoo@ugent.be; Wim.Notebaert@ugent.be

In this paper, we aimed to investigate the role of self-generated predictions in the flexible control of behavior. Therefore, we ran a task switching experiment in which participants were asked to try to predict the upcoming task in three conditions varying in switch rate (30, 50, and 70%). Irrespective of their predictions, the color of the target indicated which task participants had to perform. In line with previous studies (Mayr, 2006; Monsell and Mizon, 2006), the switch cost was attenuated as the switch rate increased. Importantly, a clear task repetition bias was found in all conditions, yet the task repetition prediction rate dropped from 78 over 66 to 49% with increasing switch probability in the three conditions. Irrespective of condition, the switch cost was strongly reduced in expectation of a task alternation compared to the cost of an unexpected task alternation following repetition predictions. Hence, our data suggest that the reduction in the switch cost with increasing switch probability is caused by a diminished expectancy for the task to repeat. Taken together, this paper highlights the importance of predictions in the flexible control of behavior, and suggests a crucial role for task repetition expectancy in the context-sensitive adjusting of task switching performance.

Keywords: task switching, proactive cognitive control, expectancy bias, switch cost, prediction

INTRODUCTION

A hallmark of human cognition lies in the ability to proactively anticipate relevant future events and steer both action and perception accordingly. Current influential theories of cognition advance this proactive prediction generation ability as a central mechanism of brain functioning, marking a shift away from the view of the brain passively reacting to incoming stimulation. Predictive representations of both visual (e.g., Bar, 2007; Summerfield and Egner, 2009), auditory (Kumar et al., 2011), and olfactory (Zelano et al., 2011) information have been shown to guide and prepare the brain for a forthcoming stimulus, aiding information processing in a noisy and unpredictable environment. By continuously generating predictions about the environment, the cognitive system is also able to learn and associate specific actions or stimuli with specific outcomes. Learning on the basis of these prediction-driven outcomes is ascribed a central role in optimizing action selection and response execution in recent modeling work (Alexander and Brown, 2011; Silvetti et al., 2011). In line with the conception of the predictive brain, this paper aimed to investigate how self-generated predictions can flexibly steer attentional control through advance preparation, by referring to recent empirical work in the Stroop conflict task (Duthoo et al., submitted) and providing new evidence from a task switching experiment.

Attentional control is typically studied by means of a conflict paradigm, such as the Stroop conflict task (see MacLeod, 1991, for a review). In this task, participants are asked to respond to the color of a color word while ignoring its meaning. As the color and word dimension of the stimulus can either overlap or not, easy (congruent) and difficult (incongruent) stimulus conditions are created, respectively. Optimal task performance requires adaptively adjusting attention to the relevant (color) and irrelevant (word meaning) dimension. In general, these attentional adjustments can be grouped into two categories based on the underlying mechanism and the moment in time they are implemented by the cognitive system (Egner, 2007; Wühr and Kunde, 2008). According to a *reactive* control account, adjustments to the control settings occur in response to the target, corresponding to the metaphor of the reactive brain. Current models typically assume that it is the conflict on a given trial that triggers subsequent control up-regulation, characterized by a strengthening of task-relevant associations (Botvinick et al., 2001; Verguts and Notebaert, 2008, 2009). This theoretical framework has been successfully applied to many attentional control phenomena, including the reduction of the congruency effect following high-conflict incongruent trials in single-task paradigms (i.e., the Gratton effect; Gratton et al., 1992; for a review, see Egner, 2007), but also the increase of the switch cost following high-conflict incongruent stimuli in dual-task paradigms (Goschke, 2000; Braem et al., submitted). Alternatively, control adjustments can also be triggered in anticipation of the upcoming task or target, biasing the task or attentional set proactively. These *proactive* control adjustments, captured by the metaphor of the predictive brain described above, have received considerably less attention in the cognitive control literature.

In order to investigate this type of expectancy-induced control, two different strategies have been pursued. On the one hand, subjects' expectancies can be manipulated implicitly. Studies on attentional control have, for example, manipulated the proportion of incongruent trials (Logan and Zbrodoff, 1979) or congruency level transitions (Duthoo and Notebaert, 2012) to induce preparatory strategic control adjustments. Whereas the first manipulation successfully triggered anticipatory control, reflected in faster reactions to highly expected incongruent trials than to unexpected congruent trials, the second, more subtle manipulation appeared not strong enough to elicit expectancy-induced adaptation effects that were clearly dissociable from reactive, conflict-induced adjustments (see also Jiménez and Méndez, 2012). Alternatively, a more common and widespread experimental tool to probe anticipatory control adjustments is to cue participants explicitly about the upcoming stimulus event (for some early experimental work with the cueing paradigm, see Neill, 1978; Logan and Zbrodoff, 1982; Harvey, 1984). More recently, Aarts and Roelofs (2011) applied a probabilistic cueing procedure to a Stroop-like task to point out that the anticipation of upcoming conflict (or lack of conflict) can trigger similar sequential adjustments as experienced conflict (or lack thereof) on the previous trial, both behaviorally and in the activation pattern of the anterior cingulate cortex (ACC). In similar vein, Correa et al. (2009) found that anticipating conflict in a cued congruency task sped up both conflict detection and conflict resolution.

However, investigating proactive control by means of a cueing paradigm is not really testing the implications of a predictive brain, as it is assumed that we constantly generate predictions ourselves. Compared to the large amount of studies concerning cue-induced attentional control, few studies have centered on the effect of self-generated predictions on subsequent processing. Yet, human predictive behavior itself has been the focus of much experimental work outside the field of cognitive control. Interestingly, an influential line of research revealed that people's predictions and expectancies are often strongly biased (e.g., Kahneman et al., 1982), as they either overestimate or underestimate the actual probability of events to occur (see also Ayton and Fischer, 2004). When confronted with a random run of stimuli, participants will typically indicate that longer runs of a particular event have to be balanced out by the occurrence of the alternative event, a phenomenon known as the gambler's fallacy. This tendency for negative recency is also typically found when people are asked to generate or identify a random sequence (see Nickerson, 2002, for a review). However, other studies have shown that people can also display the opposite expectancy bias, the tendency to predict positive recency. A study of Kareev (1995), for example, in which participants were asked to predict the next item on the list, revealed that subjects typically overestimate repeating events. According to Kareev, this repetition bias stems from a persistent tendency to perceive or find patterns and causality in the environment (note, however, that the same tendency, seen from another perspective, can also result in probability matching behavior at the outcome level, the strategy to predict the events in proportion to their probability of occurrence; see Gaissmaier and Schooler, 2008). Apart from its impact on simple serial two-choice reaction time tasks (Remington, 1969; Soetens et al., 1985), the impact of this expectancy bias on information processing and attentional control remains still relatively uninvestigated. Given both these persistent prediction biases and the cognitive system's inherent prepotency to generate predictions and evaluate its outcomes, investigating self-generated expectancies and comparing their impact on subsequent processing to that of exogenously triggered expectancies might reveal new insights into how the brain implements proactive control.

In a previous study (Duthoo et al., submitted), we undertook a first attempt to measure these biased predictions explicitly and verify their influence on cognitive control by subjecting participants to a Stroop task and letting them predict the congruency level of the upcoming Stroop stimulus. Interestingly, after recoding participants' predictions ("Do you expect a congruent or incongruent trial?") relative to the congruency level of the previous trial, results revealed a clear repetition bias in the prediction pattern: in line with Kareev (1995), participants expected the congruency level to repeat from one trial to the next in 65% of all cases, even though congruency level repetition probability was set at 50%. Moreover, attentional adjustments (i.e., a Gratton effect) were only found when they anticipated a congruency level repetition. Participants showed both a reduced interference of repeating conflict trials (by proactively narrowing attention to the stimulus color) and increased facilitation of repeating non-conflicting trials (by proactively allowing the word meaning to influence response selection). In case of an unexpected congruency level alternation, these preparatory adjustments backfired and longer reaction times were registered, resembling the results of Aarts and Roelofs (2011) in a probabilistic cueing experiment. Interestingly, analyses of the congruency alternation predictions also suggested that in anticipation of an alternation, participants seemed to switch to a default control mode, as no sequential adjustments were found. In sum, the study revealed a clear bias toward predicting repeating events, and an optimization of control processes (i.e., a Gratton effect) in anticipation of such repeating events. Alternation expectancies, on the other hand, did not induce preparatory control.

Contrary to the literature on conflict control, the contribution of a preparatory component in task switching research has played a central role in the theoretical debate (e.g., see Karayanidis et al., 2010 and Kiesel et al., 2010 for recent overviews), overshadowing research on the reactive priming effects of the previous task-set on current task performance. In order to investigate these proactive adjustments, similar strategies have been implemented, aimed at inducing expectancies either implicitly or explicitly. As an example of the former strategy, fixed (predictable) task sequences (i.e., the alternating-runs paradigm; Rogers and Monsell, 1995) have been introduced to compare predictable task switch trials to predictable repetition trials. Even though two simple tasks were used and the task sequence was entirely predictable, this paradigm consistently evoked increased reaction times and higher error rates on switch compared to repetition trials (i.e., robust switch costs). To probe the impact of explicit expectancies on these switch costs, the explicit cueing paradigm (Meiran, 1996) was developed, in which cues specified the required task in a random run of task repetitions and switches. This cueing paradigm has been extensively used to evidence preparatory reductions in switch costs (e.g., Meiran, 1996; Koch, 2003), albeit not without its own set of methodological pitfalls (see Logan and Bundesen, 2003; but see also De Baene and Brass, 2011 and Jost et al., submitted).

Whereas a previous single-task study (Duthoo et al., submitted) suggested that alternation expectancies did not induce preparatory control adjustments, research on task switching has convincingly shown how increasing the preparation interval prior to an anticipated task alternation led to more controlled processing (i.e., a reduced switch cost). Monsell et al. (2003), for example, reported performance benefits for predictable compared to unpredictable task switches, suggesting that participants can strategically control their task-set readiness in function of their expectation, and, more precisely, in function of the probability of encountering a task switch on the upcoming trial. In similar vein, further research has robustly found a reduced switch cost with increasing switch probability (Mayr, 2006; Monsell and Mizon, 2006; Schneider and Logan, 2006; Bonnin et al., 2011). Others have pointed out that not only when expecting a task alternation, but also in anticipation of an expected task repetition, task-set readiness can be adjusted for optimal task performance, resulting in strong repetition benefits (Dreisbach et al., 2002). In sum, more so than in single-task paradigms, dual-task performance seems to rely on a strong anticipatory control component.

Even though the theoretical debate about this anticipatory control component is still ongoing, a key role is usually attributed to repetition expectancy. For example, the smaller difference between switch and repeat trials in a context with a 50% compared to a 30% switch probability is sometimes explained by the fact that participants match their task preparation to the probability of the switch and repeat conditions, thus equally preparing both tasks in a 50% switch probability context (Dreisbach et al., 2002; Brass and von Cramon, 2004; Monsell and Mizon, 2006). Alternatively, other authors suggested that people prepare the other task on part of the trials (e.g., Monsell and Mizon, 2006), resulting in extra preparation and thus longer reaction times on task repetition trials (when their guess was wrong) and less preparation and thus faster reactions to task switch trials (when their guess was right). Importantly, both explanations stress the importance of expectancies about the upcoming task. However, as indicated above, past research has consistently found that people's predictions are biased and therefore often do not match the actual probability in a given context (especially in the context of a random sequence of events; but see the work of Gaissmaier and Schooler, 2008, showing that the search for patterns can also result in probability matching at the outcome level). Moreover, the abovementioned studies never measured expectancies themselves, so that it remains a question for further research how expectancies can steer task preparation.

To shed some light on this issue, as well as to compare self-generated predictions in a dual-task paradigm to previous findings in a single-task paradigm, we decided to apply a similar procedure as our previous study on prediction-driven adjustments in the Stroop task (Duthoo et al., submitted). Therefore, we asked participants to try to predict the upcoming task on a trial-by-trial basis in one of three between-subjects conditions varying in switch rate (30, 50, and 70%), and probed both how these contexts affected the prediction pattern and how these predictions themselves influenced the task switch cost. Similar to our previous findings in the Stroop task, we expected predictions to evoke advance preparation for the upcoming target. More specifically, we expected repetition predictions to induce a strong reaction time benefit when a task repetition was actually presented, and a huge cost when one had to unexpectedly switch tasks, irrespective of condition. In contrast with the strong switch costs (and repetition benefits) following repetition predictions, we expected that alternation predictions evoke less strong preparatory effects (Duthoo et al., submitted), thereby reducing the switch cost, irrespective of condition. Consequently, assuming that participants' tendency to predict task repetitions is attenuated with increasing switch probability, we predicted to replicate the finding of a reduced switch cost in contexts of higher switch probabilities (Mayr, 2006; Monsell and Mizon, 2006; Schneider and Logan, 2006).

MATERIALS AND METHODS

PARTICIPANTS

Forty-eight Ghent University students (14 males; age: 17–28) signed up to participate in one of the three conditions ($n = 16$) of the experiment, lasting approximately 45 min. They received a monetary payment in return. Prior to the testing, participants provided written informed consent.

STIMULI AND APPARATUS

A program written with T-scope software (Stevens et al., 2006) controlled the experiment. All stimuli were displayed on a 17-inch monitor, with a viewing distance of approximately 50 cm. The numbers 1–9, with the exclusion of 5, served as the target stimuli, presented in Arial, font size 32. These stimuli were presented centrally on a black background in yellow (for the magnitude task) or blue (for the parity task). Responses were registered by means of a QWERTY keyboard.

DESIGN AND PROCEDURE

Participants were randomly assigned to one of the three experimental conditions, differing only in the amount of task switches during the three blocks where an explicit task prediction was registered. In the *repetition condition*, the task switch probability was restricted to 30%. In the *intermediate condition*, participants were confronted with an equal amount of task repetitions and alternations (50%). The *alternation condition* increased the task switch probability to 70%.

Throughout all blocks of the experiment, each target number was equally often presented in blue and yellow, implying that within each block participants performed an equal amount of magnitude and parity judgments. Selection of the target number was pseudo-random, with the restriction that each of the eight possible number targets appeared an equal amount of times in each of the two possible colors within one block. In all dual-task blocks, consisting of 80 trials, each target number was thus presented five times in both blue and yellow. Participants had to respond by pressing the E or U keyboard key for small or even target numbers and the R or I keyboard key for large or odd target numbers. The mapping of the task (magnitude or parity) to the middle and index finger of the left hand (keys E and R, respectively) or index and middle finger of the right hand (keys U and I, respectively) was counterbalanced across participants. In order to indicate which of the two tasks they expected, participants had to press the V or N key with their thumbs. The mapping of these keys to either a magnitude or parity task prediction was compatible with the mapping of the left or right hand to one of the two tasks.

In all conditions, participants were first trained on each of the two tasks separately during 40 trials of first magnitude and then parity judgments, adding up to 80 single-task practice trials. Hereafter, the two tasks where combined during two blocks of 80 trials, as to familiarize participants with the dual-task procedure. For these dual-task training blocks the task switch probability was kept at 50% in all three conditions. The color in which targets were presented indicated the task participants had to perform. A yellow number target asked for a magnitude judgment, whereas a blue target required a parity response. In the final phase of the experiment, three blocks of 80 trials were presented during which participants first had to predict which of the two tasks they expected to come next. Irrespective of their choices, the color in which the upcoming target was presented again indicated which of the two task participants had to perform, thereby serving as a feedback signal for their task predictions. For their performance on the target numbers no error feedback was provided. A store coupon was promised to the participant who performed best in the three last blocks for each condition, taking into account both the amount of correct predictions and mean reaction times and error percentages. In between blocks, participants took a short, self-paced break. After completing the experiment, participants filled in a short questionnaire, probing their awareness of the switch probability manipulation and their use of strategies in predicting the task sequence.

Each trial started with the presentation of a fixation cross for 500 ms. In the training blocks, this was followed by the target, which appeared on the screen until a response was registered, with the maximal reaction time restricted to 2500 ms. Next, the screen turned black for 500 ms, serving as the inter-trial-interval. In trials in which participants also had to predict the task on the next trial, a fixation cross was first presented for 500 ms, after which an instruction appeared on the screen ("Next trial?") that remained visible on the screen until participants clicked one of the two designated keyboard keys. Hereafter, a fixation cross was again displayed for 500 ms, after which a number target appeared on the screen, with identical timing values as described above.

RESULTS

In the results section, we focus on the three experimental blocks in which predictions were also registered. Two participants who did not engage in the prediction task (by "predicting" the same task throughout at least one of the three experimental blocks) were removed from the analysis, restricting the number of participants in the intermediate and alternation condition to 15. Non-responses and badly recorded data (adding up to 1.6%) were excluded from both the reaction time and performance error analysis. We applied the multiple comparison correction method put forward by Holm (1979) in order to control for the family wise error rate, adjusting the p-values of the post tests in the reaction time and error analysis accordingly.

REACTION TIMES AND PREDICTIONS

Before conducting the reaction time analysis, the data were subjected to a trimming procedure. We first excluded the trials on which participants committed an error (8.1% of the remaining data; distributed equally over the three conditions). Hereafter, the first trial of each block and RT outliers (±2.5 SD, calculated separately per condition, subject, and task) were removed (another 3.9%). Taken together, the analysis was thus carried out on 86.9% of the complete data.

First, a mixed-design analysis of variance with the between-subjects variable Condition (three levels: repetition, intermediate, and alternation) and the within-subjects variables Task (two levels: magnitude and parity) and Sequence (two levels: repetition and alternation) was carried out. Results revealed main effects of Task, $F(1, 43) = 57.36$, $p < 0.0001$, reflecting faster magnitude than parity judgments (757 and 877 ms, respectively) and Sequence, $F(1, 43) = 116.95$, $p < 0.0001$, indicating the presence of a switch cost of 106 ms, but not a main effect of Condition, $F(2, 43) < 1$, ns. The two-way interaction between Task and Sequence turned out significant as well, $F(1, 43) = 5.47$, $p < 0.05$, reflecting a larger switch cost for the parity task compared to the magnitude task (120 and 93 ms, respectively), irrespective of Condition, $F(2, 43) < 1$, ns. Most importantly, the analysis revealed a two-way interaction between Sequence and Condition, $F(2, 43) = 11.05$, $p < 0.0001$, implying that the size of the switch cost was significantly affected by the transitional manipulation. Further independent-samples t-tests showed that, compared to the switch cost of 112 ms in the intermediate condition, the switch cost was significantly reduced to 52 ms by increasing the switch probability in the alternation condition, $t(28) = 3.5$, $p < 0.01$. Decreasing the switch probability to 30% in the repetition condition significantly increased the switch cost to 166 ms compared to the alternation condition, $t(29) = 4.5$, $p < 0.0001$. The increase in switch cost of 54 ms in the repetition compared to the intermediate condition was only marginally significant, $t(29) = 2.0$, $p = 0.056$. These differences in the switch cost over conditions are depicted in **Figure 1**.

Next, we took a deeper look into participants' task prediction patterns. Irrespective of condition, participants predicted the magnitude and parity task equally often (i.e., 50%, on average,

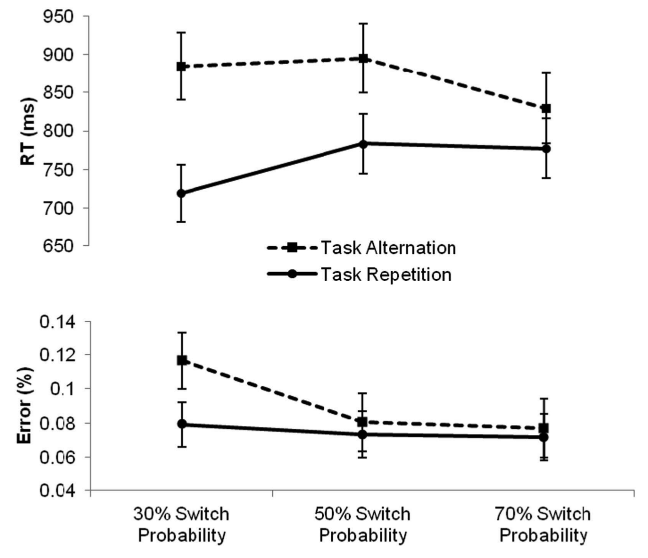

FIGURE 1 | Mean reaction times (RTs, in milliseconds) and error percentages for task repetitions (full line) and task alternations (dashed line) in the three conditions varying in switch probability. Error bars represent 95% confidence intervals around the mean.

SD = 5.4%). These task predictions were then recoded into repetition or alternation predictions, relative to the task presented on the previous trial. In line with our manipulation of task switch probability, participants in the repetition condition predicted more task repetitions (78%), both compared to participants in the intermediate [66%, independent-samples $t(29) = 3.1$, $p < 0.001$] and participants in the alternation condition [51%, independent-samples $t(29) = 7.63$, $p < 0.0001$]. Remarkably, in all three conditions a task repetition bias was found, as comparisons between the task switch prediction rate and the actual task switch probability indicated that both in the intermediate condition [66% compared to 50%, $t(14) = 6.77$, $p < 0.0001$], repetition condition [78% compared to 70%, $t(15) = 2.86$, $p < 0.05$] and alternation condition [51% compared to 30%, $t(14) = 8.39$, $p < 0.0001$] the amount of task repetitions was consistently overpredicted.

Finally, we examined the effect of these task predictions on task performance, by investigating how repetition and alternation expectations impacted the switch cost. To this end, we ran a mixed-design analysis of variance with the between-subject variable Condition (three levels: repetition, intermediate, and alternation) and the within-subjects variables Prediction and Sequence (two levels: repetition and alternation)[1]. Apart from the main effect of Sequence, $F(1, 43) = 59.89$, $p < 0.0001$, reflecting a switch cost, the analysis also revealed a marginally significant main effect of Prediction, $F(1, 43) = 3.87$, $p = 0.056$, indicating that number targets were responded to 17 ms slower following alternation predictions than following repetition predictions. Importantly, a significant interaction between Prediction and Sequence was also found, $F(1, 43) = 88.75$, $p < 0.0001$. The three-way interaction with Condition did not reach significance, $F(2, 43) < 1$, ns, suggesting that participants' predictions influenced the switch cost similarly in all three conditions. Following an alternation prediction, the switch cost, calculated as the difference between an expected task alternation and an unexpected task repetition, disappeared completely. Even though inspection of the reaction times suggested a switch benefit numerically (24, 31, and 32 ms in the repetition, intermediate, and alternation condition, respectively), post tests indicated that this difference did not reach statistical significance in any of the conditions (all ps > 0.62). Following a repetition prediction, a huge and significant repetition benefit, calculated as the difference between an unexpected task alternation and an expected task repetition, was found in all conditions (222, 116, and 147 ms in the repetition, intermediate, and alternation condition, respectively; all ps < 0.0001). This pattern of reaction times is visualized in **Figure 2**.

ERROR RATES

First, we ran a repeated-measures ANOVA with the between-subjects variable Condition (three levels: repetition, intermediate, and alternation) and the within-subjects variables Task (two levels: magnitude and parity) and Sequence (two levels: repetition and alternation) on the aggregated error scores. Similarly to the reaction time analysis, we found main effects of Task, $F(1, 43) = 36.61$, $p < 0.0001$, reflecting worse performance on parity than magnitude judgments (12 and 4.6%, respectively) and Sequence, $F(1, 43) = 9.51$, $p < 0.01$, indicating higher error rates on task alternations than on task repetitions (9.2 and 7.4%, respectively), but no main effect of Condition, $F(2, 43) < 1$, ns. The two-way interaction between Task and Sequence also reached significance, $F(1, 43) = 5.07$, $p < 0.05$, indicating that switching to a parity task (compared to repeating this task) increased the error rate (3.2%), whereas switching to a magnitude task did not. Most importantly, we again found a significant interaction between Sequence and Condition, $F(2, 43) = 3.22$, $p < 0.05$, indicating that the size of the error switch cost differed significantly between the three conditions, irrespective of task, $F(2, 43) < 1$, ns. Further independent-samples t-tests revealed that this interaction was brought about by a significant increase in the error switch cost (3.7%) in the repetition condition compared to the intermediate condition, $t(29) = 2.09$, $p < 0.05$, whereas the error switch cost was not statistically lower in the alternation condition compared to the intermediate condition, $t(28) < 1$, ns. The error rates for task repetitions and task alternations in each of the three conditions are visualized in **Figure 1**.

In order to investigate how participants' predictions had an impact on the error rates, we conducted another repeated-measures ANOVA with the between-subjects variable Condition (three levels: intermediate, repetition, and alternation) and the within-subjects variables Prediction and Sequence (two levels: repetition and alternation). This analysis revealed only a main effect of Prediction, $F(1, 43) = 5.73$, $p < 0.05$, indicating that an alternation prediction produced more erroneous responses compared to a repetition prediction (9.6 and 7.8%, respectively). The two-way interaction between Prediction and Sequence was only marginally significant, $F(1, 43) = 3.3$, $p = 0.076$. The data pattern closely resembled the reaction pattern, showing a trend for the error switch cost to be absent following alternation predictions, and present following repetition predictions. Again, this pattern did not differ significantly between the three conditions, $F(2, 43) < 1$, ns. The error rates for task repetitions and task alternations following repetition and alternation predictions in each of the three conditions are presented in **Figure 2**.

DISCUSSION

In the present study, we aimed to investigate how self-generated predictions influence conflict and task control, expanding previous research on expectancy-induced proactive control. To do so, we inserted explicit task predictions into a task switching procedure, thereby complementing as well as elaborating on a previous experiment in which the influence of congruency level predictions on subsequent Stroop performance was put to the test (Duthoo et al., submitted). Results revealed three interesting findings.

Firstly, analysis of participants' prediction patterns exposed a bias toward predicting task repetitions in all three conditions. In the intermediate condition, in which the two tasks alternated in 50% of all transitions, participants displayed a clear task repetition

---

[1] We did not include the variable Task in this analysis, as this would cause some of the cells of the ANOVA to be calculated on a very limited amount of observations (for instance: a switch to the parity task following an alternation prediction in the repetition condition). We therefore collapsed observations over the two tasks. Still, running the analysis with the Task variable included did not change the pattern of the results. Importantly, the Task variable did not interact significantly with any of the other variables (all ps > 0.14).

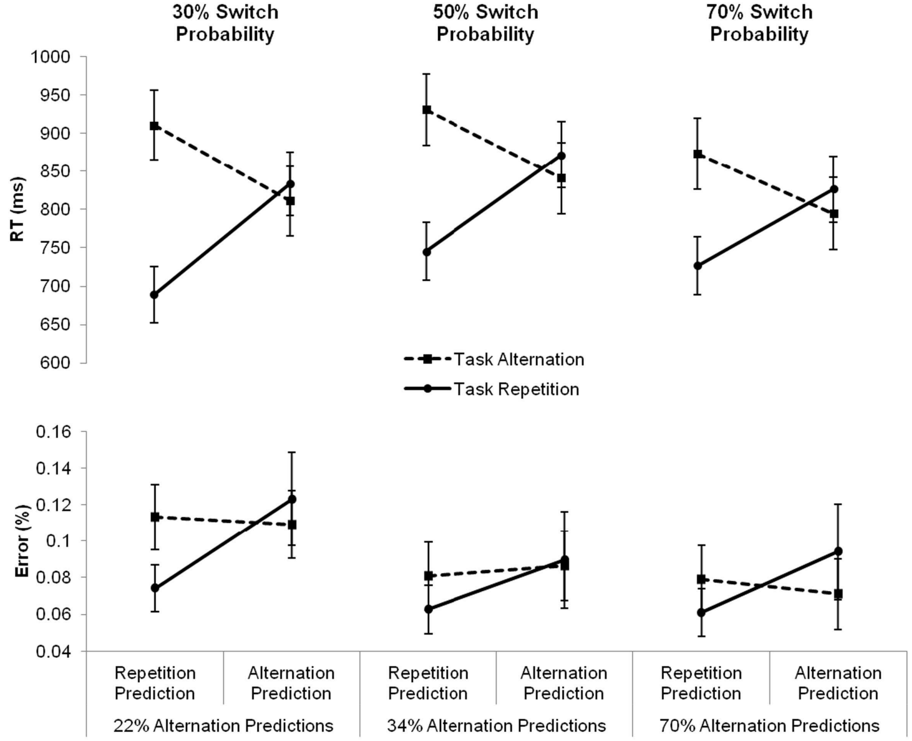

FIGURE 2 | Mean reaction times (RTs, in milliseconds) and error percentages for task repetitions (full line) and task alternations (dashed line) following repetition and alternation predictions, separately for the three conditions varying in switch probability. Under each of the graphs, the corresponding overall percentage of alternation predictions is presented. Error bars represent 95% confidence intervals around the mean.

bias (66%). Also in the alternation condition, participants still predicted a task repetition in 51% of all transitions, when only 30% were actually presented. Moreover, reaction times and error rates showed that irrespective of condition, reactions following a task alternation prediction were slower and more error-prone. At first sight, this tendency to predict repeating stimulus events, or "hot hand fallacy," might seem at odds with the literature on probability matching (Gaissmaier and Schooler, 2008), revealing participants' tendency to match their choice behavior to the actual probability of two stimuli that are not equally likely to be presented. Yet, given that participants in the current experiment were asked to predict the upcoming task rather than the task transition, participants matched probabilities quite well, as irrespective of condition the two tasks were predicted equally often (i.e., 50%). Still, further insight into the transitional probabilities could help them predicting the upcoming task more accurately. Yet, these transitional probabilities were less readily picked up, since the experiment revealed a clear bias toward expecting repetitions. Interestingly, participants' prediction error rate only dropped from 50 to 38% in the repetition condition [$t(15) = 8.9, p < 0.0001$], in which transitional probability was in line with their repetition expectancy bias.

Secondly, our manipulation of switch probability affected the switch cost as predicted: compared to the switch cost in the intermediate condition with a 50% switch probability, increasing this switch probability decreased the switch cost significantly, whereas decreasing the switch probability strongly amplified the switch cost. Put differently, the switch cost is attenuated under conditions of high switch probability, replicating previous studies (Mayr, 2006; Monsell and Mizon, 2006; Schneider and Logan, 2006; Bonnin et al., 2011). Moreover, results also revealed that switching to the parity task came at a greater cost than switching to the magnitude task, both in reaction time and accuracy. This corresponds well with previous research on asymmetries in switch costs showing that separating the response set of the two tasks results in greater costs in switching to the more difficult task (Yeung and Monsell, 2003). In the current experiment, response set overlap was reduced in terms of response decisions (parity versus magnitude judgments) and stimulus-response mapping (both tasks were mapped to separate hands). Most importantly, this task asymmetry did not interact with predictions, which formed the main focus of this study.

Thirdly, by inserting explicit predictions into the dual-task procedure, we were able to identify a potential mechanism underlying the finding of reduced switch costs in conditions with high switch probability. In all three conditions, the same prediction-driven behavioral adjustments were found: following an alternation prediction, the difference between repetition and switch trials disappeared, whereas repetition predictions were followed by a large switch cost (or a large repetition benefit). Participants in the alternation condition expected more alternations, thereby reducing the switch cost significantly. In other words, the reduction in switch cost in a context of high switch probability might stem from proactively switching to a more controlled processing strategy when expecting task alternations. However, preparing for a task alternation still comes at a cost, as comparisons between correctly predicted task repetitions and alternations revealed a significant residual task switch cost (all *p*s < 0.001). This finding is in line with studies using the explicit cueing paradigm that consistently show that even validly cued task alternations robustly slowed down responses compared to validly cued task repetitions (Meiran, 1996).

On an important note, part of the speed-up in reaction time following correct predictions might reflect an effect of hand priming, as in the current design correct predictions involved the finger of the same hand needed for subsequent task execution, whereas incorrect predictions entailed a switch of hand (e.g., Cooper and Marí-Beffa, 2008). Still, this definitely cannot account for the whole pattern of findings, since predicting the other task relative to the task on the previous trials correctly (i.e., a task alternation in which the same hand was used for predicting and responding to the target) did not produce reactions that were significantly faster than following incorrect task alternation predictions, in which the task repeated but the hand used for predictions differed from the hand used for responding to the target. Taken together, this study suggests that in a dual-task environment, participants expect the task to repeat, leading to improved performance when it does and a large cost when it alternates. Still, in anticipation of a task alternation, participants respond equally fast to a task alternation as to a task repetition. These conclusions are clearly in line with a proactive, expectancy-based account of task switching.

Moreover, the current findings allow drawing interesting parallels between this experiment and the aforementioned previous Stroop experiment, both in the patterns of self-generated expectancies as in their effect on subsequent processing. Compellingly, we found a robust bias toward overpredicting repeating events that was also present in congruency level predictions in the Stroop task. This bias toward expecting task repetitions coincided with a clear processing benefit for these repetition predictions, as alternation predictions typically induced higher errors rates and increased reaction times, irrespective of condition. Interestingly, the observation of reaction time benefits following repetition expectations but not after alternation expectations also bears a striking resemblance to findings within the voluntary task switching paradigm (Arrington and Logan, 2004). In this paradigm, participants can choose which task to perform on a series of bivalent stimuli, with the instruction to perform both tasks equally often. In line with the inherent bias toward repetitions defended in this paper, Arrington and Logan found that the subjects produced more task repetitions (i.e., 68%) than expected if the tasks were performed in a pure random sequence. Moreover, deliberately choosing to switch tasks slowed down task performance significantly (i.e., a significant *switch cost* was found). Taken together, the experiment revealed that participants displayed a clear reluctance to switch tasks.

Similar to the voluntary task repetition and switch decisions, repetition and alternation predictions clearly produced a differential effect on subsequent processing: repetition predictions were followed by a strong reaction time benefit when an actual task repetition was presented, and a large cost when one then had to (unexpectedly) switch. Again, this pattern closely resembled findings in our previous Stroop study (Duthoo et al., submitted), where a clear congruency level repetition benefit and congruency level alternation cost were found following repetition predictions. Yet, whereas congruency level alternation predictions were not followed by behavioral adaptations in the Stroop task, the current experiment showed that following task alternation predictions the difference between an actually presented task alternation and an unexpected task repetition disappeared.

Crucially, this pattern of results did not differ between the three conditions varying in switch probability. Therefore, the present experiment suggests an explanation for the often replicated finding of reduced switch costs in conditions with a higher switch probability (Mayr, 2006; Monsell and Mizon, 2006; Schneider and Logan, 2006): increasing the switch probability increases the expectancy for task alternations, which was found to be followed by a reduction in the switch cost. However, the interpretation of this reduced switch cost in anticipation of a task alternation is still open to debate.

One possible explanation, as was also put forward by Monsell and Mizon (2006), is that participants adopt a "neutral control state," right in between the two task-sets. When the color of the target then indicated which of the two task-sets was appropriate, reactions to either one of the two tasks would be equally fast. This is exactly the pattern of results we found following alternation predictions, and it emerged in all three conditions. Moreover, this corresponds well with the absence of sequential modulations of the Stroop effect following congruency level alternation predictions, which was also explained by participants adopting a "neutral control mode" (Duthoo et al., submitted).

Alternatively, one can assume that both repetition and alternation predictions lead to advance preparation of the upcoming task, yet preparation for task alternations is never complete (i.e., there is a residual switch cost, e.g., Meiran et al., 2000). Also in our experiment, correctly predicted task alternations were responded to much slower than correctly predicted task repetitions, irrespective of condition. In case of a correctly predicted task alternation, advance preparation speeds up responding compared to an unexpected task alternation (i.e., following a task repetition prediction). Yet, because of a residual switch cost, these reactions are not significantly faster than those to unexpected task repetitions (i.e., following a task alternation prediction), where preparation misfires, but no residual switch cost affects performance. The same logic holds if one assumes the difference between switch and repeat trials to arise from adaptation to the task-set on repetition trials, reflected in a repetition benefit, rather than from reconfiguration of the task-set on switch trials, reflected in a (residual) switch cost (De Baene et al., 2012). In the case of an unexpected task repetition following a task alternation prediction, reaction times will be relatively slower than for expected task repetitions, yet equally fast to an expected task alternation, where no task-set adaptation benefit was present. However, the current data do not allow differentiating between the adaptation and reconfiguration view, as both predict the same data pattern: following correct repetition predictions, both preparation and task-set adaptation (or lack of reconfiguration) will speed-up an actual task repetition, whereas following correct alternation predictions, preparation, and the lack of task-set adaptation (or need for reconfiguration) have effects in opposite directions, explaining the intermediate reaction times. Whether this explanation in terms of equal preparation for switch and repeat trials following both types of predictions is to be favored over an explanation in terms of a lack of specific preparation for alternation predictions (i.e., a neutral control mode) is an interesting question for future research.

Yet, the current experiment applied a 1:1 mapping between the cue (i.e., the color of the target) and the task (i.e., a magnitude or parity judgment), so that task repetitions were confounded with repetitions of the cue. Therefore, this design does not allow teasing apart the facilitatory effect of repeated-cue-encoding in task repetitions from the effect of executive control processes reconfiguring the cognitive system in task alternations. In order to disentangle cue repetitions from task repetitions, some previous studies have introduced multiple cues per task (e.g., Logan and Bundesen, 2003; Mayr and Kliegl, 2003; see Schneider and Logan, 2011, for a comparison between 1:1 and 2:1 cue-to-task mappings). This approach has led to a rich body of empirical evidence showing that repetition priming of cue encoding is indeed an important component of task switching. Note, however, that these studies have also demonstrated that there are usually also substantial "true" task switch costs remaining (for a review of this evidence, see Jost et al., submitted).

Important in the light of the current results is a study of Schneider and Logan (2006), in which this 2:1 cue-to-task mapping was combined with a transitional probability manipulation similar to ours. In line with the current findings, switch costs were smallest in the condition with a high switch probability and largest when the amount of task repetitions was increased. Modeling of their data led these authors to conclude that the difference in the switch costs between different frequency conditions reflected (automatic or strategic) priming of cue encoding for the frequent transitions. Therefore, an interesting avenue for future research lies in combining a 2:1 mapping strategy with our prediction manipulation to elucidate whether the prediction-driven adjustments in task switching performance reported in this paper were driven by facilitating the speed of cue encoding rather than by promoting advance configuration of task-set.

Given the emphasis recent theories place on prediction-driven adjustments in brain functioning, the paradigm to assess self-generated predictions and probe their impact presented in the current article seems a particularly promising tool for further research. Applying this method, we were able to pinpoint structural biases in human predictions and measure their influence on subsequent processing in a direct way, rather than inferring explanations in terms of expectancy indirectly from the data. Yet, one outstanding question remains whether participants will make similar predictions when they are not explicitly asked to generate them, and, consequently, to what extent these expectancy-driven attentional adjustments can also be found in "normal" Stroop or dual-task behavior.

In conclusion, the research presented in this paper advocated viewing the brain as a predictive rather than a purely reactive device. In this light, the overestimation of repeating events (also referred to as "the hot hand fallacy") should not necessarily be considered as a weakness of our predictive brain. In real life, there is a much stronger correlation between sequential events than in our artificial lab tasks. For instance, when the road is slippery because of wet conditions in one turn, it is usually a good idea to predict that also the next turn will be slippery and adjust accordingly. It therefore appears adaptive that the cognitive system is more readily optimized in anticipation of a repeating event. This is reflected in a strong repetition benefit for both congruency level and task repetitions. Yet, when interpreting the lack of conflict adaptation and the reduced difference between task repetition and alternations following alternation predictions in terms of participants adopting a neutral control mode, it remains an extremely interesting question to what extent our brain can also prepare for expected changes.

ACKNOWLEDGMENTS

The research reported in this article was supported by grant no. 3F009109 and FWO10/PDO/234 of the Research Foundation Flanders (FWO).

REFERENCES

Aarts, E., and Roelofs, A. (2011). Attentional control in anterior cingulate cortex based on probabilistic cueing. *J. Cogn. Neurosci.* 23, 716–727.

Alexander, W. H., and Brown, J. W. (2011). Medial prefrontal cortex as an action-outcome predictor. *Nat. Neurosci.* 14, 1338–1344.

Arrington, C. M., and Logan, G. D. (2004). The cost of a voluntary task switch. *Psychol. Sci.* 15, 610–615.

Ayton, P., and Fischer, I. (2004). The hot hand fallacy and the gambler's fallacy: two faces of subjective randomness? *Mem. Cognit.* 32, 1369–1378.

Bar, M. (2007). The proactive brain: using analogies and associations to generate predictions. *Trends Cogn. Sci. (Regul. Ed.)* 11, 280–289.

Bonnin, C. A., Gaonac'h, D., and Bouquet, C. A. (2011). Adjustments of task-set control processes: effect of task switch frequency on task-mixing and task-switching costs. *J. Cogn. Psychol.* 23, 985–997.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., and Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychol. Rev.* 108, 624–652.

Brass, M., and von Cramon, D. Y. (2004). Decomposing components of task preparation with functional magnetic resonance imaging. *J. Cogn. Neurosci.* 16, 609–620.

Cooper, S., and Marí-Beffa, P. (2008). The role of response repetition in task switching. *J. Exp. Psychol. Hum. Percept. Perform.* 34, 1198–1211.

Correa, Á., Rao, A., and Nobre, A. C. (2009). Anticipating conflict facilitates controlled stimulus–response selection. *J. Cogn. Neurosci.* 21, 1461–1472.

De Baene, W., and Brass, M. (2011). Cue-switch effects do not rely on the same neural systems as task-switch effects. *Cogn. Affect. Behav. Neurosci.* 11, 600–607.

De Baene, W., Kühn, S., and Brass, M. (2012). Challenging a decade of brain research on task switching: brain activation in the task-switching paradigm reflects adaptation rather than reconfiguration of task sets. *Hum. Brain Mapp.* 33, 639–651.

Dreisbach, G., Haider, H., and Kluwe, R. H. (2002). Preparatory processes in the task-switching paradigm: evidence from the use of probability cues. *J. Exp. Psychol. Learn. Mem. Cogn.* 28, 468–483.

Duthoo, W., and Notebaert, W. (2012). Conflict adaptation: it is not what you expect. *Q. J. Exp. Psychol.* doi: 10.1080/17470218.2012.676655. [Epub ahead of print].

Egner, T. (2007). Congruency effects and cognitive control. *Cogn. Affect. Behav. Neurosci.* 7, 380–390.

Gaissmaier, W., and Schooler, L. J. (2008). The smart potential behind probability matching. *Cognition* 109, 416–422.

Goschke, T. (2000). "Intentional reconfiguration and involuntary persistence in task set switching," in *Control of Cognitive Processes XVIII*, eds S. Monsell and J. Driver (Cambridge, MA: MIT Press), 331–355.

Gratton, G., Coles, M. G., and Donchin, E. (1992). Optimizing the use of information: strategic control of activation of responses. *J. Exp. Psychol. Gen.* 121, 480–506.

Harvey, N. (1984). The Stroop effect: failure to focus attention or failure to maintain focusing? *Q. J. Exp. Psychol. Hum. Exp. Psychol.* 36, 89–115.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scand. J. Stat.* 6, 65–70.

Jiménez, L., and Méndez, A. (2012). It is not what you expect: dissociating conflict adaptation from expectancies in a Stroop task. *J. Exp. Psychol. Hum. Percept. Perform.* doi:10.1080/17470218.2012.676655. [Epub ahead of print].

Kahneman, D., Slovic, P., and Tversky, A. (1982). *Judgement Under Uncertainty: Heuristics and Biases*. Cambridge: Cambridge University Press.

Karayanidis, F., Jamadar, S., Ruge, H., Phillips, N., Heathcote, A., and Forstmann, B. U. (2010). Advance preparation in task-switching: converging evidence from behavioral, brain activation, and model-based approaches. *Front. Psychol.* 1:25. doi:10.3389/fpsyg.2010.00025

Kareev, Y. (1995). Positive bias in the perception of covariation. *Psychol. Rev.* 102, 490–502.

Kiesel, A., Steinhauser, M., Wendt, M., Falkenstein, M., Jost, K., Philipp, A. M., and Koch, I. (2010). Control and interference in task switching – a review. *Psychol. Bull.* 136, 849–874.

Koch, I. (2003). The role of external cues for endogenous advance reconfiguration in task switching. *Psychon. Bull. Rev.* 10, 488–492.

Kumar, S., Sedley, W., Nourski, K. V., Kawasaki, H., Oya, H., Patterson, R. D., Howard, M. A., Friston, K. J., and Griffiths, T. D. (2011). Predictive coding and pitch processing in the auditory cortex. *J. Cogn. Neurosci.* 23, 3084–3094.

Logan, G. D., and Bundesen, C. (2003). Clever homunculus: is there an endogenous act of control in the explicit task-cuing procedure? *J. Exp. Psychol. Hum. Percept. Perform.* 29, 575–599.

Logan, G. D., and Zbrodoff, N. J. (1979). When it helps to be misled: facilitative effects of increasing the frequency of conflicting stimuli in a Stroop-like task. *Mem. Cognit.* 7, 166–174.

Logan, G. D., and Zbrodoff, N. J. (1982). Constraints on strategy construction in a speeded discrimination task. *J. Exp. Psychol. Hum. Percept. Perform.* 8, 502–520.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: an integrative review. *Psychol. Bull.* 109, 163–203.

Mayr, U. (2006). What matters in the cued task-switching paradigm: tasks or cues? *Psychon. Bull. Rev.* 5, 794–799.

Mayr, U., and Kliegl, R. (2003). Differential effects of cue changes and task changes on task-set selection costs. *J. Exp. Psychol. Learn. Mem. Cogn.* 29, 362–372.

Meiran, N. (1996). Reconfiguration of processing mode prior to task performance. *J. Exp. Psychol. Learn. Mem. Cogn.* 22, 1423–1442.

Meiran, N., Chorev, Z., and Sapir, A. (2000). Component processes in task switching. *Cogn. Psychol.* 41, 211–253.

Monsell, S., and Mizon, G. A. (2006). Can the task-cuing paradigm measure an endogenous task-set reconfiguration process? *J. Exp. Psychol. Hum. Percept. Perform.* 32, 493–516.

Monsell, S., Sumner, P., and Waters, H. (2003). Task-set reconfiguration with predictable and unpredictable task switches. *Mem. Cognit.* 31, 327–342.

Neill, W. T. (1978). Decision processes in selective attention: response priming in the Stroop color-word task. *Percept. Psychophys.* 23, 80–84.

Nickerson, R. S. (2002). The production and perception of randomness. *Psychol. Rev.* 109, 330–357.

Remington, R. J. (1969). Analysis of sequential effects in choice reaction times. *J. Exp. Psychol.* 82, 250–257.

Rogers, R. D., and Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *J. Exp. Psychol. Gen.* 124, 207–231.

Schneider, D. W., and Logan, G. D. (2006). Priming cue encoding by manipulating transition frequency in explicitly cued task switching. *Psychon. Bull. Rev.* 13, 145–151.

Schneider, D. W., and Logan, G. D. (2011). Task-switching performance with 1:1 and 2:1 cue-task mappings: not so different after all. *J. Exp. Psychol. Learn. Mem. Cogn.* 37, 405–415.

Silvetti, M., Seurinck, R., and Verguts, T. (2011). Value and prediction error in medial frontal cortex: integrating the single-unit and systems levels of analysis. *Front. Hum. Neurosci.* 5:75. doi:10.3389/fnhum.2011.00075

Soetens, E., Boer, L. C., and Hueting, J. E. (1985). Expectancy or automatic facilitation? Separating sequential effects in two-choice reaction time. *J. Exp. Psychol. Hum. Percept. Perform.*11, 598–616.

Stevens, M., Lammertyn, J., Verbruggen, F., and Vandierendonk, A. (2006). Tscope: A C library for programming cognitive experiments on the MS Windows platform. *Behav. Res. Methods* 38, 280–286.

Summerfield, C., and Egner, T. (2009). Expectation (and attention) in visual cognition. *Trends Cogn. Sci. (Regul. Ed.)* 13, 403–409.

Verguts, T., and Notebaert, W. (2008). Hebbian learning of cognitive control: dealing with specific and nonspecific adaptation. *Psychol. Rev.* 115, 518–525.

Verguts, T., and Notebaert, W. (2009). Adaptation by binding: a learning account of cognitive control. *Trends Cogn. Sci. (Regul. Ed.)* 13, 252–257.

Wühr, P., and Kunde, W. (2008). Precueing spatial S–R correspondence: is there regulation of expected response conflict? *J. Exp. Psychol. Hum. Percept. Perform.* 34, 872–883.

Yeung, N., and Monsell, S. (2003). Switching between tasks of unequal familiarity: the role of stimulus-attribute and response-set selection. *J. Exp. Psychol. Hum. Percept. Perform.* 29, 455–469.

Zelano, C., Mohanty, A., and Gottfried, J. A. (2011). Olfactory predictive codes and stimulus templates in piriform cortex. *Neuron* 72, 178–187.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 15 June 2012; accepted: 22 July 2012; published online: 08 August 2012.

Citation: Duthoo W, De Baene W, Wühr P and Notebaert W (2012) When predictions take control: the effect of task predictions on task switching performance. Front. Psychology ***3****:282. doi: 10.3389/fpsyg.2012.00282*

This article was submitted to Frontiers in Cognition, a specialty of Frontiers in Psychology.

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 27 July 2012
doi: 10.3389/fpsyg.2012.00265

How positive affect modulates proactive control: reduced usage of informative cues under positive affect with low arousal

Kerstin Fröber* and Gesine Dreisbach

Department of Psychology, University of Regensburg, Regensburg, Germany

Edited by:
Rico Fischer, Technische Universität Dresden, Germany

Reviewed by:
Todd S. Braver, Washington University, USA
Thomas Kleinsorge, Leibniz Research Centre for Working Environment and Human Factors, Germany

****Correspondence:***
Kerstin Fröber, Department of Psychology, University of Regensburg, Universitätsstr. 31, D-93053 Regensburg, Germany.
e-mail: kerstin.froeber@psychologie.uni-regensburg.de

An example of proactive control is the usage of informative cues to prepare for an upcoming task. Here the authors will present data from a series of three experiments, showing that positive affect along with low arousal reduces proactive control in form of a reduced reliance on informative cues. In three affect groups, neutral or positive affective picture stimuli with low and high arousal preceded every trial. In Experiments 1 and 2, using a simple response cueing paradigm with informative cues (66% cue validity), a reduced cue validity effect (CVE) was found under positive affect with low arousal. To test the robustness of the effect and to see whether reactive control is also modulated by positive affect, Experiment 3 used a cued task switching paradigm with predicitive cues (75% cue validity). As expected, a reduced CVE was again found specifically in the positive affect condition with low arousal, but only for task repetitions. Furthermore, there was no difference in switch costs between affect groups (with and without task cues). Taken together, the reduced CVE indicates that positive affect with low arousal reduces proactive control, while comparable switch costs suggest that there is no influence of positive affect on reactive control.

Keywords: cognitive control, positive affect, arousal

INTRODUCTION

Studying issues of cognitive control is of major interest for the understanding of human cognition and action. The dual mechanisms of control (DMC) framework by Braver and colleagues (Braver et al., 2007; Braver, 2012) suggests that cognitive control operates in two distinct modes, proactive control and reactive control. Reactive control is assumed to be transiently activated in a just-in-time manner as soon as a high interference event is detected. In contrast, proactive control is supposed to be activated by the anticipation of upcoming interference. So, reactive control serves as a "late correction" mechanism to solve interference after its onset, whereas proactive control serves to prevent interference before it occurs. For this purpose, goal-relevant information is actively sustained in preparation for an optimized behavior in the upcoming cognitively demanding event. The DMC framework further claims that successful cognition relies on the variability of these two cognitive control functions, and that various factors – intrapersonal, interpersonal, or situational – can lead to a bias in favor of one mode of control strategy over the other. An example for such a biasing factor is affect.

Dreisbach (2006), for example, investigated affective modulations of cognitive control with an AX Continuous Performance Task (AX-CPT). The author could show that specifically positive affect and not negative affect – manipulated via pictures from the International Affective Picture System (IAPS, Lang et al., 1999) preceding every trial – leads to a more flexible but also less stable behavior. In the AX-CPT participants have to press a prespecified target response key to the target "X" but only if it follows the cue "A." If X follows another letter (e.g., B) or A is followed by another letter than X (e.g., Y), the non-target response key has to be pressed. Critically, the cue A is highly informative about the occurrence of X (70% frequency of AX trials, whereas the other trial types BX, AY, and BY occur with 10% frequency each), therefore it can be assumed that in this task there is a strong bias in favor of a proactive control strategy with active maintenance of the cue information to optimize performance. Likewise, the cue B is also very informative, as it unequivocally predicts a non-target response. Dreisbach (2006) found improved performance in AY trials, but worsened performance in BX and BY trials under positive affect. This result was interpreted as evidence for a reduced maintenance of the cue, because subjects in the positive group showed costs when a to be maintained goal had to be executed (BX and BY trials; less stability) and benefits when a to be maintained goal unexpectedly changed (AY trials; more flexibility). According to the DMC framework (Braver et al., 2007; Braver, 2012), these results might provide evidence that proactive control is reduced under positive affect, because there is less usage of the cue to prepare the upcoming task (see also Compton et al., 2004). The increased flexibility, as indicated by the better performance on AY trials under positive affect, however, might as well be interpreted in terms of increased reactive control. In line with this interpretation, a recent study (van Wouwe et al., 2011) – also using the AX-CPT, but manipulating affect via emotional film clips before the actual experiment – more directly addressed the question whether positive affect influences proactive or reactive control by including measures of event related potentials (ERP). In line with the Dreisbach (2006) study, they found improved behavioral performance in AY trials, that is, on trials on which a cue-induced response tendency has to be overcome. However, unlike the previous study, the authors did not find impairment in BX and BY trials, where the cue unequivocally announced the non-target response. Based on these behavioral results and the supporting ERP data, van Wouwe et al. (2011) concluded that cue usage, and hence proactive control, did not differ between their positive and neutral group but that, instead, reactive control as soon as the target stimulus appeared was enhanced under positive affect. Considering these mixed results so far, one aim of the present study is to further clarify whether positive affect modulates proactive or reactive control. One obvious difference between both studies might be the specific mood induction procedure (namely, IAPS pictures vs. film clips). Related to that, it is conceivable that different arousal levels in the positive affect groups of both studies might account for the different results as arousal is an inherent and variable but often neglected feature of affect (Russell, 1980; Posner et al., 2005). Furthermore, arousal differences might help to explain the mixed results found in the literature so far. To our knowledge, there is only one study (Vogt et al., 2008) that shows that highly arousing affective stimuli increase the cue validity effect (CVE) of informative cues. This study, however, investigated attentional allocation to affective stimuli as it used these affective stimuli as cues in a spatial cueing task. Therefore, that study showed that highly arousing affective stimuli can attract and bind attention, but it could not answer whether or not affective arousal influences the reliance on neutral informative cues. Thus, another aim of the present study was to explore the role of arousal on positive affect effects, when affect is not confounded with the cues.

In sum, following the results of the previous study from our lab (Dreisbach, 2006) we wanted to show that positive affect with similar (low) arousal levels as used before reduces proactive control in form of a reduced usage of informative cues. This positive affect group (positive$_{\text{low}}$ hereafter) was compared to a neutral control group and another positive affect group with higher arousal levels (positive$_{\text{high}}$ hereafter). In Experiments 1 and 2, we used a spatial response cueing task with spatially congruent target response mappings. A bias in favor of a proactive control strategy was induced by using informative cues, that is, the probability of validly cued trials was more than 50% but less than 100%. In this response cueing task, a peripheral, informative cue indicated a possible target location and thereby primed the congruent response. Furthermore, the higher probability of valid cues (66%) should promote the usage of a proactive control strategy resulting in a reliable CVE, that is, faster responses and fewer errors in validly cued trials. A reduction of proactive control should consequently reduce the CVE, because less usage of the cues would minimize the benefits in validly cued trials as well as the costs in invalidly cued trials.

For more direct evidence that specifically proactive control and not reactive control is influenced by positive affect, Experiment 3 used a task switching paradigm. Comparing task switching performance with and without informative task cues enabled the investigation of affective influences on reactive control (as measured by switch costs) and proactive control (as measured by the CVE) in a single experiment.

EXPERIMENT 1

Following previous results (Compton et al., 2004; Dreisbach, 2006) we expected to find a decrease in proactive control in form of a reduced CVE in the positive$_{\text{low}}$ group as compared to the neutral group in the response cueing task with informative cues of Experiment 1. Because arousal differences were not considered in previous studies[1] or were confounded with the cues (Vogt et al., 2008) we had no *a priori* expectations concerning different outcomes in the positive$_{\text{low}}$ and positive$_{\text{high}}$ group.

METHOD

Participants

Sixty-six undergraduate students of Regensburg University participated in the experiment for course credit or 5 Euro. Sixty-two subjects (see Results for exclusion criteria) were included into the final data analysis (Mean age = 24.13 years, SD = 3.95, range = 20–38, 53 female). Participants were assigned randomly to the three affect groups (19 neutral, 21 positive$_{\text{low}}$, 22 positive$_{\text{high}}$). All participants signed informed consent and were debriefed after the session.

Apparatus and stimuli

A computer with a 17″-monitor (display resolution at 1024 × 768 pixel), running E-Prime 2.0 (Psychology Software Tools, Sharpsburg, USA) was used for experiment presentation and data acquisition. Viewing distance was held constant at 50 cm by using a chin rest. Responses were collected via a QWERTZ-keyboard, with the y- and m-key serving as left and right response keys.

To be able to manipulate valence and arousal independently we used pictures from the IAPS (Lang et al., 1999) as affect induction procedure. These pictures are known to reliably elicit specific affective reactions even with short presentation durations (Codispoti et al., 2009), and the elicited emotional reactions maintain and even sensitize – but do not habituate – with repetitive exposure to pictures of the same valence (Bradley et al., 1996; Smith et al., 2005). For each affect condition we chose 10 pictures: The neutral picture set had medium valence levels ($M = 4.99$), and low arousal levels ($M = 2.45$), whereas both positive picture sets were high in valence (M positive$_{\text{low}} = 7.99$; M positive$_{\text{high}} = 7.25$) but differed in arousal levels (M positive$_{\text{low}} = 4.55$; M positive$_{\text{high}} = 6.30$). Neutral pictures included household objects like plates or cups, positive$_{\text{low}}$ pictures showed babies and families, and in the positive$_{\text{high}}$ group sport and adventure pictures were displayed. It should be noted, that no erotica were used in the positive$_{\text{high}}$ group to prevent different gender influences, and because erotica seem to be a special category with effects differing from non-sexual positive, highly arousing pictures (Most et al., 2007). All pictures were presented in landscape format and color, adjusted to a size of 800 × 600 pixel, and positioned centered on a gray background.

The fixation cross, cue and target were all displayed in black ink and bold on gray background. The fixation cross was presented at the center of the screen in font size 32 pt. The target (a single dot) and the cue (the "§"-symbol) appeared 8.64 cm to the left or right of the fixation cross in font size 55 pt.

---

[1] Note that in Dreisbach (2006) negative IAPS pictures were chosen to match the arousal levels of the positive affect pictures.

Procedure

Each trial started with the presentation of the fixation cross for 500 ms, followed by an IAPS picture for 350 ms. After another short fixation period (200 ms) the cue was presented left or right of the fixation cross for 200 ms. The target appeared after a variable inter stimulus interval of 50 or 150 ms, which was included to reduce premature responses to the cue, and remained visible until the participant pressed the spatially congruent response key. Participants were instructed to react as fast as possible while avoiding errors. In case of an error, the German word for error ("Fehler") was presented for 1000 ms as feedback.

To assure that all participants started with a similar mood, all participants passed a 5-min relaxation exercise – comprised of relaxing music and spoken instructions for muscle relaxation – prior to the actual experiment. These instructions were standardized mp3-files presented via stereo headphones. Subsequently, 12 practice trials without IAPS pictures enabled the participants to get used to the cueing task. These practice trials were followed by two experimental blocks, in which an IAPS picture preceded every trial. Both blocks consisted of 120 trials (80 valid and 40 invalid), separated by a short break. The trial procedure within each block was pseudo-random: Each block consisted of 10 sequences of 12 trials and within these 12 trials the only constraint was that cues and targets appeared equally often on the left and the right side. Affective pictures were drawn from the set of the picture pool at random without replacement until all pictures had been presented once and then the procedure started all over again.

Design

A 3 (affect: neutral vs. positive$_{\text{low}}$ vs. positive$_{\text{high}}$) × 2 (Cue validity: valid vs. invalid) mixed factors design was used. Affect was manipulated between, and Cue validity varied within participants.

RESULTS

Data analysis

The practice trials as well as the first trial of each experimental block were excluded from analyses. In addition, error trials, trials following an error, and trials with reaction times (RT) below 150 ms or above 1500 ms were excluded (4.31% of the data). Furthermore, RTs differing more than 3 SD from individual means were considered as outliers and also removed prior analysis (1.21% of the trials). The data of two participants were excluded from further analyses, because of too many errors (individual mean error rates 11 and 14% while overall error rate was 2.23%). Another two subjects had to be excluded due to untypical RTs throughout the experiment. One was exceptionally slow ($M = 492$ ms) in comparison to mean RTs of his affect group (M positive$_{\text{low}} = 344$ ms), and the other participant got continuously slower throughout the experiment and also had high mean RTs ($M = 411$ ms, while M neutral $= 349$ ms). Of the remaining data, mean RTs and error rates of each design cell (see **Table 1**) were entered into a 3 (Affect: neutral vs. positive$_{\text{low}}$ vs. positive$_{\text{high}}$) × 2 (Cue validity: valid vs. invalid) mixed factors analysis of variance (ANOVA).[2]

Error data, overall analysis

The overall ANOVA for the error data brought up a main effect of Cue validity, $F(1, 59) = 90.35, p < 0.001, \eta_p^2 = 0.605$. Fewer errors were made in valid than invalid trials (0.17 vs. 4.27%). The main effect Affect, $F(2,59) = 2.68, p = 0.077, \eta_p^2 = 0.083$, , did not prove reliable. But we found a significant interaction of Affect × Cue validity, $F(2, 59) = 3.45$, $p < 0.05$, $\eta_p^2 = 0.105$. Planned comparisons showed a reduced CVE in the positive$_{\text{low}}$ group compared to the positive$_{\text{high}}$ group ($F = 6.73, p < 0.05$). The CVE in the neutral group was descriptively between both positive groups, but did not differ significantly from either group (Fs < 2.51, ps > 0.118). The overall error rate was 2.23% (SD = 1.85).

RT data, overall analysis

We found a significant main effect of Cue validity, $F(1, 59) = 88.86$, $p < 0.001$, $\eta_p^2 = 0.601$. Participants responded significantly faster after valid than after invalid trials (328 vs. 363 ms), resulting in an overall CVE of 35 ms. The main effect of affect as well as the interaction of Affect × Cue validity did not prove reliable (all $F < 2.08$, all $p > 0.133$). Even though we did not find a significant interaction of Affect × Cue validity in the RT analysis, the descriptive data resembles the results found in error rates (see **Figure 1**). CVE was smallest in the positive$_{\text{low}}$ group (25 ms), intermediate in the neutral group (35 ms), and largest in the positive$_{\text{high}}$ group (43 ms).

Table 1 | Mean RTs (in ms) and error rates (in %) in the spatial response cueing task of experiment 1 as a function of Affect group and Cue validity.

	Affect group					
	Neutral		**Positive$_{\text{low}}$**		**Positive$_{\text{high}}$**	
	Valid	**Invalid**	**Valid**	**Invalid**	**Valid**	**Invalid**
RT (SD)	332 (21.4)	367 (35.5)	332 (33.3)	357 (40.69)	320 (26.0)	363 (43.9)
Errors (SD)	0.24 (0.38)	4.11 (3.1)	0.19 (0.3)	3.05 (3.16)	0.09 (0.23)	5.64 (3.95)

---

[2] Although previous studies (Compton et al., 2004; Dreisbach, 2006) already indicated a valence specific effect, we still included two negative affect groups – one with low and one with high arousal – in Experiment 1 to disentangle valence and arousal effects on proactive control. It turned out that negative affect in combination with high arousal led to a general increase in RT compared to the neutral group and both positive groups (Fs > 4.13, ps < 0.05). Both negative groups showed greater cue validity effects than the positive$_{\text{low}}$ group (M negative$_{\text{high}} = 46$ ms, M negative$_{\text{low}} = 37$ ms, M positive$_{\text{low}} = 25$ ms). There was no significant interaction of negative affect with low or high arousal and cue validity. The negative groups therefore contributed no information concerning the special topic of anticipation (and the special topic of this Frontiers issue). To enhance the readability of the present article results of the negative groups are therefore not presented here but can be reported on request.

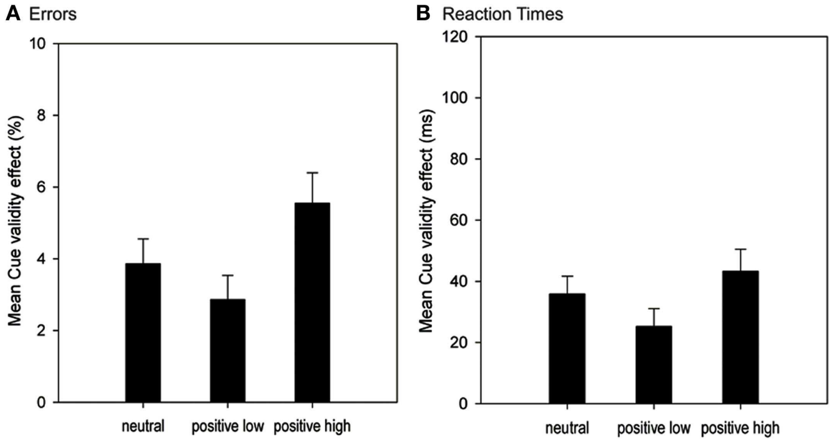

FIGURE 1 | Mean cue validity effects (CVE) in the spatial response cueing task of Experiment 1 as a function of Affect group. The **(A)** represents CVE differences in error rates (in %), the **(B)** represents CVE differences in RTs (in ms). Error bars represent 1 standard error of the mean.

Because the neutral group was more of a descriptive baseline – it differed on both valance and arousal levels from the positive groups – we conducted an additional analysis without the neutral group to search more directly for a possible arousal effect on proactive control.

Arousal effect, positive$_{low}$ vs. positive$_{high}$

A 2 (Arousal: positive$_{\text{low}}$ vs. positive$_{\text{high}}$) × 2 (Cue validity: valid vs. invalid) mixed factors ANOVA revealed a significant main effect of Cue validity, $F(1, 41) = 54.19$, $p < 0.001$, $\eta_p^2 = 0.569$. Participants responded faster after valid trials (326 vs. 360 ms), resulting in a CVE of 34 ms. The interaction of Arousal × Cue validity, $F(1,41) = 3.74$, $p = 0.059$, $\eta_p^2 = 0.084$, was on the threshold of significance. Therefore, we additionally calculated the JZS-Bayes factor (Rouder et al., 2009), which gives information about the probability of a hypothesis conditionally on observed data. JZS-Bayes factor was 0.895, which means that there is indeed some evidence in favor of a difference in CVEs between positive$_{\text{low}}$ and positive$_{\text{high}}$ group. The main effect of Arousal did not prove reliable ($F < 1$, $p = 0.787$).

DISCUSSION

Experiment 1 resulted in preliminary evidence for a reduction of proactive control under positive affect. The positive$_{\text{low}}$ group had the smallest CVE, an effect that was significant in the error data and just at the threshold of significance in the RT data. Interestingly, the CVE was increased in the positive$_{\text{high}}$ group indicating an increase of proactive control under positive affect with high arousal. However, there were only descriptive but no statistically significant differences between the neutral group and either positive group. Experiment 2 was run to collect more empirical support for the modulation of the CVE by positive affect with differing arousal levels.

EXPERIMENT 2

We used a very simple response cueing task in Experiment 1, which resulted in very fast overall RTs ($M = 345$ ms) and a low overall error rate (2.23%). Therefore, marginally significant differences between groups might be due to a floor effect. To increase variance and thereby provide room for affective modulations, we increased task difficulty in Experiment 2. To assure that both experiments were still comparable we used the same cueing task with informative cues (66% Cue validity) in combination with a concurrent math task. Based on the results of Experiment 1 we expected to find a reduced CVE in the positive$_{\text{low}}$ group, but an increased CVE in the positive$_{\text{high}}$ group, compared to the neutral group.

METHOD

Participants

Another 60 students of Regensburg University participated in the experiment for course credit or 5 Euro. Fifty-five subjects (see Results for exclusion criteria) were included into the final data analysis (Mean age = 22.86 years, SD = 3.79, range = 19–45, 40 female). Participants were assigned randomly to the three Affect groups (18 neutral, 19 positive$_{\text{low}}$, 18 positive$_{\text{high}}$). All participants signed informed consent and were debriefed after the session.

Apparatus and stimuli

Apparatus and stimuli were the same as in Experiment 1 except for the numbers presented in the math task. The numbers 1–5 were presented centrally, in black ink and in size 32 pt. Responses in the math task had to be typed in with the number keys of the first row of the keyboard.

Procedure

Procedure in Experiment 2 was the same as in Experiment 1 with the following exceptions: First, in each trial of the cueing task the first fixation was replaced by random numbers 1–5 for 800 ms. These numbers were part of the additional math task. Participants performed the cueing task, and at the same time had to add up the random numbers. Every 12 trials subjects were asked to type in the result of the summation task, which was followed by an informative feedback (3500 ms). Second, the actual experiment was preceded by a math test to assure that the Affect groups did not differ according to their calculating skills. To this end, we used a subtest of the Leistungsprüfsystem (a German IQ-test; L-P-S, Horn, 1983), which requires adding up lines of 10 random numbers from 2 to 9 under speeded conditions and is therefore similar to the actual experimental situation. And third, because of the increased task difficulty we added additional practice blocks. The first block comprised 12 trials of the spatial response cueing task. The next practice block (24 trials) introduced the math task in addition to the response cueing task. It included two complete math task cycles of 12 trials with feedback. In a final practice block (12 trials) an IAPS picture preceded every cueing trial. Data acquisition took part in the following three experimental blocks with 120 trials each (80 valid and 40 invalid trials, 10 math task cycles per block).

Design

A 3 (Affect: neutral vs. $\text{positive}_{\text{low}}$ vs. $\text{positive}_{\text{high}}$) × 2 (Cue validity: valid vs. invalid) mixed factors design was used. Affect was manipulated between, whereas Cue validity varied within participants.

RESULTS

Data analysis

We checked for group differences in calculating skills before the experiment and during the experiment with an ANOVA on performance in the L-P-S subtest as well as in the additional math task. For analysis of error rates and RTs in the cueing task, trials with math task responses differing more than two from the correct result were excluded from analysis (6.31% of the data).[3] Further preprocessing was the same as in Experiment 1, which resulted in the exclusion of another 6.83% of the trials. Furthermore one participant of the neutral group was excluded because he did not follow the instructions. Also two subjects of the $\text{positive}_{\text{low}}$ group had to be excluded. The first made too many errors in the math task (76.7%, while mean error rate was 14.7%), and the second made too many errors in the cueing task (14.8%, while mean error rate was 1.3%). Finally, two participants of the $\text{positive}_{\text{high}}$ group were excluded from further analysis, because they were exceptionally slow (715 and 894 ms, while mean RTs were 448 ms). Of the remaining data, mean RTs and error rates of each design cell (see **Table 2**) were entered in to a 3 (Affect: neutral vs. $\text{positive}_{\text{low}}$ vs. $\text{positive}_{\text{high}}$) × 2 (Cue validity: valid vs. invalid) mixed factors ANOVA.

Math performance

There were no differences in the performance in the L-P-S subtest between Affect groups before the experiment, $F(2, 52) = 2.62$, $p = 0.082$, $\eta_p^2 = 0.092$. Also, no significant differences between the three Affect groups were found in the additional math task during the experiment ($F < 1$, $p = 0.395$).

Error data, overall analysis

The overall error rate was 1.3% (SD = 1.5), and individual mean error rates were below 7.5% for all subjects. The overall ANOVA for the error data brought up a main effect of Cue validity, $F(1, 52) = 36.63$, $p < 0.001$, $\eta_p^2 = 0.413$, with fewer errors in valid than in invalid trials (0.10 vs. 2.51%). The main effect affect as well as the interaction of Affect × Cue validity did not prove reliable ($F\text{s} < 1.37$, $p\text{s} > 0.263$).

RT data, overall analysis

The ANOVA yielded a significant main effect of Cue validity, $F(1, 52) = 142.39$, $p < 0.001$, $\eta_p^2 = 0.732$. Participants responded significantly faster after valid than after invalid trials (418 vs. 478 ms), resulting in an overall CVE of 60 ms. More importantly, we found a significant interaction of Affect × Cue validity, $F(2, 52) = 3.51$, $p < 0.05$, $\eta_p^2 = 0.119$, which is depicted in **Figure 2**. Planned comparisons showed a reduced CVE in the $\text{positive}_{\text{low}}$ group (41 ms) as compared to the neutral group (72 ms; $F = 5.49$, $p < 0.05$) and the $\text{positive}_{\text{high}}$ group (70 ms; $F = 4.94$, $p < 0.05$). There was no significant difference between neutral group and $\text{positive}_{\text{high}}$ group ($F < 1$, $p = 0.904$). Also, the main effect of Affect was not significant ($F < 1$, $p = 0.578$).

DISCUSSION

An increase in mean RTs from Experiment 1 to 2 (345 vs. 437 ms) indicates that we succeeded in increasing task difficulty. With this adapted paradigm we found clear-cut evidence of a reduced CVE in the $\text{positive}_{\text{low}}$ group compared to the neutral and the $\text{positive}_{\text{high}}$ group. This suggests that specifically positive affect with low arousal reduces proactive control in form of a reduced reliance on informative cues. In contrast to a proactive control strategy, participants in the $\text{positive}_{\text{low}}$ group show behavioral costs in expected events (valid trials) and benefits in unexpected

Table 2 | Mean RTs (in ms) and error rates (in %) in the spatial response cueing task of experiment 2 as a function of Affect group and Cue validity.

	Affect group					
	Neutral		$\text{Positive}_{\text{low}}$		$\text{Positive}_{\text{high}}$	
	Valid	**Invalid**	**Valid**	**Invalid**	**Valid**	**Invalid**
RT (SD)	405 (77.6)	477 (105.8)	445 (111.9)	487 (111.2)	401 (60.3)	471 (90.7)
Errors (SD)	0.21 (0.33)	3.32 (3.61)	0.09 (0.18)	1.86 (2.38)	0.0 (0.0)	2.35 (2.84)

---

[3] We wanted to be sure that participants were truly engaged with both the response cueing task and the additional math task. Therefore, we controlled for performance in the math task. But we chose a rather moderate criterion (correct response ±2) to minimize data loss, because exclusion due to math task performance meant to exclude a complete cycle of 12 cueing trials. With an absolute criterion (only correct responses included) 17.7% of all trials would have been excluded. By using the moderate criterion, we aimed to include all trials where participants genuinely tried to follow instructions.

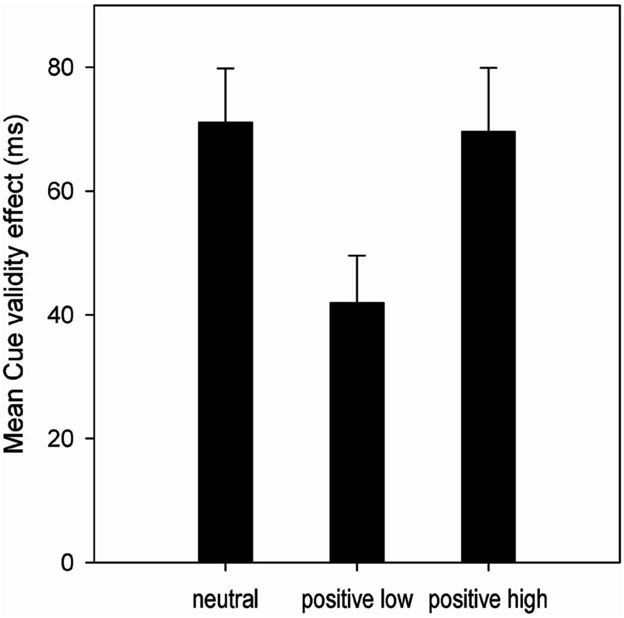

FIGURE 2 | Mean Cue validity effects (in ms) in the spatial response cueing task of Experiment 2 as a function of Affect group. Error bars represent 1 standard error of the mean.

events (invalid trials). A problem in our simple cueing paradigm is, however, that we cannot completely rule out that the reduced CVE might also be a sign of increased *reactive* control: participants in the positive$_{\text{low}}$ group might have used the cues just as the other affect groups but they might have been better able to overcome the pre-activated response in invalidly cued trials. This would be in line with the results by van Wouwe et al. (2011) that showed an enhancement in reactive control but no influence of positive affect on proactive control. To rule out this alternative explanation, we conducted an additional control experiment, using again a response cueing paradigm but this time employing non-informative cues.[4] With this modification, participants could not optimize their performance with a proactive control strategy. Again, we found a significant CVE, $F(1, 56) = 474.0$, $p < 0.001$, $\eta_p^2 = 0.894$, but *no* affective modulation thereof ($F = 1.13$, $p = 0.330$). JZS-Bayes factors for CVE comparisons between the affect groups (neutral, positive$_{\text{low}}$, and positive$_{\text{high}}$) ranged from 1.66 to 3.75, which means the null hypothesis – no difference in CVE – was indeed more likely. So, in sum the results of the response cueing experiments speak in favor of an affective modulation of proactive control only, with a reduced reliance on informative cues under positive affect with low arousal. However, it would be even better proof, if we could show that the affective modulation of the CVE is restricted to proactive control and is not present for reactive control in a unique experiment. Therefore, we conducted Experiment 3.

EXPERIMENT 3

The main aim of Experiment 3 was to gather more direct evidence that specifically proactive control and not reactive control is influenced by positive affect. Furthermore, we wanted to know, whether the affective modulation of proactive control can also be found for task cues (instead of response cues, as was the case with the response cueing paradigm used here and the AX-CPT in previous studies). To address these issues we employed a task switching paradigm. Task switching (for recent reviews, see, e.g., Kiesel et al., 2010; Vandierendonck et al., 2010) with univalent stimuli (e.g., digits and letters) is well suited to investigate reactive control in form of differences in switch costs. Using univalent stimuli (a given stimulus is only associated with one of the two possible tasks) and no precues, variations in switch costs can be taken as a direct indicator for reactive control processes. Furthermore, it has been shown that participants are generally very sensitive to probability cues (i.e., informative, but not 100% valid) in task switching (Dreisbach et al., 2002; Hübner et al., 2004; Miniussi et al., 2005; Dreisbach and Haider, 2006; Wendt et al., under review). Therefore, a cued task switching paradigm with valid and invalid cues allows not only the investigation of reactive control but also proactive control in form of differences in the CVE (like in Experiments 1 and 2). Thus, in Experiment 3 we used a task switching paradigm with a digit and a letter task that started without task cues. After the first experimental block without precues, informative task cues with a Cue validity of 75% preceded each trial. If positive affect with low arousal reduces proactive control – as Experiments 1 and 2 suggest – we should again find a reduced CVE. If positive affect, however, increases reactive control we should find a reduction of switch costs – especially so in blocks without precues.

METHOD

Participants

Sixty undergraduate students from the Regensburg University (age $M = 22.53$ years, $SD = 4.02$, range = 18–36, 53 female) participated in the experiment for course credit or 5 Euro. Participants were assigned randomly to the three affect groups (20 positive$_{\text{low}}$, 20 positive$_{\text{high}}$, 20 neutral). All participants signed informed consent and were debriefed after the session. Because we were interested in a possible modulation of the switch costs, participants with negative switch costs were excluded and replaced (two in the neutral, three in the positive$_{\text{low}}$, and two in the positive$_{\text{high}}$ group).

Apparatus and stimuli

Apparatus was the same as in Experiments 1 and 2. Also the same IAPS picture sets were used for the three affect groups.

Eight digits (1, 2, 3, 4, 6, 7, 8, and 9) written in green and eight letters (A, E, O, U, C, K, G, and T) written in purple served as target

---

[4] We tested 59 subjects (20 neutral, 20 positive$_{\text{low}}$, and 19 positive$_{\text{high}}$) in the cueing experiment with non-informative cues (50% validity). Stimuli and procedure were basically the same as in Experiment 1 except for the following changes: The fixation cross as well as the two possible target locations were always enclosed by black bordered boxes (size 100 × 100 pixels). As a cue the border of one peripheral box changed border width from 1 to 5 pt. The stimulus onset asynchrony between cue and target was changed to 50 or 150 ms, which was shorter than in Experiment 1 but equaled the inter stimulus interval of Experiment 1. Furthermore, the cue remained on screen with the target until the participant pressed a response key, and the cue validity percentage was reduced to 50%, which resulted in 60 valid and 60 invalid trials per block. All these changes were made to induce a bias in favor of a reactive control strategy. Mean error rate was 1.26% and mean RTs were 339 ms. Mean CVE were comparable in all affect groups with 40 ms in the neutral group, 43 ms in the positive$_{\text{low}}$ group, and 35 ms in the positive$_{\text{high}}$ group.

stimuli and were presented at the center of the screen in font size 52. The color coding of the digit and letter task was counter balanced across participants. Odd numbers and vowels were always assigned to one response key, even numbers and consonants to the other, while response mapping to the left and right response key (y- and m-key on a QWERTZ-keyboard) was also counterbalanced between participants. In experimental blocks 2–4, a color coded fixation cross (purple or green) served as informative task cue.

Procedure

The experiment comprised one task switching block without task cues followed by three blocks including informative task cues. In the first block each trial started with an IAPS picture (350 ms) followed by a blank screen (150 ms) and a black fixation cross (1000 ms). Then the target stimulus appeared and remained on screen until the participant responded. Subjects had to decide whether a number was odd or even (digit task) or whether a letter was a vowel or consonant (letter task). Participants were instructed to react as fast as possible while avoiding errors. Feedback was given for errors only (2000 ms), each trial ended with an intertrial interval of 500 ms. Procedure in the following blocks with informative task cues was the same as in the first block except that the fixation cross was now color coded and served as a task cue for the following task. In valid trials (75% of all trials) the colored fixation cross was followed by a target stimulus in the same color, thereby enabling the preparation of the upcoming task in a proactive manner. In contrast, in invalid trials (25% of all trials) the fixation color incorrectly predicted the upcoming target color, and can therefore mislead to prepare the wrong task.

The experiment started with the same relaxation exercise that was used before in Experiments 1 and 2. Subsequently, 16 practice trials (random presentation of all target stimuli) without IAPS pictures enabled the participants to get used to the task switching procedure. This practice block was followed by 64 trials with an IAPS picture preceding every trial. Data acquisition took place in the following four experimental blocks – the first without informative task cues – with 128 trials each. Each block contained 64 digit tasks (4 × 8 numbers) and 64 letter tasks (4 × 8 letters). Stimulus presentation was pseudo-randomized with the following constraints: repeat and switch trials were evenly distributed. Immediate repetitions of target stimuli or IAPS pictures were not allowed. Task cues (96 valid, 32 invalid) were counterbalanced across all trial types.

Design

A 3 (Affect: neutral vs. $\text{positive}_{\text{low}}$ vs. $\text{positive}_{\text{high}}$) × 2 (Trial type: repeat vs. switch) design with affect as between and Trial type as within factor was used in the first block without task cues. The experimental blocks including informative task cues had a 3 (Affect) × 3 (Block: 2 vs. 3 vs. 4) × 2 (Trial type) × 2 (Cue validity: valid vs. invalid) repeated measures design.

RESULTS

Data analysis

Practice trials as well as the first trial of each experimental block were excluded from analyses. In addition, error trials, trials following an error, and trials with RTs differing more than 3 SD from individual means were also removed prior analysis (9.34% of all trials). Separate analyses were conducted for task switching performance (mean error rates and RTs) in the first experimental block without task cues and for performance in experimental blocks 2–4 with informative task cues.

Task switching performance, block 1 without task cues

Mean RTs (see **Table 3**) were entered into a 3 (Affect: neutral vs. $\text{positive}_{\text{low}}$ vs. $\text{positive}_{\text{high}}$) × 2 (Trial type: repeat vs. switch) mixed factors ANOVA. We found a significant main effect of Trial type, $F(1, 57) = 106.45$, $p < 0.001$, $\eta_p^2 = 0.651$, with faster responses in repeat trials (655 vs. 733 ms). The main effect of Affect as well as the interaction of Affect × Trial type did not prove reliable (all $F < 1.97$, all $p > 0.150$). The same analysis for mean error rates (see **Table 3**) also resulted in a significant main effect of Trial type, $F(1, 57) = 26.82$, $p < 0.001$, $\eta_p^2 = 0.319$, with less errors in repeat trials (2.28 vs. 5.97%). Again, no significant Affect effects were found (all $F < 1.19$, all $p > 0.31$). JZS-Bayes factors for differences in switch costs between the Affect groups ranged from 2.95 to 4.04, which means that it is more likely that there are indeed equal switch costs in all three groups.

Task switching performance, blocks 2–4 with informative task cues

To check the effectiveness of the cues over time, we conducted a 3 (Affect: neutral vs. $\text{positive}_{\text{low}}$ vs. $\text{positive}_{\text{high}}$) × 3 (Block: 2 vs. 3 vs. 4) × 2 (Trial type: repeat vs. switch) × 2 (Cue validity: valid vs. invalid) mixed factors ANOVA for the three experimental blocks with informative task cues (see **Tables 4** and **5** for mean RTs and error rates). The analysis of mean error rates resulted in significant main effects of Block, $F(2,114) = 8.65$, $p < 0.001$, $\eta_p^2 = 0.072$, Trial type, $F(1, 57) = 37.34$, $p < 0.001$, $\eta_p^2 = 0.397$, and Cue validity, $F(1, 57) = 4.40$, $p < 0.05$, $\eta_p^2 = 0.072$, as well as an interaction of Trial type × Cue validity, $F(1, 57) = 4.19$, $p < 0.05$, $\eta_p^2 = 0.069$. Planned comparisons showed significantly more errors in Block 2 (3.61%) as compared to Block 3 (2,81%, $F(1, 57) = 7.97$, $p < 0.01$) and Block 4 (2.42%, $F(1, 57) = 14.20$, $p < 0.001$). Blocks 3 and 4 did not differ significantly ($F = 2.01$, $p = 0.162$). Cue validity had no significant influence an error rates in task repetitions (2.09 vs. 2.12%, $F < 1$, $p = 0.915$), but there was a significant negative CVE in task switches ($F(1,57) = 6.41$, $p < 0.05$) with more errors in valid trials (4.30 vs. 3.27%). The interaction of Block and Trial type did not prove reliable ($F = 2.82$, $p = 0.064$). There was no significant main effect of Affect or significant interactions with affect (all $F < 1.68$, all $p > 0.185$). In the RT analysis we found significant main effects for Block, $F(2, 114) = 19.83$, $p < 0.001$, $\eta_p^2 = 0.258$, Trial type, $F(1, 57) = 98.88$, $p < 0.001$, $\eta_p^2 = 0.634$, and Cue validity, $F(1,57) = 19.53$, $p < 0.001$, $\eta_p^2 = 0.255$, which were further qualified by a significant three-way interaction of these factors, $F(2, 114) = 11.28$, $p < 0.001$, $\eta_p^2 = 0.165$. Planned comparisons showed a significant interaction of Trial type × Cue validity specifically in the first block with informative task cues, $F(1, 57) = 2.54$, $p < 0.001$ (Blocks 3 and 4: all $F < 0.07$, all $p > 0.41$). Further analysis of Block 2 showed a significant CVE with faster RTs after valid cues in repeat trials (590 vs. 644 ms, $F(1, 57) = 32.28$, $p < 0.001$), but not in switch trials (667 vs. 659 ms, $F = 1.32$,

Table 3 | Mean RTs (in ms) and error rates (in %) in the in the first experimental block of experiment 3 (task switching without task cues) as a function of Affect group and Trial type.

	Affect group					
	Neutral		**Positive$_{\text{low}}$**		**Positive$_{\text{high}}$**	
	Repeat	**Switch**	**Repeat**	**Switch**	**Repeat**	**Switch**
RT (SD)	646 (76.9)	731 (116.7)	705 (170.5)	774 (202.0)	615 (96.9)	693 (133.5)
Errors (SD)	2.7 (2.89)	5.89 (4.68)	1.52 (2.09)	4.83 (3.72)	2.64 (2.73)	7.18 (8.19)

Table 4 | Mean RTs (in ms, SD in parentheses) in experimental blocks 2–4 of Experiment 3 (task switching with informative task cues) as a function of Affect group, Trial type, and Cue validity.

Cue	**Affect group**					
	Neutral		**Positive$_{\text{low}}$**		**Positive$_{\text{high}}$**	
	Repeat	**Switch**	**Repeat**	**Switch**	**Repeat**	**Switch**
BLOCK 2						
Valid	568 (81.1)	639 (133.9)	588 (61.1)	661 (83.4)	615 (119.9)	702 (142.4)
Invalid	613 (111.9)	626 (92.3)	617 (103.2)	665 (92.9)	702 (176.9)	685 (125.63)
BLOCK 3						
Valid	561 (73.4)	617 (110.2)	590 (91.9)	648 (105.5)	600 (99.8)	665 (141.0)
Invalid	558 (82.7)	643 (151.9)	600 (104.3)	654 (125.5)	603 (123.3)	656 (124.0)
BLOCK 4						
Valid	557 (86.8)	595 (108.1)	566 (71.4)	608 (140.0)	591 (102.5)	635 (133.8)
Invalid	567 (109.2)	631 (151.1)	579 (81.2)	645 (140.0)	602 (115.6)	631 (126.9)

Table 5 | Mean error rates (in %, SD in parentheses) in experimental blocks 2–4 of experiment 3 (task switching with informative task cues) as a function of Affect group, Trial type, and Cue validity.

Cue	**Affect group**					
	Neutral		**Positive$_{\text{low}}$**		**Positive$_{\text{high}}$**	
	Repeat	**Switch**	**Repeat**	**Switch**	**Repeat**	**Switch**
BLOCK 2						
Valid	2.76 (2.3)	5.55 (4.7)	2.27 (2.0)	4.89 (4.6)	1.87 (2.1)	4.88 (4.0)
Invalid	2.29 (3.7)	5.67 (4.8)	2.5 (4.8)	3.17 (4.6)	2.5 (5.5)	5.01 (5.3)
BLOCK 3						
Valid	2.27 (2.8)	4.62 (5.6)	2.39 (3.4)	4.44 (4.6)	1.67 (2.8)	4.15 (3.9)
Invalid	1.91 (3.2)	2.96 (4.2)	2.15 (3.3)	2.28 (4.0)	2.15 (3.9)	2.73 (5.2)
BLOCK 4						
Valid	1.7 (2.3)	4.2 (3.8)	2.39 (2.7)	2.8 (2.5)	1.52 (2.0)	3.2 (3.3)
Invalid	2.95 (3.6)	1.96 (3.2)	0.59 (2.6)	2.49 (4.2)	2.06 (4.4)	3.21 (4.9)

$p = 0.26$). So, there was a strong cueing effect only in the first block with informative task cues, and specifically in repeat trials. The main effect of Affect as well as all other interactions did not prove reliable (all $F < 3.36$, all $p > 0.067$). With respect to our hypotheses, also in these blocks with informative task cues the Affect groups did not differ significantly in switch costs (M neutral $= 54$ ms, M positive$_{\text{low}} = 57$ ms, M positive$_{\text{high}} = 44$ ms). JZS-Bayes factors for single comparisons of switch costs ranged from 2.46 to 4.24, which further supports that switch costs were indeed comparable in all three groups. Since we were interested in the affective modulation of the CVE, we reran the analysis, this time only including Block 2 (i.e., the first block with informative task cues), the only block where the CVE was significant.

Affect effects, first task switching block with informative task cues only

A 3 (Affect: neutral vs. $\text{positive}_{\text{low}}$ vs. $\text{positive}_{\text{high}}$) × 2 (Trial type: repeat vs. switch) × 2 (Cue validity: valid vs. invalid) mixed factors ANOVA revealed significant main effects for Trial type, $F(1, 57) = 39.46$, $p < 0.001$, $\eta_p^2 = 0.409$, and Cue validity, $F(1, 57) = 18.07$, $p < 0.001$, $\eta_p^2 = 0.241$. Participants responded faster in repeat trials (617 vs. 663 ms) as well as in valid trials (629 vs. 651 ms). Furthermore, we found a significant interaction of Trial type × Cue validity, $F(1, 57) = 22.54$, $p < 0.001$, $\eta_p^2 = 0.283$. Planned comparisons showed a significant CVE in repeat trials (590 vs. 644 ms, $F(1, 57) = 32.28$, $p < 0.001$), but not in switch trials (667 vs. 659 ms, $F = 1.32$, $p = 0.26$). Most important with respect to our hypothesis, there was a significant interaction of Affect × Trial type × Cue validity, $F(2, 57) = 3.08$, $p = 0.05$, $\eta_p^2 = 0.098$, which is depicted in **Figure 3**. CVE was significantly smaller in the $\text{positive}_{\text{low}}$ compared to the $\text{positive}_{\text{high}}$ group (29 vs. 87 ms, $F(1,57) = 6.32$, $p < 0.05$). The CVE in the neutral group (45 ms) was descriptively between both positive groups but did not differ significantly from either group ($F\text{s} < 3.35$, $p\text{s} > 0.072$). The main effect Affect and all other interactions did not prove reliable (all $F < 1.94$, all $p > 0.15$). The same analysis for mean error rates resulted only in a significant main effect of Trial type, $F(1, 57) = 25.06$, $p < 0.001$, $\eta_p^2 = 0.306$, with less errors in repeat trials (2.36 vs. 4.86%). No further significant main effects or interactions were found (all $F < 1$, all $p > 0.47$).

DISCUSSION

In Experiment 3 switch costs did not differ between affect groups, neither in the first experimental block without task cues nor in the following blocks with informative cues. Strong cueing effects were found only in the first block with informative task cues and specifically in repeat trials. In this block we also found an affect effect similar to the results of Experiments 1 and 2: the CVE in repeat trials was reduced in the $\text{positive}_{\text{low}}$ group as compared to the positive high group, while the CVE was descriptively in between both positive groups in the neutral group. It is not surprising that we found an affective modulation only in Block 2, because block wise analysis of all three blocks including cues showed that the informative task cues only had an impact on performance while they were new, whereas their influence diminished with more practice in the task (RTs and error rates declined throughout the experiment, see **Tables 4** and **5**). The generally reduced reliance on cues over blocks might be due to the fact that the task cues were neither necessary (because univalent stimuli were used) nor entirely useful (e.g., Sudevan and Taylor, 1987). The fact that the CVE is restricted to repeat trials only was also found by Miniussi et al. (2005), and might be a consequence of anticipatory backward inhibition (Mayr and Keele, 2000; Hübner et al., 2003; Li and Dupuis, 2008): in task switching, backward inhibition refers to the phenomenon that preparation for a task switch leads to inhibition of the just executed task set, and is hence also a form of proactive control. There is plenty of evidence that the foreknowledge about an upcoming task switch suffices to trigger the inhibition of the preceding task (Mayr and Keele, 2000, Experiment 5; Hübner et al., 2003; Li and Dupuis, 2008; Wendt et al., under review). Applied to our data, an invalidly cued repetition already caused inhibition of the previous task resulting in performance costs when this very task unexpectedly repeats. In invalidly cued switches, on the other hand, the cue predicts a repetition and as such does not trigger backward inhibition resulting in typical switch costs – like in validly cued switches. In sum, Experiment 3 succeeded in showing that specifically proactive control and not reactive control is modulated by positive affect: Switch costs – as a measure of reactive control – were comparable in all three affect groups in the first block without task cues. Positive affect along with high or

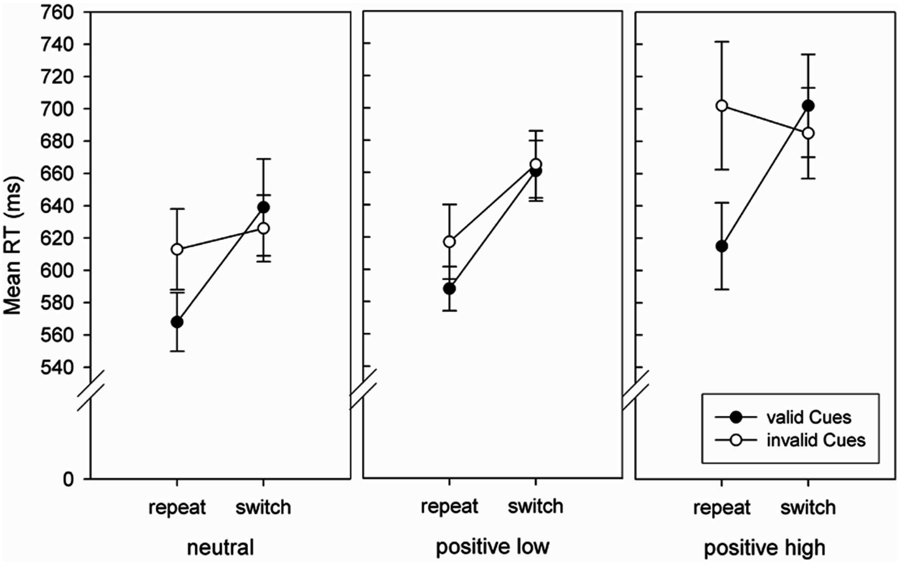

FIGURE 3 | Mean RTs (in ms) in the first task switching block with informative task cues of Experiment 3 as a function of Affect group, Trial type, and Cue validity. Error bars represent 1 standard error of the mean.

low arousal did neither improve nor impair the adaption to a (unexpected) task switch. In contrast, the CVE – as a measure of proactive control – was again modulated by affect, and indicated a reduction of proactive control in the positive$_{\text{low}}$ group.

Together with results from Experiments 1 and 2, we thus found converging evidence that performance under positive affect with low arousal is less dependent on informative cues, indicating a reduction in proactive control. Positive affect with high arousal, on the other hand, seems to increase the usage of informative cues.

GENERAL DISCUSSION

Purpose of the present study was to investigate the influence of positive affect on processes of proactive control under different arousal conditions. According to the DMC framework (Braver et al., 2007; Braver, 2012) cognitive control can be divided into proactive and reactive control: proactive control means sustained preparation for an upcoming event – for example, by using informative cues to optimize performance, while reactive control means a just-in-time activation of control as soon as a demanding event appears. In three experiments with different paradigms and kinds of informative cues, we found converging evidence that positive affect with low arousal – induced via short presentation of affective pictures – reduces the CVE. These results replicate and extend previous findings (Compton et al., 2004; Dreisbach, 2006) by showing that only positive affect with low arousal but not positive affect with high arousal reduces the usage of informative cues, and by showing that these effects are not limited to response cues but can be generalized to task cues (for limitations see Discussion of Experiment 3). Furthermore, results on task switching performance in Experiment 3 strongly suggest that positive affect does *not* modulate reactive control (see also Discussion of Experiment 2 and Footnote 4) by showing that switch costs were not manipulated by affect. Taken together, the results of Experiments 1–3 support the assumption that specifically positive affect with low arousal leads to a reduction in proactive control.

In Experiment 3, we found comparable switch costs in both positive groups and the neutral group suggesting that reactive control was not modulated by affect. At first sight, this seems to be at odds with findings by Dreisbach and Goschke (2004), who found interactions of switch costs and positive affect. Their study, however, did not use a classical task switching paradigm but a cognitive set-switching paradigm. In this paradigm, participants did not have to switch between different tasks, but performed a single task only: they had to categorize a target presented in one color, while ignoring a simultaneously presented distractor in another color. Dreisbach and Goschke investigated two switching conditions of cognitive sets: After the switch, either the targets appeared in a new color, while the former target color become the distractor color (perseveration condition), or the distractors appeared in a new color, while the former distractor color became the target color (learned irrelevance condition). Positive affect diminished switch costs when switching to a new cognitive set (perseveration condition), but increased switch costs and interference by distractors in the learned irrelevance condition. Dreisbach and Goschke interpreted these very specific interactions between positive affect and switch costs as evidence for increased cognitive flexibility accompanied by costs of increased distractibility under positive affect. The task switches in Experiment 3 of this study, however, can not be differentiated by these two switching conditions. Therefore, it is no surprise that no affective modulation of switch costs was found here. However, the reduced CVE might just as well be interpreted as an index of increased cognitive flexibility. For example, Compton et al. (2004) argued that a reduced CVE can be interpreted in terms of more flexibility because the behavior is less dependent on the cue information.

The CVE was significantly smaller in the positive$_{\text{low}}$ compared to the positive$_{\text{high}}$ group in both the response cueing (Experiments 1 and 2) as well as the cued task switching paradigm (Experiment 3), while the CVE in the neutral group was roughly between both positive groups (see **Figure 4**). But in spite of these descriptive differences between the neutral group and both positive groups, there was only once – in Experiment 2 – also a significant reduction of the CVE in the positive$_{\text{low}}$ group compared to the neutral group (while the magnitude of the CVE was equally high in the neutral and the positive$_{\text{high}}$ group). This lack of significant differences might be a byproduct of our procedure: each experiment started with a short relaxation exercise to create a similar baseline mood in all participants. This procedure, however, might already have resulted in a mild positive affect induction, thereby possibly reducing the differences between the neutral group and the positive group especially with low arousal. Admittedly, what speaks against this assumption is that in Experiment 2, the CVE of the neutral group actually resembled the positive$_{\text{high}}$ group. It is, however, conceivable that the higher task demands due to the additional math task have counteracted the relaxed mood in the neutral group. Thus, the significant difference found in Experiment 2 might in fact be closer to the actual difference between neutral affect and positive affect with low arousal. Also, it can be assumed that everyday mood is generally rather mildly positive than truly neutral. Therefore, it might not be too surprising that differences between mild positive affect and neutral affect are not easily detected. However, with these constraints in mind, the observed differences in the CVE between positive affect with low arousal and neutral and positive affect with high arousal provide sufficient evidence for the conclusion that positive affect with low arousal decreases proactive control, while positive affect with high arousal seems to increase proactive control compared to neutral affect.

The reduced CVE in the positive$_{\text{low}}$ group converges with findings from previous studies by Compton et al. (2004) and Dreisbach (2006). Compton et al. (2004) investigated associations between baseline mood state – assessed via the Profile of Mood States (McNair et al., 1971) – and performance in an attentional orienting task with informative cues. Self-reported negative affect was unrelated to attentional orienting performance. High positive affect, however, was associated with a reduced CVE, with slower responses after validly cued targets and faster responses following invalidly cued targets, compared to low positive affect. Dreisbach (2006) used the AX-CPT and found enhanced performance in AY trials, that is, in invalidly cued trials, but impaired performance in BX and BY trials, that is, in validly cued trials, under positive affect as compared to neutral or negative affect. Thus in both studies, positive affect resulted in a benefit in expected events, but also in costs in unexpected events. These findings – like our

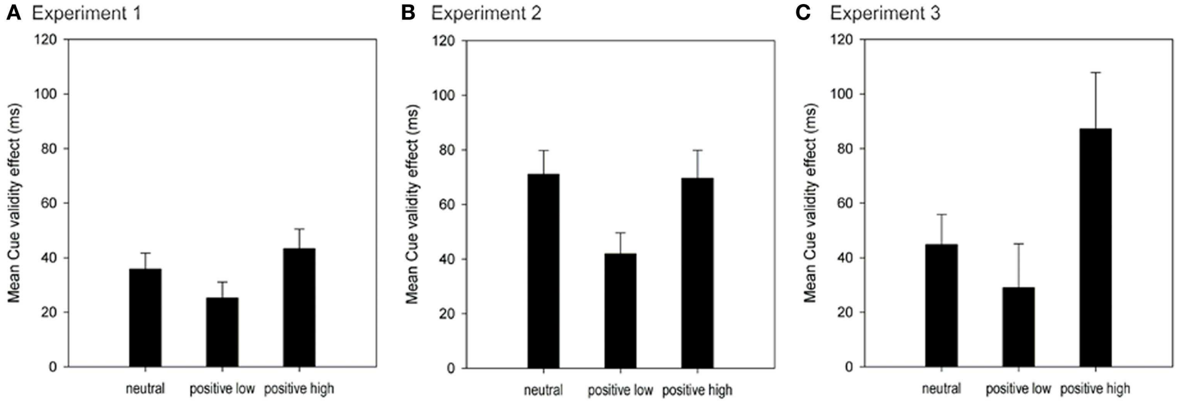

FIGURE 4 | Mean Cue validity effects (in ms) as a function of Affect group for Experiment 1 (A), Experiment 2 (B), and Experiment 3 (C). Error bars represent 1 standard error of the mean.

results – can be explained by a reduced usage of informative cues, which indicates a reduction in proactive control. But unlike Dreisbach (2006) a recent study by van Wouwe et al. (2011) – also using the AX-CPT – found no influence of positive affect on cue usage (no impairment in BX and BY trials), and hence proactive control, but, instead, differences between their positive and neutral group in reactive control: participants in the positive affect group showed a performance benefit and ERP differences in AY trials only, where a pre-dominant response tendency has to be overcome. In line with these results are also several studies by Kuhl and colleagues (Kuhl and Kazen, 1999; Baumann and Kuhl, 2005; Kazén and Kuhl, 2005) that used paradigms without informative cues, which means that there is not much room for proactive control. They used the Stroop task and a global-local task and found a reduction in Stroop interference and a reduced global precedence under positive affect (again no consideration of arousal differences) indicating also an enhanced ability to overcome pre-dominant response tendencies. So overall, there is evidence for increased flexibility in form of a reduction in *proactive* control (this study; Compton et al., 2004; Dreisbach, 2006), but also evidence for increased flexibility in form of a modulation of *reactive* control (Kuhl and Kazen, 1999; Baumann and Kuhl, 2005; Kazén and Kuhl, 2005; van Wouwe et al., 2011). One reason for these mixed results might be the differential affect induction procedures: the current study – like the AX-CPT study by Dreisbach – manipulated affect in a between groups design with affective pictures preceding every trial, Compton et al. investigated differences in baseline mood state, van Wouwe et al. used emotional film clips previous to the actual experiment (for a more detailed discussion on differences between the two AX-CPT studies see van Wouwe et al., 2011), and Kuhl and colleagues used a within design with random presentation of positive, negative, or neutral prime words preceding every trial. So, Compton et al. as well as van Wouwe et al. were concerned with effects of a sustained mood state – in the former case the currently existing mood state, in the latter case an induced mood state – whereas Kuhl and colleagues investigated influences of rather transient affective reactions. The affect induction procedure used in our lab (this study; Dreisbach, 2006) – affective pictures preceding every trial in a between groups design – most likely resulted in both transient and sustained affective reactions. IAPS pictures very quickly elicit typical emotional reactions with changes in cortical, autonomic, and facial activity, as well as evaluative ratings even with short presentation durations (Codispoti et al., 2001, 2009). Furthermore, repetitive exposure to pictures of the same valence leads to maintained or even sensitized affective reactions and can therefore be seen as a mood induction procedure (Bradley et al., 1996; Smith et al., 2005). For the studies reviewed here, however, the difference between sustained mood states vs. transient affective reactions does not seem to be a crucial factor to explain the different outcomes in the affective modulation of cognitive control. For example, also van Steenbergen et al. (2009) found consistent positive affect effects on the sequential modulation of response conflicts using either randomized affective signals between trials (smilies) or specific mood induction in a between groups design (van Steenbergen et al., 2010). In fact, there are other procedural factors aside from different affect induction procedures that might as well be crucial. For example, the reduced Stroop interference found by Kuhl and colleagues was restricted to conditions when intention memory is activated, that is, in the first of two consecutive Stroop tasks in a single trial (Kuhl and Kazen, 1999) or when using specific positive primes related to achievement (Kazén and Kuhl, 2005). Also, none of the above-quoted studies considered differences in arousal levels. But note that in the Dreisbach (2006) study the positive IAPS pictures had low arousal levels comparable to the ones used here. In sum, the existing literature is characterized by mixed results, which might be explained to some extent by different affect induction procedures – pictures vs. film clips vs. words, between vs. within –, differences in intention memory load, as well as different arousal levels. Therefore, future studies are clearly needed to further clarify under which conditions positive affect influences proactive or reactive control.

The fact, that we found a reliable difference in the CVE between the positive groups with low and high arousal, demonstrates that it is most important to consider both dimensions of affect – valence and arousal (cf., Russell, 1980; Posner et al., 2005). Whether there is less attention to the cues or a reduced maintenance of the cue information in the $\text{positive}_{\text{low}}$ group cannot be answered based on behavioral results alone. But nonetheless, it remains an interesting question why positive affect in combination with low arousal reduces proactive control, whereas positive affect along with high arousal seems to increase proactive control. Reduced proactive control under positive affect with low arousal seems to converge with our everyday experience: When being in a relaxed, mildly positive mood one tends to enjoy the moment without looking ahead. This would also be in line with Carver's (2003) coasting theory. This theory assumes a feedback function of affect: more precisely, positive affect signals better progress than necessary, and consequently reduces the effort invested in the ongoing task (=coasting). Proactive control in this sense is associated with more effort than reactive control, because it involves sustained maintenance of informative cues or task goals for an optimized behavior (Braver et al., 2007; Braver, 2012). Thus, a reduction of proactive control could be a sign of coasting: Participants in the $\text{positive}_{\text{low}}$ group apply less effort in sustained task preparation, and instead rely on reactive control alone as soon as the target appears. This might also explain why the effects of reduced proactive control were restricted to the positive affect group with low arousal and were not found with high arousal. Obviously, coasting might not be a reasonable strategy under high arousal as any high arousal signal might rather serve as a warning or alertness signal. For example, Fuentes and Campoy (2008) showed in an attention network task that alerting tones increase the CVE, and inferred that alerting enhances the effect of informative cues. A similar explanation presents the integrative theory of locus coeruleus-norephinephrine function (LC-NE) by Aston-Jones and Cohen (2005). Arousal is associated with NE activity, and according to the integrative LC-NE theory specifically phasic LC-NE activity promotes exploitative behavior that helps to optimize task performance. Applied to our data, the short presentation of highly arousing positive pictures might have triggered phasic NE activity and thereby resulted in increased proactive control in form of a stronger usage of the informative cue, and, as a consequence, an increased CVE.[5]

CONCLUSION

The DMC framework (Braver et al., 2007; Braver, 2012) assumes that there are various factors that induce a bias in favor of one type of control strategy over the other. Taken together, Experiments 1–3 resulted in converging evidence that positive affect is such a factor. Specifically, positive affect with low arousal led to a reduction in proactive control in form of a reduced reliance on informative cues. On the other hand, positive affect in combination with high arousal increased the CVE and therefore seems to promote proactive control. Reactive control, in contrast, was not influenced by positive affect.

ACKNOWLEDGMENT

The publication fee for this article is funded by Deutsche Forschungsgemeinschaft DFG.

---

[5] We are aware of the fact that in the NE-LC model (Aston-Jones and Cohen, 2005) the relation between phasic NE activity and exploitative behavior vs. tonic NE activity and explorative behavior is correlational in nature and based on animal studies. That is, even if we assume that the short presentation of highly arousing picture stimuli triggers NE activity, it is still an open question, whether the activation of phasic NE activity triggers exploitative behavior in humans. Actually, the first pharmacological study addressing this issue failed to find evidence for the NE-LC involvement in human action control (Jepma et al., 2010).

REFERENCES

Aston-Jones, G., and Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. *Annu. Rev. Neurosci.* 28, 403–450.

Baumann, N., and Kuhl, J. (2005). Positive affect and flexibility: overcoming the precedence of global over local processing of visual information. *Motiv. Emot.* 29, 123–134.

Bradley, M. M., Cuthbert, B. N., and Lang, P. J. (1996). Picture media and emotion: effects of a sustained affective context. *Psychophysiology* 33, 662–670.

Braver, T. S. (2012). The variable nature of cognitive control: a dual mechanisms framework. *Trends Cogn. Sci. (Regul. Ed.)* 16, 106–113.

Braver, T. S., Gray, J. R., and Burgess, G. C. (2007). "Explaining the many varieties of working memory variation: dual mechanisms of cognitive control," in *Variation in Working Memory*, eds A. R. A. Conway, C. Jarrold, M. J. Kane, A. Miyake, and J. N. Towse (Oxford: Oxford University Press), 76–106.

Carver, C. S. (2003). Pleasure as a sign you can attend to something else: placing positive feelings within a general model of affect. *Cogn. Emot.* 17, 241–261.

Codispoti, M., Bradley, M. M., and Lang, P. J. (2001). Affective reactions to briefly presented pictures. *Psychophysiology* 38, 474–478.

Codispoti, M., Mazzetti, M., and Bradley, M. M. (2009). Unmasking emotion: exposure duration and emotional engagement. *Psychophysiology* 46, 731–738.

Compton, R. J., Derrick, W., Golnaz, P., Eric, C., and Wendy, H. (2004). Association between positive affect and attentional shifting. *Cognit. Ther. Res.* 28, 733–744.

Dreisbach, G. (2006). How positive affect modulates cognitive control: the costs and benefits of reduced maintenance capability. *Brain Cogn.* 60, 11–19.

Dreisbach, G., Haider, H., and Kluwe, R. H. (2002). Preparatory processes in the task-switching paradigm: evidence from the use of probability cues. *J. Exp. Psychol. Learn Mem. Cogn.* 28, 468–483.

Dreisbach, G., and Haider, H. (2006). Preparatory adjustment of cognitive control in the task switching paradigm. *Psychon. Bull. Rev.* 13, 334–338.

Dreisbach, G., and Goschke, T. (2004). How positive affect modulates cognitive control: reduced perseveration at the cost of increased distractibility. *J. Exp. Psychol. Learn Mem. Cogn.* 30, 343–353.

Fuentes, L. J., and Campoy, G. (2008). The time course of alerting effect over orienting in the attention network test. *Exp. Brain Res.* 185, 667–672.

Horn, W. (1983). *Leistungsprüfsystem: L-P-S; Handanweisung für die Durchführung, Auswertung und Interpretation*, 2nd Edn. Göttingen: Hogrefe.

Hübner, M., Dreisbach, G., Haider, H., and Kluwe, R. H. (2003). Backward inhibition as a means of sequential task-set control: evidence for reduction of task competition. *J. Exp. Psychol. Learn Mem. Cogn.* 29, 289–297.

Hübner, M., Kluwe, R. H., Luna-Rodriguez, A., and Peters, A. (2004). Task preparation and stimulus-evoked competition. *Acta Psychol. (Amst.)* 115, 211–234.

Jepma, M., Te Beek, E. T., Wagenmakers, E. J., van Gerven, J. M., and Nieuwenhuis, S. (2010). The role of the noradrenergic system in the exploration-exploitation trade-off: a pharmacological study. *Front. Hum. Neurosci.* 4:170. doi:10.3389/fnhum.2010.00170

Kazén, M., and Kuhl, J. (2005). Intention memory and achievement motivation: volitional facilitation and inhibition as a function of affective contents of need-related stimuli. *J. Pers. Soc. Psychol.* 89, 426–448.

Kiesel, A., Steinhauser, M., Wendt, M., Falkenstein, M., Jost, K., Philipp, A. M., and Koch, I. (2010). Control and interference in task switching – a review. *Psychol. Bull.* 136, 849–874.

Kuhl, J., and Kazen, M. (1999). Volitional facilitation of difficult intentions: joint activation of intention memory and positive affect removes stroop interference. *J. Exp. Psychol. Gen.* 128, 382–399.

Lang, P. J., Margaret, M. B., and Cuthbert, B. N. (1999). *International Affective Picture System (IAPS): Technical Manual and Affective Ratings.* Gainesville, FL: University of Florida Center for Research Psychophysiology.

Li, K. Z., and Dupuis, K. (2008). Attentional switching in the sequential flanker task: age, location, and time course effects. *Acta Psychol. (Amst.)* 127, 416–427.

Mayr, U., and Keele, S. W. (2000). Changing internal constraints on action: the role of backward inhibition. *J. Exp. Psychol. Gen.* 129, 4–26.

McNair, D., Lorr, M., and Droppleman, L. (1971). *Profile of Mood States.* San Diego: Educational and Industrial Testing Service.

Miniussi, C., Marzi, C. A., and Nobre, A. C. (2005). Modulation of brain activity by selective task sets observed using event-related potentials. *Neuropsychologia* 43, 1514–1528.

Most, S. B., Stephen, D. S., Cooter, A. B., Levy, B. N., and Zald, D. H. (2007). The naked truth: positive, arousing distractors impair rapid target perception. *Cogn. Emot.* 21, 964–981.

Posner, J., Russell, J. A., and Peterson, B. S. (2005). The circumplex model of affect: an integrative approach to affective neuroscience, cognitive development, and psychopathology. *Dev. Psychopathol.* 17, 715–734.

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., and Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychon. Bull. Rev.* 16, 225–237.

Russell, J. A. (1980). A circumplex model of affect. *J. Pers. Soc. Psychol.* 39, 1161–1178.

Smith, J. C., Bradley, M. M., and Lang, P. J. (2005). State anxiety and affective physiology: effects of sustained exposure to affective pictures. *Biol. Psychol.* 69, 247–260.

Sudevan, P., and Taylor, D. A. (1987). The cuing and priming of cognitive operations. *J. Exp. Psychol. Hum. Percept. Perform.* 13, 89–103.

van Steenbergen, H., Band, G. P., and Hommel, B. (2009). Reward counteracts conflict adaptation: evidence for a role of affect in executive control. *Psychol. Sci.* 20, 1473–1478.

van Steenbergen, H., Band, G. P., and Hommel, B. (2010). In the mood for adaptation: how affect regulates conflict-driven control. *Psychol. Sci.* 21, 1629–1634.

van Wouwe, N. C., Band, G. P., and Ridderinkhof, K. R. (2011). Positive affect modulates flexibility and evaluative control. *J. Cogn. Neurosci.* 23, 524–539.

Vandierendonck, A., Liefooghe, B., and Verbruggen, F. (2010). Task switching: interplay of reconfiguration and interference control. *Psychol. Bull.* 136, 601–626.

Vogt, J., De Houwer, J., Koster, E. H., Van Damme, S., and Crombez, G. (2008). Allocation of spatial attention to emotional stimuli depends upon arousal and not valence. *Emotion* 8, 880–885.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 25 May 2012; accepted: 09 July 2012; published online: 27 July 2012.
Citation: Fröber K and Dreisbach G (2012) How positive affect modulates proactive control: reduced usage of informative cues under positive affect with low arousal. Front. Psychology ***3****:265. doi: 10.3389/fpsyg.2012.00265*
This article was submitted to Frontiers in Cognition, a specialty of Frontiers in Psychology.

APPENDIX

Numbers of affective picture stimuli (Lang et al., 1999).

Neutral: 7000, 7004, 7006, 7009, 7035, 7040, 7080, 7090, 7175, 7233.
$\text{Positive}_{\text{low}}$: 1440, 1710, 1750, 1920, 2057, 2150, 2260, 2311, 2340, 2530.
$\text{Positive}_{\text{high}}$: 5260, 5621, 5623, 5626, 5629, 8161, 8180, 8190, 8200, 8490.

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 01 October 2012
doi: 10.3389/fpsyg.2012.00378

Does explicit expectation really affect preparation?

Valentin J. Umbach[1]*, Sabine Schwager[1], Peter A. Frensch[1,2] and Robert Gaschler[1,2]

[1] Department of Psychology, Humboldt-Universität zu Berlin, Berlin, Germany
[2] Berlin Research Center Image Knowledge Gestaltung, Humboldt-Universität zu Berlin, Berlin, Germany

Edited by:
Rico Fischer, Technische Universität Dresden, Germany

Reviewed by:
Wim Notebaert, Ghent University, Belgium
Wilfried Kunde, Julius-Maximilians-Universitaet Wuerzburg, Germany

***Correspondence:**
Valentin J. Umbach, Department of Psychology, Humboldt-Universität zu Berlin, Rudower Chaussee 18, 12489 Berlin, Germany.
e-mail: valentin.umbach@hu-berlin.de

Expectation enables preparation for an upcoming event and supports performance if the anticipated situation occurs, as manifested in behavioral effects (e.g., decreased RT). However, demonstrating coincidence between expectation and preparation is not sufficient for attributing a causal role to the former. The content of explicit expectation may simply reflect the present preparation state. We targeted this issue by experimentally teasing apart demands for preparation and explicit expectations. Expectations often originate from our experience: we expect that events occurring with a high frequency in the past are more likely to occur again. In addition to expectation, other task demands can feed into action preparation. In four experiments, frequency-based expectation was pitted against a selective response deadline. In a three-choice reaction time task, participants responded to stimuli that appeared with varying frequency (60, 30, 10%). Trial-by-trial stimulus expectations were either captured via verbal predictions or induced by visual cues. Predictions as well as response times quickly conformed to the variation in stimulus frequency. After two (of five) experimental blocks we forced participants by selective time pressure to respond faster to a less frequent stimulus. Therefore, participants had to prepare for one stimulus (medium frequency) while often explicitly expecting a different one (high frequency). Response times for the less frequent stimulus decreased immediately, while explicit expectations continued to indicate the (unchanged) presentation frequencies. Explicit expectations were thus not just reflecting preparation. In fact, participants responded faster when the stimulus matched the trial-wise expectations, even when task demands discouraged their use. In conclusion, we argue that explicit expectation feeds into preparatory processes instead of being a mere by-product.

Keywords: explicit expectation, action control, anticipation, preparation, task goals

INTRODUCTION

"You have to expect things of yourself before you can do them," as stated by basketball legend Michael Jordan (http://www.biography.com/people/michael-jordan-9358066). Expectation is elemental in many types of behavior. It allows us to predict and prepare for an upcoming event. It can be implicit, as when we are not aware of it, or explicit. Here we focus on explicit expectations pertaining to an upcoming stimulus. These expectations can be either based on experienced stimulus frequency (made explicit through verbal predictions) or based on cues providing advance information.

Many researchers stress the role of expectation in controlling our behavior (e.g., Kunde et al., 2007; Duthoo et al., 2012). The quote above is just one example of how we take for granted that expectations influence how we go about a task. However, there are prominent findings on action control, which demonstrate that the colloquial notion of expectations influencing preparation needs empirical support. For instance, a recent brain imaging study by Soon et al. (2008) found brain activity reflecting the preparation for a free choice up to 10 s before it entered awareness (mirroring the classic "free will" experiment by Libet et al., 1983). Conscious intention might thus only be an epiphenomenon of preparatory processes in the brain (but see Trevena and Miller, 2010, for opposing evidence). Similarly, when asking someone to verbalize their expectation (about a future event that they will have to respond to) it is unclear whether the verbalized expectation simply *reflects* a preparatory state or whether it can in addition influence task processing. According to the latter view, an explicit expectation (which might be rooted in preparatory processes to some extent) feeds back into task processing. For instance, preparatory processes might be slightly stronger for one vs. another stimulus at the moment an explicit expectation is generated. The explicit expectation might feature just one of the stimuli and preparation for this option might be amplified in a winner-takes-it-all manner, because an explicit expectation had to be generated.

While the notion of expectation as a distinct construct has served as an example for redundant theorizing by critics of early cognitive psychology (e.g., Skinner, 1950) it has gained considerable support through cognitive modeling, where prediction error terms are at the core of many learning models (e.g., Sutton and Barto, 1981), as well as through the discovery of neural correlates (e.g., Schultz et al., 1997). According to Gallistel (2005) expectations have a causal role in human behavior in many economic theories and are the driving force of fast adaptation in animals to changed reinforcement schedules. The concept of expectation is discussed under various labels such as anticipation (e.g., Kunde et al., 2007), expectancy (e.g., Perruchet et al., 2006), and prediction (e.g., Sutton and Barto, 1981). Expectation encompasses both the act of looking forward as well as the thing looked forward to. In the current study, we refer to expectation as the explicit verbal prediction (or descriptive cue) of an upcoming stimulus in a sequential choice task.

In the current study, we wish to put the notion that explicit expectations have a causal role in preparation to the test. As in the work by Soon et al. (2008) we use a broad concept of preparation, encompassing any process, or state of the cognitive system that promotes the (speedy and accurate) execution of a certain action. This can take place anywhere along the cognitive processing chain, from attentional preparation (perception) to response selection (decision) to motor preparation (action). Faster responding has been shown if orientation of attention is possible in advance and facilitates perception (e.g., Posner and Petersen, 1990). On the other hand, processes of response selection and execution also benefit from preparation based on available advance information (e.g., Rosenbaum and Kornblum, 1982), which then results in faster responding. Wherever the facilitation takes place, a prepared action should be executed faster (as measured by RT). Here, we talk about *match effects* when comparing cases in which the required response matches the expectation, vs. cases in which it does not.

Expectations often originate from our experience: we expect that events occurring with a high frequency in the past are more likely to occur again in the future (e.g., Fitts et al., 1963). According to information theory (Shannon, 1948), information gain is low if an event encountered frequently before re-occurs. On the one hand, in this case little can be learned. On the other hand, the occurrence of the expected event usually boosts performance, whereas unexpected events can cause cognitive conflict and impair performance (e.g., Bernstein and Reese, 1965; Posner and Snyder, 1975). In line with the view that explicit expectation can feed back into action preparation, Miller and Anbar (1981) have suggested two routes for the impact of event frequency on action preparation: directly by strengthening S-R associations and indirectly by subjective expectations.

However, in many task situations explicit expectations and other aspects of task preparation favor the same behavior. This renders it difficult to demonstrate that explicit expectation is influencing task processing above and beyond these other aspects. For instance, a frequent S-R connection might be favored both by the high strength of the S-R association as by an explicit expectation, but it is difficult to demonstrate that the latter is actually feeding back into preparatory processes in such a situation. Therefore, we developed a paradigm in which participants can be made to expect one event (by event frequency) while another task demand (severe time constraint on a stimulus which is not the most frequent one) at the same time requires that they are preparing for a different event. If explicit expectations have an effect on task processing in a situation in which one would be better off preparing for a different event than the one expected, this would considerably strengthen the view that explicit expectations are feeding back into preparatory processes. This approach borrows its rationale from Perlman and Tzelgov (2006) who suggested scrutinizing effects that are *not* adaptive. Often, cognitive psychology builds on concepts that lend their credibility to adverse performance effects. If the effect of interest disturbs efficient performance, it is hard to explain it away. In their case, the concept of implicit learning (as distinct from controlled learning processes that in some cases might run in parallel) could be considerably supported by showing that implicit learning takes place even when it hampers performance – more learning led to worse performance. Similarly, our notion of explicit expectation as a distinct source of task processing could be backed by demonstrating dysfunctional performance effects.

In line with our perspective, a recent study by Duthoo et al. (2012) points toward the use of expectation even when it is invalid. We want to extend this finding. If, for example, people expect an event they know is very unlikely to occur, are they still preparing for it? Finding performance gains in such a case (if the unlikely event does occur) would suggest a functional role of expectation (being translated into preparation), despite the largely dysfunctional effects. As a stronger test for the impact of explicit expectation on preparation we introduced a conflicting task demand promoting the preparation of an option different from the one expected. Preparation in terms of "response readiness" (Rosenbaum and Kornblum, 1982) should be susceptible to other influences besides advance information or stimulus expectation. For example, the reinforcement of a certain response should increase its preparation state even if expectation based on past experience or situational cues favors a different response. Significant match effects in this case would suggest an influence of explicit expectation even when it is maladaptive. On the other hand, following the view of conscious intention as epiphenomenon of unconscious determinants of behavior (Libet et al., 1983; Soon et al., 2008), explicit expectations in our study should change in line with changes in preparation. If explicit expectation is merely reflecting rather than influencing task preparation, then explicit expectation should change when task preparation is experimentally changed. There is evidence, however, that subjective expectations can deviate from action preparation based on priming or associative learning (Perruchet et al., 2006). If explicit expectation is assumed to have a function in cognitive processing (as opposed to being a mere by-product) it should not be altered by a task demand that selectively manipulates preparation.

In addition to past experience, expectation can also be based on situational cues. The distinction between these two sources of expectation has been largely overlooked in research on expectation effects (but see Acosta, 1982). Results from our lab (Kemper et al., 2012) point to significant differences: self-generated predictions are accompanied by a distinctive expectation state visible in the contingent negative variation of the electroencephalogram and have a stronger effect on sensoric potentials compared to external cues, resulting in larger behavioral effects. In order to target the role of explicit expectations in preparation on a broad basis, we used both types of explicit expectations in the current study.

MATERIALS AND METHODS

In a series of four experiments, we used a three-choice reaction time task. Stimuli were displayed with different frequencies, with one stimulus being presented in 60% of all trials, another one in 30%, and the last in 10% of all trials. Participants responded to each stimulus by pressing one of three keys. As a measure of trial-wise subjective expectation we asked participants to verbally predict the upcoming stimulus on each trial (Experiment 1: *verbal* *predictions*). To control for effects of this verbalization procedure, we ran a variant where no predictions were required (Experiment 2: *no predictions*). In two additional experiments, we replaced the self-generated predictions with external cues indicating the upcoming stimulus. Cues were either not predictive of the subsequent stimulus presentation (Experiment 3: *non-informative cues*), or they correctly indicated the upcoming stimulus on 80% of all trials (Experiment 4: *informative cues*). In order to test for effects of explicit expectation when it is not perfectly in line with other demands for task preparation, we introduced a response deadline for the medium frequency stimulus toward the second half of all experiments.

EXPERIMENT 1: VERBAL PREDICTIONS

Responses to the more frequent stimuli should generally be faster because of stronger S-R associations and because they are expected more often (Miller and Anbar, 1981). Subjective predictions (in Experiment 1) should also reflect this frequency pattern, with participants more often predicting the more frequent stimuli. A common phenomenon in this context is the tendency of people to match their predictions to the observed probabilities, resulting in fewer correct predictions compared to an optimal strategy (i.e., always predict the most frequent event). This phenomenon has been described as probability matching (e.g., Gaissmaier and Schooler, 2008). Participants should display the same tendency in our task if they really try to predict the upcoming stimulus. Therefore, finding a frequency effect in explicit expectations provides a manipulation check to ensure that participants are in fact correctly performing the task of verbalizing their expectations in our experiment. While actual stimulus presentation was unrelated to these subjective predictions, responses should be faster after (coincidental) correct predictions if people use their predictions to prepare for task execution.

Faster responses to correctly predicted stimuli (match effects) would point toward a mandatory use of subjective expectation in action preparation. Since there is no relation between participants' predictions and the actual stimulus they have to respond to, there is no reliable gain for them in following their predictions. This holds in particular for predictions of the two less frequent stimuli. To challenge the assumption of a mandatory use of explicit expectations even further, we introduced an additional task demand with the goal of diverting preparatory processes away from the response to the expected stimulus. After two of five experimental blocks participants were instructed to give particularly fast responses to occurrences of the medium frequency stimulus (30%). Slow responses on these trials were punished by presenting an unpleasant noise which acted as a negative reinforcement. This additional task demand was therefore at odds with the pattern set up by the stimulus frequencies. While stimulus frequency and subjective expectations should lead to faster responses for the most frequent stimulus, the additional task goal (avoid the unpleasant noise) should lead to a stronger preparation for the medium frequency stimulus. It makes preparation on the basis of frequency expectations less useful because preparing for the predicted response may result in hearing the aversive sound in some cases (i.e., when the frequent stimulus is predicted and prepared and the medium frequent stimulus occurs and is responded to too slowly). Still finding match effects under these conditions would be further evidence for the mandatory use of explicit expectation in preparing for an upcoming task. To the extent participants are able to adjust their preparation to the requirements of the actual task one could expect reduced expectation match effects in blocks three to five: participants should rely less on their stimulus predictions if the medium frequent response is reinforced.

Match effects (faster responses following correct predictions) are in line with our idea that people use their explicit subjective expectations in action preparation. However, there is the possibility that these expectations are simply a by-product of preparation without functional use. In this case, participants should adjust their predictions in line with the changes in action preparation once the additional task demand is established. If participants in fact prepare to respond to the medium frequency stimulus, and if their stimulus expectations are inseparably linked to this preparation (as in "reading out" an internal preparation state determined by the strength of specific S-R associations), this should be reflected in their prediction frequencies. In this case, match effects might not be reduced (see above), as both preparation and prediction would follow the altered task demands. If, on the other hand, people generate expectations independently of action preparation that is fueled by a second task demand, the frequency pattern should remain intact in their subjective predictions.

EXPERIMENT 2: NO PREDICTIONS

In Experiment 1 verbal predictions were required before each stimulus occurrence resulting in a dual-task like situation: to generate verbal predictions and to perform the manual choice reaction task. This could have resulted in different processing of the choice task as compared to solely producing choice reactions. In order to verify the results found for frequency and, particularly, the effect of selective reinforcement of the medium frequent stimulus, we repeated the experiment without verbal predictions.

EXPERIMENT 3: NON-INFORMATIVE CUES

Expectation effects are most often investigated by using external advance information (provided by cues, e.g., Posner and Snyder, 1975; Miller and Anbar, 1981; Mattler, 2004). It has been shown, however, that expectations induced by cues affect performance differently from predictions generated by participants themselves (Kemper et al., 2012). Against this background we repeated Experiment 1 and replaced verbal predictions with visual, non-verbal cues that announced one of the three stimuli in advance before the imperative stimulus was presented. The probability of match was kept at approximately the same level as in the prediction experiment by presenting the cues with the same frequencies as the stimuli (10, 30, and 60%) but randomized independently of stimulus presentation. The general effect of stimulus frequency should be similar to the previous experiments, as well as the impact of the selective response deadline. In line with previous studies (Acosta, 1982; Kemper et al., 2012) we expect a smaller match effect with cues than with predictions.

EXPERIMENT 4: INFORMATIVE CUES

We conducted Experiment 4 for two reasons. First, the use of non-informative cues is quite atypical for investigating expectation effects by the help of external advance information. Usually, cueing effects on preparation appear only with highly reliable cues (e.g., Alpay et al., 2009; Scheibe et al., 2009). The reason for finding an effect under such unfavorable conditions as in Experiment 3 might lie in feature overlap between cue and stimulus. Second, we wanted to explore an idea that could explain the difference in effectiveness between explicit expectations generated by the individual or provided by external advance information. As the overall real validity of predictions (Experiment 1) and cues (Experiment 3) was comparable the difference might in fact go back to the degree to which participants rely on their expectation, depending on its source. One possible mechanism could be that participants weight self-generated predictions stronger and that external information has to be of a much higher validity to be included into controlled action preparation, or, alternatively, predictions and cues differ in subjective usefulness.

Therefore, in Experiment 4 we increased the probability of match between cue and stimulus feature to 80%. Under these conditions a much larger effect of expectation match than in Experiment 3 should be observable. We expect comparable effects of stimulus frequency as in the previous experiments, as well as an effect of selectively reinforcing the medium frequent stimulus by use of a deadline.

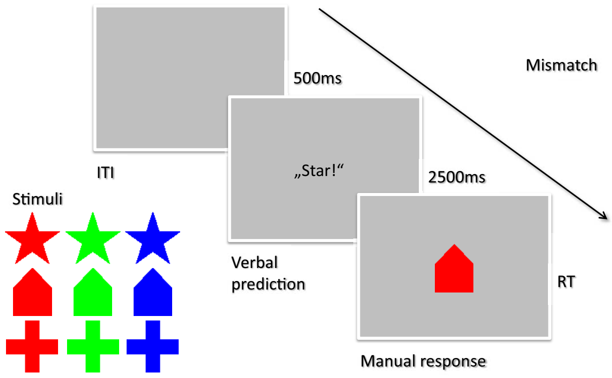

FIGURE 1 | Task used in Experiment 1. On each trial, participants had to verbalize their prediction for the upcoming stimulus (in this case "star"). After 2500 ms the stimulus appeared on screen (in this case a house, signifying a mismatch) and participants had manually respond by pressing one of three keys. The next trial started 500 ms after the response. For any given participant, only one of the two stimulus features (shape, color) was relevant throughout the task (in this case, both predictions and responses pertained to the shape of a stimulus).

Participants

One hundred five undergraduate students of psychology and other fields (74 women, mean age = 24.9 years) participated in individual sessions lasting approximately 90 min (Experiments 1 and 2) or 60 min (Experiments 3 and 4). Participants either received partial course credit or were paid 8–12 euros for their time. They provided written informed consent, particularly to the exposure to aversive sounds.

Design and procedure

In all of the experiments reported here, we used a three-choice reaction time task. Three different shapes served as stimuli – star, house, and cross – that were presented in one of three colors, red, green, or blue. Each stimulus could be named by a monosyllabic word in order to provide for approximately equal verbalization times (for Experiment 1; German "Stern," "Haus," "Kreuz," or "rot," "grün," "blau"). Stimuli were displayed centrally on a 17″ CRT computer monitor with a light gray background and occupied approximately 2.2 cm in width and height (corresponding to a visual angle of about 6.4° at a viewing distance of 60 cm). Three keys (V, B, and N) on a standard Windows keyboard were mapped by instruction either to the three shapes or the three colors, with the relevant feature varying between participants. The task and stimuli are shown in **Figure 1**.

Frequency of the three possible shapes or colors, respectively, was predetermined in the stimulus set to yield three frequency classes, frequent (60%), medium (30%), and infrequent (10%) for the relevant stimulus feature. Occurrence of the irrelevant feature was equally distributed and co-occurrence was balanced across features. Half of the participants used shapes as relevant feature for predictions and response selection and the others used color. The irrelevant feature was not used in this task.

Participants completed five blocks of 120 trials for a total of 600 trials. The frequent stimulus occurred in 72 trials, the medium frequent in 32 trials, and the infrequent stimulus in 12 trials per block. After the first two experimental blocks the additional task demand was introduced. Participants were informed that their reactions to the medium frequency stimulus (which was simply described by its label) had to be extra fast if they wanted to avoid the annoying sound on their headphones. This aversive auditory stimulus, a white noise burst of about 75 dB, had been demonstrated to participants at the beginning of the session before they gave their consent to the procedure. The response deadline for the medium frequency stimulus was individually determined at the median reaction time for the frequent stimulus in the preceding Block 2 and kept constant over the remaining three blocks. If participants exceeded this deadline on any given trial with the medium frequency stimulus, the aversive sound was immediately presented on their headphones and ended 500 ms after their (late) reaction.

At the end of the session participants were asked to estimate the frequency of the relevant stimulus feature.

EXPERIMENT 1: VERBAL PREDICTIONS

On every trial, participants were asked for their subjective expectation regarding the upcoming stimulus. According to the relevant stimulus feature, the prompt "Farbe?" or "Form?" (German for color or shape) were displayed on the screen. Participants then had 2500 ms to verbalize their expectation. If voice onset was registered more than 1500 ms after the onset of the prompt, participants were reminded to speak as soon as the prompt is shown on the next trial. In addition, participants were randomly reminded in 10% of all trials to speak loudly and clearly. After this expectation interval (2500 ms after the prompt onset) the stimulus was shown and participants had to press the corresponding key on the keyboard. The following trial started 500 ms after the response. The experimental blocks were preceded by three practice blocks of 18 trials each in which manual responses and verbal expectations were first trained separately and then combined. Frequency and combinations of relevant and irrelevant stimulus feature were equally distributed in the practice blocks.

Verbal expectations were captured with a microphone headset and identified using a real-time speech recognition program implemented in Matlab (Donkin et al., 2009). At the beginning of the experimental session, the software was trained to the individual voice with the participant repeating the words in the response set 10 times. This was followed immediately by an accuracy check with 10 additional exemplars per word. If recognition accuracy was below 95% (i.e., more than one misidentification) the original training was restarted, otherwise the additional exemplars were added to the pool of training exemplars and the experiment commenced. Recognition accuracy was tested again at the end of the session.

EXPERIMENT 2: NO PREDICTIONS

The task was the same as in Experiment 1, with the only difference that participants were not instructed to generate verbal predictions at the beginning of each trial. Instead of the prompts used in Experiment 1 a fixation dot was displayed for 2500 ms to keep the timing equivalent to Experiment 1.

EXPERIMENT 3: NON-INFORMATIVE CUES

Again, the task was largely the same as in Experiment 1. Instead of prompting participants to verbalize their subjective expectations on each trial symbolic cues were presented predicting the upcoming stimulus. These cues were similar to the imperative stimuli but only varied in the relevant feature: if a participant had to respond to the shape of a (colored) stimulus the cues consisted of black shapes, if color was the relevant feature colored circles were used as cues. Participants did not have to verbalize the cues. Cues were displayed 1000 ms after the last response and remained visible for 1000 ms followed by a blank screen for another 1000 ms, after which the imperative stimulus appeared. Thus, the response-stimulus interval was the same as in the other experiments (3000 ms) and the timing of the cues was similar to the verbal predictions in Experiment 1. Importantly, cue presentation was randomized independently and was not related to the subsequent stimulus presentation. Therefore cues exhibited the same low overall validity as the predictions in Experiment 1: on only 46% of all trials was a cue followed by the corresponding stimulus (60% for the frequent stimulus, 30% for the medium, and 10% for the infrequent stimulus).

EXPERIMENT 4: INFORMATIVE CUES

The task was the same as in Experiment 3, except that the validity of cues was 80% for all frequencies. Thus, in 80% of all trials a cue was followed by the corresponding stimulus.

RESULTS

EXPERIMENT 1: VERBAL PREDICTIONS

Three participants were excluded for producing too many false responses (>10%), another two participants were excluded because of problems with the speech recognition software (<75% accuracy in the post-experiment test). Data of the remaining 19 participants were analyzed. For the following analyses all trials were recruited, including those with immediate stimulus repetitions. The proportion of stimulus repetitions naturally were related to stimulus frequency, with 60% repetition trials for the frequent stimulus, and 29 and 9% respectively for the medium and infrequent stimuli. All results reported here remain unaltered if stimulus repetitions, i.e., 46% of all trials, are excluded. RT analyses are based on correct responses only, excluding error trials. The response deadline, representing the median reaction time for the frequent stimulus in Block 2, was on average set at 424 ms (SD = 76 ms), with individual participants ranging between 303 and 633 ms. In 28% of the trials with the reinforced stimulus, participants passed this deadline and were consequently exposed to the aversive sound (32% in Block 3, 24% in Block 4, 27% in Block 5).

Our experiments, except Experiment 2 with no predictions, included three within-subjects factors: *match* (testing the effectiveness of explicit expectation), *block* (mirroring the effect of training and, more importantly, of the introduction of the response deadline from block 2 to block 3), and *frequency*. A three-factorial repeated measures ANOVA could not be run as participants did not contribute enough data points to one of the cells (match trials for the infrequent stimulus occurred too rarely to get reliable medians per block). Therefore, three two-way ANOVAs were run over the response times and error rates of all experiments: one with *frequency* and *block* to examine the general effect of selectively reinforcing the medium frequent response, one with *match* and *frequency* to look for a potential dependency of the size of expectation effects on experienced stimulus frequency, and one with *match* and *block* to examine the interaction of expectation and the deadline manipulation. In the context of a Bonferroni correction we divided the critical significance level (alpha = 0.05) by three in order to account for repeated tests on one and the same data set.

Before the introduction of the response deadline, RTs and errors followed stimulus frequency. The infrequent stimulus led to the slowest and most error prone reactions and the responses to frequent stimuli were the fastest and most accurate. The medium frequency stimuli lay in between. With the response deadline, in the last three blocks, responses to the medium frequency (reinforced) stimulus became faster than responses to the more frequent stimulus, while response times for all stimuli decreased. A two-way repeated measures ANOVA with the factors *frequency* and *block* revealed main effects for both *frequency*, $F(2, 36) = 81.63$, $p < 0.001$, and *block*, $F(4, 72) = 82.27$, $p < 0.001$, as well as an interaction, $F(8, 144) = 15.91$, $p < 0.001$. Importantly, the selective speedup of responses to the medium frequent stimulus was not achieved at the expense of a higher error rate for the frequent stimulus (see **Figure 2**, top left). The same effects as in RT were found in the error rates (all $p < 0.001$).

Verbal predictions already reflected the frequency differences in the first block and approached the actual values over the course of the experiment. Importantly, this pattern was not altered with the introduction of the response deadline in the third block (see **Figure 3**). Therefore, participants continued to expect the most frequent stimulus most often but reacted fastest to the medium frequency stimulus. The three different stimuli were predicted in the order of their frequency of occurrence (most often the most frequent stimulus, less often the medium frequent stimulus,

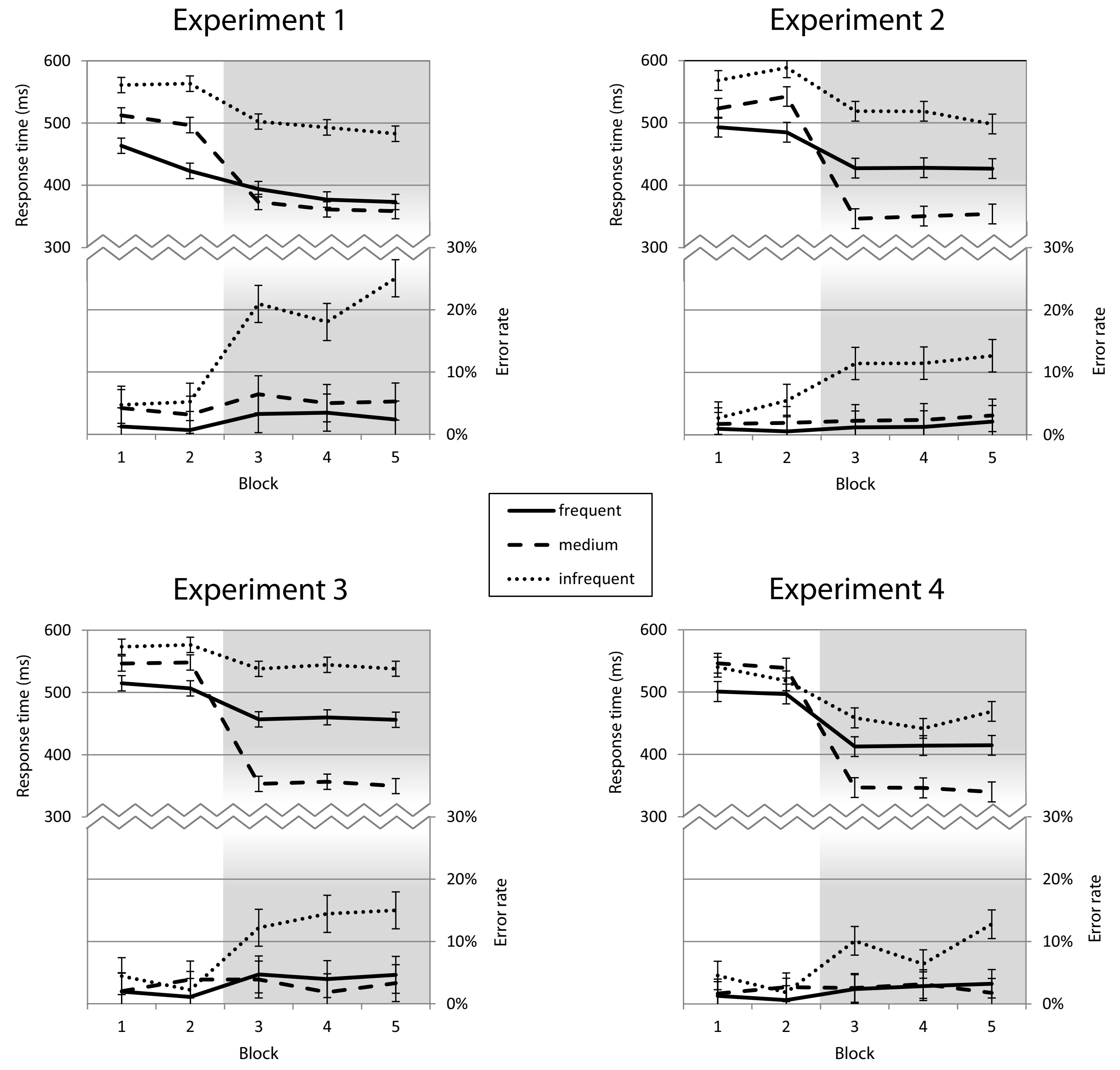

FIGURE 2 | Response times and error rates for Experiments 1–4. Response times (on the top of each panel) exhibit an effect of stimulus frequency (with the frequent stimulus, marked by the solid line, leading to faster responses than the medium and infrequent stimuli) in the first two blocks, as well as an effect of the additional task goal starting in the third block (with the reinforced medium stimulus getting faster responses). The same pattern is visible in conditions with verbal predictions (Experiment 1) and without (Experiment 2) and also with low and high validity non-verbal cues (Experiments 3 and 4).

and least often the rare stimulus). This rank order of prediction frequencies stayed the same over the experiment, so that prediction behavior was highly correlated over blocks (correlation of ranks between successive blocks: $\tau = 0.74$, 0.79, 0.92, and 0.83, all $p < 0.001$), regardless of the changed pattern in choice performance.

Stimuli matched predictions in 42% of all trials (with a minimum of 40% in Block 1 and a maximum of 44% in Block 4; 51% matches for the frequent stimulus, 30% for the medium, and 18% for the infrequent stimulus). Response times were shorter for trials in which the stimulus matched the participant's prediction, as compared to mismatch trials. This match effect was visible for all stimulus frequencies. The ANOVA with the factors *match* and *frequency* revealed main effects on RT for *match*, $F(1, 18) = 130.72$, $p < 0.001$, and *frequency*, $F(2, 36) = 74.55$, $p < 0.001$, but no interaction *match* × *frequency*, $F(2, 36) = 2.77$, ns. After introducing the response deadline for the medium stimulus, the mean difference between match and mismatch trials declined from 110 ms in Block 2 to 60 ms in Block 3 (see **Figures 4** and **5**, top left).The ANOVA with the factors *match* and *block* revealed main effects on RT for *match*, $F(1, 18) = 107.63$, $p < 0.001$, and *block*, $F(4, 72) = 81.14$, $p < 0.001$, as well as an interaction *match* × *block*, $F(4, 72) = 30.56$, $p < 0.001$. The same effects were found in the error rates (all $p < 0.001$).

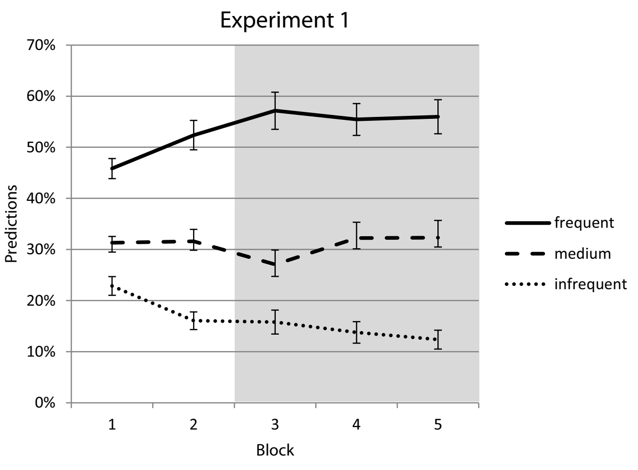

FIGURE 3 | Predictions in Experiment 1 already reflect the varying stimulus frequencies in the first block and approach the actual values (60, 30, and 10%) over the course of the experiment. The additional task goal introduced in the third block does not change this pattern.

The *post hoc* estimates of stimulus occurrence in percent made by the participants also provided a good approximation of the actual frequencies, with the frequent stimulus at 63%, the medium at 24%, and the infrequent stimulus at 13%.

EXPERIMENT 2: NO PREDICTIONS

One participant was excluded from analyses for producing too many false responses (>10%). Data of the remaining 21 participants were analyzed. The response deadline was on average fixed to 491 ms (SD = 89 ms), with individual participants ranging between 333 and 693 ms. On 8% of the trials with the reinforced stimulus, participants exceeded this deadline and were consequently exposed to the aversive sound (9% in Block 3, 7% in Block 4, 9% in Block 5).

The same pattern emerged as in Experiment 1: Responses were faster and more accurate to the more frequent stimuli in the first two experimental blocks, before the introduction of the response deadline. With the deadline, in the last three blocks, responses to the reinforced medium frequent stimulus became faster than responses to the frequent stimulus, while response times for all stimuli decreased (see **Figure 2**, top right). A two-way repeated measures ANOVA on RTs with the factors *frequency* and *block* revealed main effects of both stimulus *frequency*, $F(2, 40) = 71.87$, $p < 0.001$, and *block*, $F(4, 80) = 58.96$, $p < 0.001$, as well as an interaction, $F(8, 160) = 25.02$, $p < 0.001$. The same effects were found in the error rates (all $p < 0.001$).

The *post hoc* estimates again provided a good approximation of the actual frequencies, with the frequent stimulus at 64%, the medium at 25%, and the infrequent stimulus at 11%.

EXPERIMENT 3: NON-INFORMATIVE CUES

Seven participants were excluded from analyses for producing too many false responses (>10%). Data of the remaining 30 participants were analyzed. The response deadline was on average fixed at 502 ms (SD = 75 ms), with individual participants ranging between 383 and 695 ms. On 8% of the trials with the reinforced stimulus participants exceeded this deadline and were consequently exposed to the aversive sound (9% in Block 3, 7% in Block 4 and 5).

Similar to Experiment 2, RTs followed stimulus frequency in Blocks 1 and 2, but the medium frequency stimulus elicited the fastest responses when the reinforcement procedure started after Block 2 (compare **Figure 2**, bottom left). A two-way repeated measures ANOVA with *frequency* and *block* revealed main effects on RT for *frequency*, $F(2, 58) = 107.33$, $p < 0.001$, and *block*, $F(4, 116) = 71.45$, $p < 0.001$, as well as an interaction *frequency* × *block*, $F(8, 232) = 61.28$, $p < 0.001$. The same effects were found in the error rates (all $p < 0.001$). The RT effect of cue match was clearly present for all frequencies as well, but smaller than the effect of expectation match in Experiment 1. The ANOVA with *match* and *frequency* revealed main effects on RT for *match*, $F(1, 29) = 21.57$, $p < 0.001$, and *frequency*, $F(2, 58) = 80.11$, $p < 0.001$, but no interaction *match* × *frequency*, $F(2, 58) = 0.43$, ns. In the error rates, only frequency yielded a significant effect, $F(2, 58) = 16.64$, $p < 0.001$. After the introduction of the response deadline the match effect was diminished from 35 ms in Block 2 to 19 ms in Block 3 (see **Figures 4** and **5**, bottom left). The third ANOVA with *match* and *block* revealed main effects on RT for *match*, $F(1, 29) = 23.41$, $p < 0.001$, and *block*, $F(4, 116) = 44.47$, $p < 0.001$, as well as an interaction *match* × *block*, $F(4, 116) = 13.74$, $p < 0.001$. The same effects were found in the error rates (all $p < 0.01$).

The *post hoc* estimates again provided a good approximation of the actual frequencies, with the frequent stimulus at 57%, the medium at 30%, and the infrequent stimulus at 13%.

EXPERIMENT 4: INFORMATIVE CUES

Four participants were excluded from analyses for producing too many false responses (>10%). Data of the remaining 18 participants were analyzed. The response deadline was on average fixed to 497 ms (SD = 120 ms), with individual participants ranging between 311 and 708 ms. On 11% of the trials with the reinforced stimulus participants passed this deadline and were consequently exposed to the aversive sound (13% in Block 3, 12% in Block 4, and 9% in Block 5).

As shown in **Figure 2** (bottom right), RTs followed stimulus frequency in the first two blocks until the onset of the reinforcement of the medium frequency stimulus at the beginning of Block 3 led to faster responses to this stimulus. The two-way repeated measures ANOVA with *frequency* and *block* revealed main effects on RT for *frequency*, $F(2, 34) = 45.83$, $p < 0.001$, and *block*, $F(4, 68) = 34.74$, $p < 0.001$, as well as an interaction *frequency* × *block*, $F(8, 136) = 22.99$, $p < 0.001$. The same effects were found in the error rates (all $p < 0.001$). RT effects of match between cue and stimulus were much more pronounced than in the low validity variant explored in the previous experiment and were not reduced after the introduction of the response deadline (102 ms in Block 2, 98 ms in Block 3, see **Figures 4** and **5**, bottom right). Accordingly, the ANOVA with *match* and *block* revealed main effects on RT for *match*, $F(1, 17) = 110.14$, $p < 0.001$, and *block*, $F(4, 68) = 41.30$, $p < 0.001$, but no interaction *match* × *block*, $F(4, 68) = 0.55$, ns. The third ANOVA with *match* and *frequency* revealed main effects on RT for *match*, $F(1, 17) = 113.20$, $p < 0.001$, and *frequency*, $F(2,$

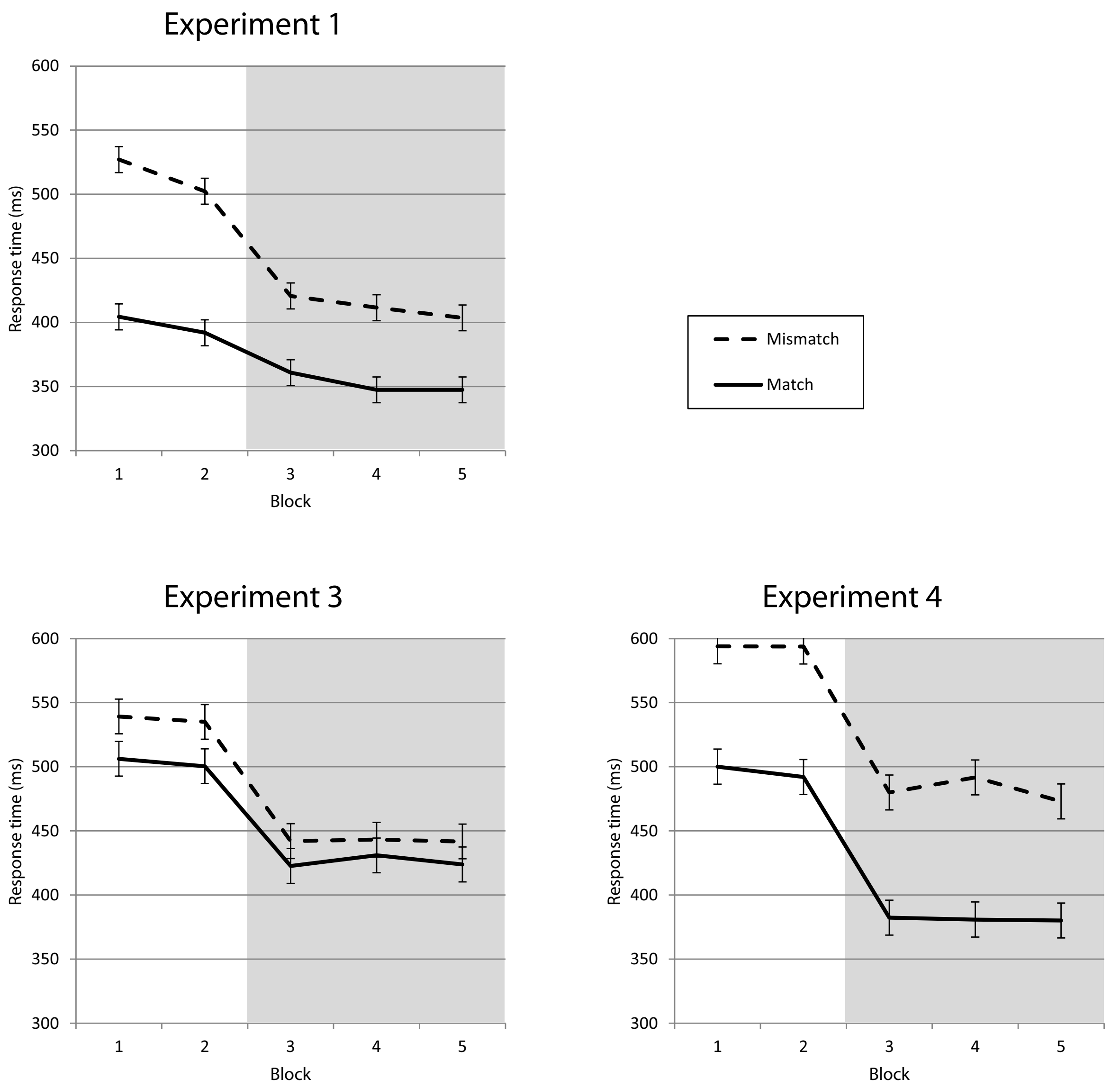

FIGURE 4 | Match vs. mismatch trials for Experiments 1, 3, and 4 (pooled over frequencies). Responses following correct predictions (match) in Experiment 1 are 117 ms faster on average compared to incorrect predictions (mismatch) in the first two blocks; after the introduction of the selective response deadline for the medium frequency stimulus this difference is reduced to 60 ms on average. In Experiment 3, using invalid cues (similar to the predictions of Experiment 1), the difference between match and mismatch trials averages 34 ms at the beginning and is down to 17 ms with the additional task demand. Experiment 4 shows no reduction in this mismatch effect, with 98 ms before and 100 ms after the introduction of the deadline on average.

$34) = 76.56$, $p < 0.001$, but no interaction *match* × *frequency*, $F(2, 34) = 7.56$, ns.

The *post hoc* estimates again provided a good approximation of the actual frequencies, with the frequent stimulus at 56%, the medium at 30%, and the infrequent stimulus at 14%.

DISCUSSION

In all four experiments reported here, stimulus frequencies (60, 30, 10%) were reflected in response times and error rates, with the most frequent stimulus producing the fastest and most accurate responses. While discussion about the role of conscious intention in controlling behavior (Libet et al., 1983; Soon et al., 2008; Trevena and Miller, 2010; see Introduction) might be taken to suggest that explicit expectations merely reflect other preparatory processes but do not influence them, our results suggest that explicit expectations feed back into task processing and thus have a causal role. We disentangled explicit expectation from other forms of preparation by adding a secondary task demand. With instruction and a response deadline combined with an aversive sound, participants were encouraged to prepare for a different stimulus (i.e., the medium frequency stimulus) than the one they were expecting most often (i.e., the high frequency stimulus). Explicit

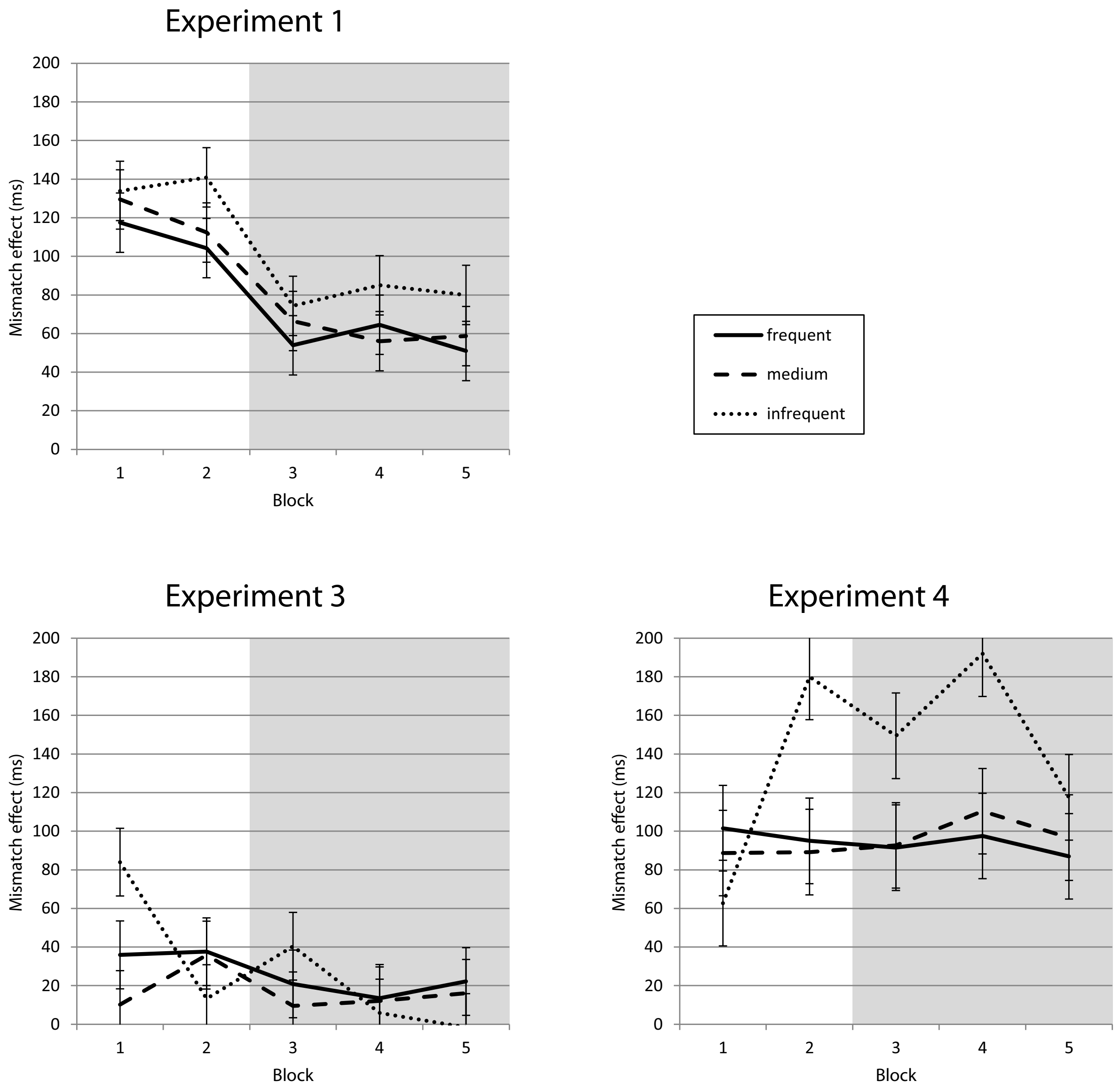

FIGURE 5 | Mismatch effect for Experiments 1, 3, and 4 (displayed by frequency). No differences in mismatch effect between frequencies, similar reduction (Experiment 1) or stability (Experiment 3) with the additional task goal introduced in Block 3. Values from infrequent stimuli are statistically unreliable because of the low number of match trials.

expectations affected task processing even when it would have been beneficial not to rely on them: On the one hand, effects of expectation conflicted with the requirement to respond faster than the response deadline on the medium frequency stimulus. This could have largely been avoided if participants had either not have turned verbalized expectation into task preparation or, alternatively, would have started to explicitly expect the medium frequency stimulus in most or all trials. On the other hand, participants showed faster response times when their expectation matched rather than mismatched the stimulus even in case of the infrequent stimulus – which they sometimes expected. Such an expectation was mostly followed by the frequent or medium frequent rather than the infrequent stimulus. In principle one could have betted on and prepared for the frequent or medium stimulus, despite verbalizing an expectation for the infrequent one. A mismatch was much more likely than a match after such a prediction, yet matches were faster than mismatches. It would have been conceivable that participants show RT benefits of expectations matching the stimuli in case of frequent and medium frequency stimuli and a reversal of the expectation match effect in case of the infrequent (10%) stimuli. For instance, Notebaert et al. (2009) have reported that in cases with a majority of error trials RTs are prolonged after the rare correct trials rather than after error trials, suggesting that event frequency rather than match vs. mismatch of task demands and action can drive performance costs. This does not seem to count for explicit expectations, however.

Thus, neither were explicit expectations themselves chosen flexibly to boost performance, nor could the aftereffects of these expectations be flexibly regulated. The results thus suggest that explicit expectations influence rather than merely reflect other preparatory processes and do so rather inflexibly. Explicit expectations seem to count – even when they are not adaptive to current task demands.

In the current experiments we took two different approaches by measuring expectations through verbal predictions and inducing them by cues. In *Experiment 1* we asked participants to verbally predict the upcoming stimulus on each trial and then respond to the actual stimulus by pressing the corresponding key. Verbal predictions (as a measure of subjective expectation) mirrored actual stimulus frequencies already in the first experimental block, with participants predicting the most frequent stimulus on a higher proportion of trials. When the imperative stimulus matched the prediction on a given trial, participants responded much faster compared to trials on which the stimulus violated their prediction. This gain was similar for all three stimulus frequencies, suggesting that participants used their predictions to prepare the response even if it was unlikely to be fulfilled (18% for the infrequent stimulus, compared to 51% for the frequent stimulus). Introducing the response deadline for the medium frequency stimulus reduced this match effect from 117 to 60 ms, while predictions themselves were not altered.

In *Experiment 2* we replicated the effects of stimulus frequency without verbal predictions, ruling out the possibility that the response time effects found in Experiment 1 were dependent on the second task of explicitly verbalizing stimulus expectations. In *Experiment 3* we induced explicit expectations through symbolic cues. As cue presentation was not related to the subsequent stimulus, their predictive value was as low as that of the self-generated predictions in Experiment 1. There was a small match effect with faster responses following correct cues (34 ms) before the introduction of the response deadline that was diminished to a statistically non-significant difference (17 ms) with the additional task demand. In *Experiment 4*, with cues correctly predicting the upcoming stimulus in 80% of all trials, there was a large match effect that was not reduced by the response deadline (98 ms before, 100 ms after the manipulation). This deviates from the patterns found in Experiments 1 and 3, where the additional task demand (fast responses on the medium frequency stimulus to avoid the aversive tone) led to a reduction in the match effect.

DOUBLE IMPACT OF STIMULUS FREQUENCY

In addition to explicit expectations, RT was affected by stimulus frequency in all four experiments. This is in line with earlier calls to integrate associative as well as an expectancy-based accounts of action preparation. For instance, Miller and Anbar (1981) argue that frequency effects on response time can arise directly (through the strength of S-R associations) and indirectly (through subjective expectancies). Asking participants to verbalize their expectations (in Experiment 1) might have led to larger RT differences between stimuli of different frequency compared to the variants without predictions (Experiment 2) or with external cues (Experiments 3 and 4). Frequency effects might have been prominent on two rather than just one path in Experiment 1. As frequency effects remained evident after the introduction of the response deadline for the medium frequency stimulus, this is pointing toward an automatic effect of S-R frequency and as such toward an independent contribution of this source.

Subjective expectations measured as predictions in Experiment 1 closely mirrored the frequency pattern, a phenomenon also known as probability matching (e.g., Gaissmaier and Schooler, 2008, see below). Thus, performance in predicting the upcoming stimulus was also influenced by the given frequency pattern. The participants presumably made use of their prior experience represented in associations of varying strength. However, the *effect* of subjective expectation and the general effect of frequency on performance in the choice task appear to be independent from each other. Match effects were of similar size for all frequencies, or, to put it differently: the general effect of frequency proved to be the same, regardless of expectation match. This also holds for the experiments where cues instead of predictions were used. That is, the influence of explicit expectation on task processing appears to be different from other effects that arise from stimulus frequencies.

PREDICTIONS: MATCHING VS. MAXIMIZING

Predictions were generated and used in a less than optimal manner. Participants could have maximized their correct predictions (in Experiment 1) by always predicting the most frequent stimulus (which would have lead to 60% matches). Instead, they apparently tried to reproduce the observed stimulus frequencies in their predictions (resulting in only 42% matches). This behavior is in line with the probability matching phenomenon (e.g., Gaissmaier and Schooler, 2008). Trials with expectations matching the stimulus were faster than those with a mismatch. For boosting performance in the choice reaction task it would have been favorable to choose to predict the most frequent stimulus on all trials in the first part of the experiment and the medium frequency stimulus once the response deadline on this stimulus was set in place. Maximizing has been observed in the literature on strategy change in skill acquisition (e.g., Touron and Hertzog, 2004; Gaschler and Frensch, 2007, 2009) where people tend to exclusively choose the one of two processing strategies that is the most suitable on most of the trials. This however, might be an exception as in many other task contexts probability matching has proven to be a robust phenomenon (see, e.g., Gallistel, 2005, for a discussion). He suggested that probability matching is a "hard-wired" policy which is useful in dynamic environments as it guarantees continuous sampling of the options so that an agent does not run the risk of missing to notice changes in which options are currently more or less rewarding. Our results lend further support to this "hard-wired" view, as the *influence* of the probability-matched expectations appears not to be easily adapted to more promising strategies either. However, we do not know for certain what the goals of our participants in optimizing their task performance are. It is possible that they tried to find a balance between the two tasks of realistically predicting stimuli while performing rapidly and correctly on the choice task. Therefore, instructing them to increase their proportion of matches might change the pattern of results.

CONFLICTING TASK DEMAND ATTENUATES IMPACT OF EXPECTATION

The match effects we found, with faster responses following correct predictions and valid cues, are compatible with the idea that explicit expectation serves as a trigger for action preparation and thus assumes a causal role in cognitive processing. However, there are differences in the robustness of these match effects that depend on the source of expectation on the one hand and on its validity on the other hand.

The additional task demand of trying to respond quickly to the medium frequency stimulus in order to avoid hearing the unpleasant sound significantly reduced the match effects in Experiments 1 and 3, but not in Experiment 4. While in Experiment 4 the cue was highly predictive of the stimulus, explicit expectations (Experiment 1) and cues in Experiment 3 were equally unreliable. Arguably, the strong associations between cue and stimulus in Experiment 4 were still fully effective under the response deadline, whereas the impact of the unreliable predictions in the other experiments could be attenuated. Importantly, the match effect was reduced for all stimuli to a similar extent. The predictions that could have boosted the processing of the medium frequency stimulus with the deadline attached to it, were apparently not spared. Rather, participants seem to have relied somewhat less on expectations in general.

While the influence of the non-informative cues (in Experiment 3) on response time was effectively removed by the additional task demand, subjective predictions retained a significant impact. This suggests that self-generated predictions are mandatorily processed and trigger action preparation even if they are obviously unreliable and if task demands favor the preparation of a different action. As Kunde et al. (2007) argue, expectation is an integral component of action control. Expectations are always generated and translated into preparation (of perception or action) as this is usually beneficial to optimize behavior in real life. Artificial external cues do not share this processing privilege by default and have to first prove their usefulness (reliability). When they do, however, as in Experiment 4 (with 80% valid cues), they retain their influence in spite of the additional task demand.

EXPECT ONE THING, PREPARE FOR ANOTHER

The selective reinforcement of the medium frequent stimulus led to a selective speed up of responses to the reinforced stimulus. Thus, participants in our study apparently were able to predict one thing while at least partly preparing for another. A similar dissociation between explicit expectation and overt behavior has been reported before (Perruchet et al., 2006) for simple reactions in an associative learning experiment. In the "Perruchet effect," response time (as a measure of automatic activation) decreases with increasing number of repeated associations, while explicit expectation develops in the opposite direction, increasingly favoring an alternation after longer runs of repetitions (the "gambler's fallacy"). However, in contrast to the build-up of associative effects, in our study the change in performance occurred immediately after instructing the new requirement, rather than gradually. The abrupt effect of the deadline suggests that intentional control processes can influence the extent to which learned S-R connections impact behavior. The ordering of RTs by stimulus frequency was immediately altered. With the stimulus-specific deadline, the RT for the medium frequency stimulus surpassed RT for the frequent stimulus. In line with the intentional weighting principle proposed by Hommel et al. (2001), intentional control might put some extra strength on a response alternative that would have been otherwise weak and so alter the result of the competition for response selection. Put differently, if something we have learned earlier (as, e.g., expecting stimuli with a given frequency) conflicts with actual task goals (as, e.g., responding fast to a less expected stimulus), behavior will always be the result of resolving this – classical – conflict situation (see Botvinick et al., 2001). If expectations conflict with other task demands it seems feasible to prepare for something one is not expecting.

CONCLUSION

We have shown that explicit expectation affects preparatory processes and thus assumes a causal role in controlling behavior. This finding speaks against the notion of explicit expectation as a mere by-product of preparation. When we ask participants for their subjective predictions about an upcoming event they have to respond to, they are preparing for what they say (instead of telling us what they are preparing for).

ACKNOWLEDGMENTS

We thank Katharina Kanthak and all other members of our lab for their help in conducting the experiments as well as providing helpful feedback throughout the preparation of the manuscript. This research was supported by the German Research Foundation (DFG) grant FR 1471/9-2 awarded to Peter A. Frensch.

REFERENCES

Acosta, E. (1982). Subjective and objective determinants of expectancy: similarities and differences. *Am. J. Psychol.* 95, 139–160.

Alpay, G., Goerke, M., and Stürmer, B. (2009). Precueing imminent conflict does not override sequence-dependent interference adaptation. *Psychol. Res.* 73, 803–816.

Bernstein, I. H., and Reese, C. (1965). Behavioral hypotheses and choice time reaction. *Psychon. Sci.* 3, 259–260.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., and Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychol. Rev.* 108, 624–652.

Donkin, C., Brown, S. D., and Heathcote, A. (2009). ChoiceKey: a real-time speech recognition program for psychology experiments with a small response set. *Behav. Res. Methods* 41, 154–162.

Duthoo, W., De Baene, W., Wühr, P., and Notebaert, W. (2012). When predictions take control: the effect of task predictions on task switching performance. *Front. Psychol.* 3, 282. doi:10.3389/fpsyg.2012.00282

Fitts, P. M., Peterson, J. R., and Wolpe, G. (1963). Cognitive aspects of information processing: II. Adjustments to stimulus redundancy. *J. Exp. Psychol.* 65, 423–432.

Gaissmaier, W., and Schooler, L. J. (2008). The smart potential behind probability matching. *Cognition* 109, 416–422.

Gallistel, C. R. (2005). Deconstructing the law of effect. *Games Econ. Behav.* 52, 410–423.

Gaschler, R., and Frensch, P. A. (2007). Is information reduction an item-specific or an item-general process? *Int. J. Psychol.* 42, 218–228.

Gaschler, R., and Frensch, P. A. (2009).When vaccinating against information reduction works and when it does not work. *Psychol. Stud.* 54, 43–53.

Hommel, B., Müsseler, J., Aschersleben, G., and Prinz, W. (2001). The theory of event coding (TEC): a framework for perception and action planning. *Behav. Brain Sci.* 24, 849–878.

Kemper, M., Umbach, V. J., Schwager, S., Gaschler, R., Frensch, P. A., and Stürmer, B. (2012). Stronger effects of self-generated vs. cue-induced expectations in event-related potentials. (Manuscript submitted for publication).

Kunde, W., Elsner, K., and Kiesel, A. (2007). No anticipation-no action: the role of anticipation in action and perception. *Cogn. Process.* 8, 71–78.

Libet, B., Gleason, C. A., Wright, E. W., and Pearl, D. K. (1983). Time of conscious intention to act in relation to onset of cerebral activity (readiness potential): the unconscious initiation of a freely voluntary act. *Brain* 106, 623–642.

Mattler, U. (2004). Combined expectancy effects are modulated by the relation between expectancy cues. *Q. J. Exp. Psychol.* 57, 193–221.

Miller, J., and Anbar, R. (1981). Expectancy and frequency effects on perceptual and motor systems in choice reaction time. *Mem. Cognit.* 9, 631–641.

Notebaert, W., Houtman, F., Van Opstal, F., Gevers, W., Fias, W., and Verguts, T. (2009). Post-error slowing: an orienting account. *Cognition* 111, 275–279.

Perlman, A., and Tzelgov, J. (2006). Interactions between encoding and retrieval in the domain of sequence-learning. *J. Exp. Psychol. Learn. Mem. Cogn.* 32, 118–130.

Perruchet, P., Cleeremans, A., and Destrebecqz, A. (2006). Dissociating the effects of automatic activation and explicit expectancy on reaction times in a simple associative learning task. *J. Exp. Psychol. Learn. Mem. Cogn.* 32, 955–965.

Posner, M. I., and Petersen, S. E. (1990). The attention system of the human brain. *Annu. Rev. Neurosci.* 13, 25–42.

Posner, M. I., and Snyder, C. R. R. (1975). "Facilitation and inhibition in the processing of signals," in *Attention and Performance V*, eds P. M. A. Rabbitt and S. Domic (New York: Academic Press), 669–682.

Rosenbaum, D. A., and Kornblum, S. (1982). A priming method for investigating the selection of motor responses. *Acta Psychol. (Amst.)* 51, 223–243.

Scheibe, C., Schubert, R., Sommer, W., and Heekeren, H. R. (2009). Electrophysiological evidence for the effect of prior probability on response preparation. *Psychophysiology* 46, 758–770.

Schultz, W., Dayan, P., and Montague, P. R. (1997). A neural substrate of prediction and reward. *Science* 14, 1593–1599.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell Syst. Tech. J.* 27, 379–423, 623–656.

Skinner, B. F. (1950). Are theories of learning necessary? *Psychol. Rev.* 57, 193–216.

Soon, C. S., Brass, M., Heinze, H.-J., and Haynes, J.-D. (2008). Unconscious determinants of free decisions in the human brain. *Nat. Neurosci.* 11, 543–545.

Sutton, R. S., and Barto, A. G. (1981). Toward a modern theory of adaptive networks: expectation and prediction. *Psychol. Rev.* 88, 135–170.

Touron, D. R., and Hertzog, C. (2004). Strategy shift affordance and strategy choice in young and older adults. *Mem. Cognit.* 32, 298–310.

Trevena, J., and Miller, J. (2010). Brain preparation before a voluntary action: evidence against unconscious movement initiation. *Conscious. Cogn.* 19, 447–456.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 15 June 2012; accepted: 12 September 2012; published online: 01 October 2012.

*Citation: Umbach VJ, Schwager S, Frensch PA and Gaschler R (2012) Does explicit expectation really affect preparation? Front. Psychology **3**:378. doi: 10.3389/fpsyg.2012.00378*

This article was submitted to Frontiers in Cognition, a specialty of Frontiers in Psychology.

frontiers in
PSYCHOLOGY

OPINION ARTICLE
published: 11 July 2012
doi: 10.3389/fpsyg.2012.00236

Temporal orienting and alerting – the same or different?

Noam Weinbach* and Avishai Henik

Department of Psychology and Zlotowski Center for Neuroscience, Ben-Gurion University of the Negev, Beer-Sheva, Israel
**Correspondence: noam.weinbach@gmail.com*

Edited by:
Rico Fischer, Technische Universität Dresden, Germany

Reviewed by:
Andrea Kiesel, Julius-Maximilians-University, Germany

Attention helps regulate what to attend to and what to filter out. A warning cue prior to an event can be used to direct attention and improve performance when response to an imperative target is required. Various studies have suggested that warning cues may induce a change in alertness or modulate temporal anticipation of an upcoming event. The current literature presents similar effects for these two functions; hence, effects of a warning cue are sometimes attributed to changes in the state of alertness and in other cases, to voluntary orienting of attention in time. In this article we will discuss whether temporal orienting of attention and alertness are dissociable.

TEMPORAL ORIENTING

Spatial orienting of attention (Posner et al., 1980) has been studied for many years. Recently, studies demonstrated the ability to flexibly and voluntarily orient attention to moments in time – temporal orienting of attention (Coull and Nobre, 1998; for review see Correa, 2010). Many studies on voluntary temporal orienting present symbolic warning cues prior to a target that predict, with high probability, the specific time of target onset. For example, a red rectangle can be used to predict with 75% chance that the target will appear shortly, following 400 ms, and a green rectangle can be used to predict with 75% chance that the target will appear later – following 1,300 ms. The cues are considered valid when the target appears at the predicted time (i.e., 75% of the trials), and invalid when the target appears at a temporally unexpected time (i.e., 25% of the trials). Reaction times (RTs) are faster following valid cues compared with invalid cues. Another method used to study temporal preparation is manipulating the time interval between the warning cue and the target (i.e., foreperiod). For example, when using a constant foreperiod in a block of trials (i.e., within a block the target always appears following the same foreperiod), RTs will be faster for a shorter foreperiod block (e.g., 800 ms) compared with a long foreperiod block (e.g., 2,000 ms, see Rolke and Hofmann, 2007, for a typical study). The time between the cue and target allows top-down temporal preparation to develop, but it will be less accurate as time is prolonged. In contrast, when different foreperiods are intermixed within a block, expectancy builds up as time elapses and performance will be better at later SOA's, a phenomenon called "*the foreperiod effect*" (see Niemi and Näätänen, 1981).

ALERTNESS

Alertness is considered by some researchers as an attentional system that helps regulate the intensity of attention to given stimuli (Posner and Petersen, 1990; Sturm et al., 1999). Effects of alertness are attributed to a high state of arousal for a short period of time following an abrupt external event (i.e., phasic alertness). Most studies use neutral warning cues (i.e., task-irrelevant) prior to a target to induce a state of alertness. Faster RTs are observed following these cues compared with a no-cue condition, in which arousal is low. It was argued that a warning cue that elevates alertness has the optimal influence on performance at a foreperiod of 500 ms (Posner and Boies, 1971). Most studies on alerting cues use foreperiods that range roughly between 100–800 ms. Some authors use the term "accessory stimuli" rather than warning cues if the foreperiod is less than 500 ms (Hackley et al., 2009).

THE PROBLEM

Although neutral warning signals do not necessarily predict the exact onset time of the target, they may still trigger temporal expectation by indicating that a target will appear shortly. On the other hand, temporal orienting cues, which trigger voluntary modulation of attention in time, also involve, to some extent, a change in the state of alertness (Correa et al., 2004). Since both processes can be triggered by a single cue, there is difficulty in assessing to which extent the effects following a warning cue reflect benefit due to bottom-up arousal or are due to top-down temporal expectancy. It makes sense that the shorter the interval between the warning cue and the target, the less likely it is for top-down processes to develop. However, how short is "short"? Studies on phasic alerting show that alerting cues can reduce RTs even at a foreperiod of 100 ms (e.g., Fernandez-Duque and Posner, 1997). Some argue that at foreperiods below 500 ms, temporal expectancy cannot build up (Hackley et al., 2009). However, other studies on temporal preparation challenge this view and report temporal preparation effects even at foreperiods of 200 ms when manipulating temporal contingencies (Thomaschke et al., 2011) or 300 ms for effects of symbolic temporal cues (Coull and Nobre, 1998). Clearly, there is an overlap in the time course of temporal preparation and alerting.

In addition to the methodological difficulty in dissociating these effects, the literature can be sometimes confusing when considering the definitions different researchers use for alerting and temporal orienting. In fact, some authors define alerting basically in the same way as temporal orienting is defined. For example, "… *alerting is the ability to make use of a cue which provides information about the onset time of a target stimulus, and thus triggers the allocation of attention at a given point in time*" (Dye et al., 2009, p. 1780). Others use a more general definition for alerting, which does not necessarily consider temporal expectancy. For example, "… *the ability to increase response readiness for a short* *period of time subsequent to external cues or stimuli (phasic alertness)*" (Sturm and Willmes, 2001, p. S76).

Because of the overlap in definitions and due to the overlap in time-course of the effects as mentioned above, it is not surprising that many studies report significantly similar findings regarding the behavioral and neuronal features of the two processes. These findings are attributed to alerting in some studies and in others, to temporal orienting.

COMPARING SOME OF THE EFFECTS OF ALERTING AND TEMPORAL ORIENTING

Both alerting and temporal orienting cues usually produce faster motor execution of response to an imperative target compared with no-cue or invalid temporal cue conditions, respectively. A major question was whether the source of faster execution of response could be attributed only to motor preparation, or was there also a change in early perceptual and response selection processing stages? Event-related potential (ERP) studies have demonstrated that both alerting and voluntary temporal orienting modulate similar components that are related to early processing stages such as perceptual and response selection, rather than just late motor preparation (for alerting see Hackley and Valle-Inclán, 1998; Böckler et al., 2011; for temporal orienting see Correa et al., 2006a; Lange et al., 2006). This is also supported by behavioral studies demonstrating that temporal preparation can improve perceptual processing by operating at the onset of sensory information accumulation, facilitating perceptual discrimination, improving perceptual sensitivity and discrimination accuracy (e.g., Correa et al., 2005, 2006b; Rolke and Hofmann, 2007; Rolke, 2008; Seibold et al., 2011). Phasic alerting has also been found to increase perceptual processing speed, improve conscious perception, and bias perceptual processing (Matthias et al., 2010; Kusnir et al., 2011; Weinbach and Henik, 2011; Finke et al., 2012).With regard to response selection processes, both alerting and temporal orienting have been found to have similar effects. For example, alerting cues have been suggested to increase response conflict due to increased activation of the stimulus-response link (Fischer et al., 2010, 2012). Similarly, valid temporal cues have been suggested to disrupt response selection by automatically activating competing responses (Correa et al., 2010).

From a neurophysiological aspect, both the effects of alerting and of temporal cueing can be reduced by drugs such as Clonidine, which reduce norepinephrine (NE) release (Coull et al., 2001). However, the brain activity that accompanied the use of Clonidine in alerting vs. temporal orienting cues did not overlap, and it was argued that modulation of the alerting effect by Clonidine is unlikely to be due to an underlying effect on temporal orienting processes (Coull et al., 2001).

In addition, when reviewing the literature on the neural correlates of temporal orienting cues and alerting cues, a clear dissociation between them is also somewhat difficult to find. Studies show that both alerting cues and temporal orienting cues are associated with similar regional activity in the left hemisphere (Coull et al., 2001; Fan et al., 2005). This led Coull et al. to conclude that "… *alerting effect primarily indexes temporal orienting and motor preparation, rather than arousal or phasic alertness*" (p. 81). However, it is important to note that there is an ongoing debate regarding the lateralization of alerting and some uncertainties remain (Petersen and Posner, 2012).

In summary, all of the above are examples that could indicate that alerting and temporal preparation actually represent the same preparation process.

DISSOCIATING ALERTING AND TEMPORAL ORIENTING

The difficulty in dissociating the effects of alerting and temporal orienting could be because they both actually reflect the same process of preparation. Alternatively, they could represent different processes that function similarly. There are two main problems in dissociating alerting and temporal orienting; one is methodological and the second is their definitions. Methodologically, both processes can be triggered by the same warning cue and overlap in time-course. Regarding their definitions, different authors sometimes define both processes in a similar way and this could lead to similar operationalization. However, it is important to note that there is evidence that these processes are dissociable.

At first, a reduction in RTs following warning cues is observed even when the cues are not temporally predictive. When making sure that following the warning cue, there is equal probability for the target to appear at each foreperiod (a technique called "*non-aging foreperiods*," see Niemi and Näätänen, 1981) there is still reduction in RTs following the warning cue (Whitehead, 1991). This benefit cannot be understood by reduction of uncertainty regarding the temporal onset of the target (Fernandez-Duque and Posner, 1997). In addition, arousing cues (i.e., accessory stimuli) that are presented concurrently or even following the target still produce a benefit in RTs (e.g., Stahl and Rammsayer, 2005; Kiesel and Miller, 2007), even though there is no temporal preparation in this situation. In a more recent study, Hackley et al. (2009) reported a dissociation between phasic alerting and temporal expectancy after showing that alerting cues still induce benefit in RTs even when participants know in advance exactly when the target will appear, making the alerting cues completely task-irrelevant. Even more recently, Lawrence and Klein (in press) offered a clever methodological solution for examining the pure effect of alerting by demonstrating the benefit of these cues in a block of trials where there were absolutely no contingencies between the alerting signals and the target.

These examples of dissociations between the two processes lead to the conclusion that alerting and temporal orienting represent different processes.

CONCLUSION AND PRACTICAL SUGGESTIONS

We suggest a distinction between temporal orienting, in which temporal information is inherent in the cues, and arousal, which does not depend on temporal contingencies. Both types of preparation can be achieved voluntarily or automatically. Voluntary temporal orienting is best reflected by tasks using symbolic temporal cues (Coull and Nobre, 1998). Automatic temporal orienting can be observed following regular rhythms that orient attention in time involuntarily (Rohenkohl et al., 2011). Automatic arousal is reflected in phasic alertness, which has the largest effect at short foreperiods and can occur independently of temporal contingencies. Voluntary arousal is what authors name "tonic alertness," meaning the general ability to stay alert and prepared for detecting infrequent stimuli during a task (usually measured in vigilance and continuous performance tasks).

The distinction between arousal and temporal orienting should be taken into consideration when studying temporal preparation and alerting (arousal) because these processes are commonly confounded in most experimental designs. In order to examine one process only, it is necessary to control for the irrelevant process. Researchers can adopt techniques such as non-aging foreperiod distribution in order to control strategic temporal expectancy processes following the alerting cue. Note that Lawrence and Klein (in press) have recently suggested another methodology in order to reveal a pure effect of alerting cues.

Studies on cued temporal orienting can include non-informative neutral cues, which can be considered as the baseline arousal level, and compare the results achieved with these cues to those with valid cueing (i.e., high temporal expectation) and invalid cueing (i.e., low temporal expectation; see similar procedure in Coull and Nobre, 1998; Coull et al., 2001).

REFERENCES

Böckler, A., Alpay, G., and Stürmer, B. (2011). Accessory stimuli affect the emergence of conflict, not conflict control: a Simon-task ERP study. *Exp. Psychol.* 58, 102–109.

Correa, Á. (2010). "Enhancing behavioural performance by visual temporal orienting," in *Attention and Time*, eds A. C. Nobre and J. T. Coull (Oxford: Oxford University Press), 357–370.

Correa, Á., Cappucci, P., Nobre, A. C., and Lupiáñez, J. (2010). The two sides of temporal orienting. *Exp. Psychol.* 57, 142–148.

Correa, Á., Lupiáñez, J., Madrid, E., and Tudela, P. (2006a). Temporal attention enhances early visual processing: a review and new evidence from event-related potentials. *Brain Res.* 1076, 116–128.

Correa, Á., Sanabria, D., Spence, C., Tudela, P., and Lupiáñez, J. (2006b). Selective temporal attention enhances the temporal resolution of visual perception: evidence from a temporal order judgment task. *Brain Res.* 1070, 202–205.

Correa, Á., Lupiáñez, J., Milliken, B., and Tudela, P. (2004). Endogenous temporal orienting of attention in detection and discrimination tasks. *Atten. Percept. Psychophys.* 66, 264–278.

Correa, Á., Lupiáñez, J., and Tudela, P. (2005). Attentional preparation based on temporal expectancy modulates processing at the perceptual level. *Psychon. Bull. Rev.* 12, 328–334.

Coull, J., and Nobre, A. (1998). Where and when to pay attention: the neural systems for directing attention to spatial locations and to time intervals as revealed by both PET and fMRI. *J. Neurosci.* 18, 7426–7435.

Coull, J., Nobre, A., and Frith, C. (2001). The noradrenergic alpha 2 agonist clonidine modulates behavioural and neuroanatomical correlates of human attentional orienting and alerting. *Cereb. Cortex* 11, 73–84.

Dye, M., Green, C., and Bavelier, D. (2009). The development of attention skills in action video game players. *Neuropsychologia* 47, 1780–1789.

Fan, J., McCandliss, B. D., Fossella, J., Flombaum, J. I., and Posner, M. I. (2005). The activation of attentional networks. *Neuroimage* 26, 471–479.

Fernandez-Duque, D., and Posner, M. I. (1997). Relating the mechanisms of orienting and alerting. *Neuropsychologia* 35, 477–486.

Finke, K., Matthias, E., Keller, I., Müller, H. J., Schneider, W. X., and Bublak, P. (2012). How does phasic alerting improve performance in patients with unilateral neglect? A systematic analysis of attentional processing capacity and spatial weighting mechanisms. *Neuropsychologia* 50, 1178–1188.

Fischer, R., Plessow, F., and Kiesel, A. (2010). Auditory warning signals affect mechanisms of response selection: evidence from a Simon task. *Exp. Psychol.* 57, 89–97.

Fischer, R., Plessow, F., and Kiesel, A. (2012). The effects of alerting signals in action control: activation of S–R associations or inhibition of executive control processes? *Psychol. Res.* 76, 317–328.

Hackley, S. A., Langner, R., Rolke, B., Erb, M., Grodd, W., and Ulrich, R. (2009). Separation of phasic arousal and expectancy effects in a speeded reaction time task via fMRI. *Psychophysiology* 46, 163–171.

Hackley, S. A., and Valle-Inclán, F. (1998). Automatic alerting does not speed late motoric processes in a reaction-time task. *Nature* 391, 786–788.

Kiesel, A., and Miller, J. (2007). Impact of contingency manipulations on accessory stimulus effects. *Atten. Percept. Psychophys.* 69, 1117–1125.

Kusnir, F., Chica, A. B., Mitsumasu, M. A., and Bartolomeo, P. (2011). Phasic auditory alerting improves visual conscious perception. *Conscious. Cogn.* 20, 1201–1210.

Lange, k., Krämer, U. M., and Röder, B. (2006). Attending points in time and space. *Exp. Brain Res.* 173, 130–140.

Lawrence, M. A., and Klein, R. M. (in press). Isolating exogenous and endogenous modes of temporal attention. *J. Exp. Psychol. Gen.*

Matthias, E., Bublak, P., Müller, H. J., Schneider, W. X., Krummenacher, J., and Finke, K. (2010). The influence of alertness on spatial and nonspatial components of visual attention. *J. Exp. Psychol. Hum. Percept. Perform.* 36, 38–56.

Niemi, P., and Näätänen, R. (1981). Foreperiod and simple reaction time. *Psychol. Bull.* 89, 133.

Petersen, S. E., and Posner, M. I. (2012). The attention system of the human brain: 20 years after. *Annu. Rev. Neurosci.* 35, 73–89.

Posner, M. I., and Boies, S. J. (1971). Components of attention. *Psychol. Rev.* 78, 391–408.

Posner, M. I., and Petersen, S. E. (1990). The attention system of the human brain. *Annu. Rev. Neurosci.* 13, 25–42.

Posner, M. I., Snyder, C. R., and Davidson, B. J. (1980). Attention and the detection of signals. *J. Exp. Psychol. Gen.* 109, 160–174.

Rohenkohl, G., Coull, J. T., and Nobre, A. C. (2011). Behavioural dissociation between exogenous and endogenous temporal orienting of attention. *PLoS ONE* 6, e14620. doi: 10.1371/journal.pone.0014620

Rolke, B. (2008). Temporal preparation facilitates perceptual identification of letters. *Atten. Percept. Psychophys.* 70, 1305–1313.

Rolke, B., and Hofmann, P. (2007). Temporal uncertainty degrades perceptual processing. *Psychon. Bull. Rev.* 14, 522–526.

Seibold, V. C., Bausenhart, K. M., Rolke, B., and Ulrich, R. (2011). Does temporal preparation increase the rate of sensory information accumulation? *Acta Psychol. (Amst.)* 137, 56–64.

Stahl, J., and Rammsayer, T. H. (2005). Accessory stimulation in the time course of visuomotor information processing: stimulus intensity effects on reaction time and response force. *Acta Psychol. (Amst.)* 120, 1–18.

Sturm, W., Simone, A., Krause, B., Specht, K., Hesselmann, V., Radermacher, I., Herzog, H., Tellmann, L., Muller-Gartner, H. W., and Willmes, K. (1999). Functional anatomy of intrinsic alertness: evidence for a fronto-parietal-thalamic-brainstem network in the right hemisphere. *Neuropsychologia* 37, 797–805.

Sturm, W., and Willmes, K. (2001). On the functional neuroanatomy of intrinsic and phasic alertness. *Neuroimage* 14, S76–S84.

Thomaschke, R., Wagener, A., Kiesel, A., and Hoffmann, J. (2011). The scope and precision of specific temporal expectancy: evidence from a variable fore period paradigm. *Atten. Percept. Psychophys.* 73, 953–964.

Weinbach, N., and Henik, A. (2011). Phasic alertness can modulate executive control by enhancing global processing of visual stimuli. *Cognition* 121, 454–458.

Whitehead, R. (1991). Right hemisphere processing superiority during sustained visual attention. *J. Cogn. Neurosci.* 3, 329–334.

Received: 17 May 2012; accepted: 21 June 2012; published online: 11 July 2012

Citation: Weinbach N and Henik A (2012) Temporal orienting and alerting – the same or different? Front. Psychology ***3****:236. doi: 10.3389/fpsyg.2012.00236*

This article was submitted to Frontiers in cognition, a specialty of Frontiers in psychology

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 28 August 2012
doi: 10.3389/fpsyg.2012.00308

Temporal preparation driven by rhythms is resistant to working memory interference

María Dolores de la Rosa, Daniel Sanabria*, Mariagrazia Capizzi and Angel Correa*

Departamento de Psicología Experimental, Universidad de Granada, Granada, Spain

Edited by:
Rico Fischer, Technische Universität Dresden, Germany

Reviewed by:
Noam Weinbach, Ben-Gurion University of the Negev, Israel
Roland Thomaschke, University of Regensburg, Germany

****Correspondence:***
Daniel Sanabria and Angel Correa, Facultad de Psicología, Universidad de Granada, Campus Universitario de Cartuja s/n, 18071 Granada, Spain.
e-mail: act@ugr.es; daniel@ugr.es

It has been recently shown that temporal orienting demands controlled attention (Capizzi et al., 2012). However, there is current debate on whether temporal preparation guided by regular rhythms also requires the generation of endogenous temporal expectancies or rather involves a mechanism independent of executive control processes. We investigated this issue by using a dual-task paradigm in two different experiments. In Experiment 1, the single-task condition measured reaction time to respond to the onset of an auditory stimulus preceded by either a regular or an irregular auditory rhythm. The dual-task condition additionally included a working memory task, which demanded mental counting and updating. In Experiment 2, the simultaneously WM task was a variant of the Sternberg Task. We hypothesized that, if temporal preparation induced by rhythms did not involve executive processing, it would not be interfered by the simultaneous working memory task. The results showed that participants could anticipate the moment of target onset on the basis of the regular rhythm and, more important, this ability resisted the interference from the double task condition in both experiments. This finding supports that temporal preparation induced by rhythms, in contrast to temporal orienting, does not require resources of executive control.

Keywords: exogenous attention, reaction times, working memory, temporal orienting, bottom-up, stimulus-driven, dual-task

INTRODUCTION

Temporal preparation consists of the ability to direct attention to a point in time when a relevant event is expected (Coull and Nobre, 1998). The environment provides us with temporal information such as symbolic cues or temporal regularity of certain events (i.e., rhythms), which we can use to build up temporal expectations about stimulus onset and prepare an optimized response at the appropriate moment in time.

Recent studies have investigated the nature of the mechanisms involved in temporal preparation with the aim of dissociating between exogenous and endogenous components. On the one hand, endogenous temporal preparation ("temporal orienting of attention") depends on the expectations built on predictive temporal information given explicitly by symbolic cues and used to voluntary prepare the response at the expected time. It has previously been related to processes of controlled nature (Coull and Nobre, 1998; Capizzi et al., 2012). On the other hand, it has been shown that temporal preparation can be induced bottom-up, by the temporal regularities provided by regular sequences of stimuli (i.e., rhythms). Regular rhythms would orient our attentional resources in time without the implication of endogenous temporal expectancies, which is reflected by enhanced accuracy and/or faster response to target stimuli (Jones et al., 2002; Sanabria et al., 2011). A relevant issue in research on temporal preparation is to determine the similarities and differences between these two ways to orient attention within the temporal domain.

ENDOGENOUS TEMPORAL PREPARATION

Coull and Nobre (1998), based on the Cost and Benefits paradigm (Posner et al., 1980), developed a temporal orienting task adapted to study how attention can be oriented to specific points in time. The procedure consists of a symbolic cue, which explicitly indicates with high probability the time interval or foreperiod (e.g., "*early*" at 400 ms of cue onset, or "*late*" at 1400 ms) at which the target stimuli will occur. For instance, in 75% of trials the temporal cue indicated correctly the moment of target occurrence (i.e., valid trials), whereas in the remainder of trials the target appeared either before or after that cued time (i.e., invalid trial). The results typically show faster reaction times (RTs) in valid relative to invalid trials, mainly at the short foreperiod, which is known as "temporal orienting effect." This effect is usually reduced or absent at the long foreperiod (see Correa et al., 2004).

Previous research has supported the involvement of controlled processes in temporal orienting. Capizzi et al. (2012) showed that the temporal orienting effect diminished significantly in demanding dual-task conditions. Nevertheless, sequential effects (i.e., faster RTs when the previous interval had either the same or shorter duration than the current interval) did not show any modulation by performing a simultaneous working memory task. It was concluded that temporal orienting involved controlled processing, which was affected by competition for executive resources demanded by the concurrent task. Sequential effects, associated to automatic processing (Los, 1996; Los and Van den Heuvel, 2001;

Vallesi and Shallice, 2007; Vallesi et al., 2007), resisted the dual-task interference.

EXOGENOUS TEMPORAL PREPARATION

Temporal preparation can also be induced by temporal regularities of certain events. It has been shown that the presentation of regular sequences of auditory stimulus (i.e., rhythms) enhanced the performance in a pitch discrimination task when the target tone appeared at a time point corresponding to the temporal pattern of the sequence (Jones et al., 2002; see also Lange, 2010). Moreover, cuing time by means of rhythms speeded up responses to a relevant stimulus when it appeared at the moment in time matching the rhythm's pace (Sanabria et al., 2011). Jones and colleagues have suggested that rhythms induce automatic temporal preparation, since regular repetitions of tone onsets would synchronize the internal attending activity producing an improved response when target stimulus onset continues the rhythmic pattern (Barnes and Jones, 2000).

Rohenkohl et al. (2011) provided further evidence supporting the involvement of exogenous processes in temporal preparation guided by rhythms. They compared temporal preparation guided by rhythms with temporal preparation guided by symbolic cues to dissociate between exogenous and endogenous processes of temporal preparation. Specifically, participants performed a task consisting of a ball moving across the screen until reaching an occluding band. When the ball reappeared, participants were required to discriminate whether the target contained an upright or tilted cross. Participants could predict the moment of target occurrence by means of either the rhythm (i.e., the ball moved following a constant speed, regular rhythmic pace) or the meaning of the symbolic cue (i.e., the color of the ball predicted the duration of the occlusion). At the beginning of the task, participants were instructed to attend to either the rhythm or the symbolic cue to predict the target onset. Results showed that temporal regularity of rhythms enhanced responses to the target regardless of the instructions received by participants ("attend to color" or "attend to speed"). However, the effect of symbolic cues depended on the instruction to attend to color, that is, symbolic cuing was only effective in the "attend to color" but not in the "attend to speed" condition. Therefore, these findings suggested dissociation between temporal preparation driven by rhythms and temporal preparation guided by symbolic cues.

A recent neuropsychological study (Triviño et al., 2011) has shown that patients with right frontal damage could orient attention in time by means of regular rhythms, whereas deficit was observed when symbolic cues were presented. Triviño et al.'s (2011) findings further suggest that temporal preparation guided by rhythms does not depend on the endogenous building up of temporal expectancies. Thus, temporal preparation induced by rhythms would involve a more exogenous bottom-up process, such that it would not depend on the functioning of right prefrontal structures related to attentional control.

In contrast, an event-related potentials (ERPs) study (Schwartze et al., 2011) has questioned the sole involvement of exogenous bottom-up processes in temporal preparation guided by rhythms. Schwartze et al. (2011) used an auditory oddball paradigm to investigate whether regularity of rhythms influenced automatic processing (as indexed by the mismatch negativity – MMN – potential) or "attention-dependent" processing (as indexed by the P3b potential), in two sessions, "pre-attentive" and "attentive." In both sessions, the auditory sequence could be formed by either a regular or an irregular rhythm (i.e., isochronous vs. random temporal structure). In the pre-attentive session, participants had to watch a video clip while listening to an auditory rhythm that should be ignored. In the attentive session, participants should concentrate on the rhythm and count the deviant tones in each auditory sequence. The results showed that regular rhythms modulated the attention-related potential (P3b) in the attentive session, while in the pre-attentive session, the automatic processing potential (MMN) was not influenced by the rhythm. Schwartze et al. (2011) concluded that synchronization of attention by rhythms required the involvement of top-down mechanisms, such that the influence of temporal regularity was dependent on top-down attentional processing rather than on bottom-up automatic processing. This result differed from previous research suggesting that temporal preparation driven by rhythms involves exogenous bottom-up processing, since it was not necessary to attend to rhythms to orient attention in time (Rohenkohl et al., 2011; Sanabria et al., 2011, Experiment 3).

To summarize, while there is agreement on the controlled nature of the endogenous temporal preparation driven by symbolic cues, it remains currently unclear whether exogenous temporal preparation driven by regular rhythms involves only bottom-up mechanisms or it requires the development of endogenous temporal expectancies. In order to clarify this issue, in the present study, we used a dual-task paradigm to compare the effects of temporal preparation guided by rhythms between a single-task condition and a dual-task condition.

A dual-task paradigm requires performing two tasks (primary and secondary task) simultaneously. In our study, the single-task condition consisted of a temporal preparation task, in which the time of target onset was cued by means of auditory rhythms similar to Lange (2010) who found faster RTs when the target was preceded by a regular rather than by an irregular sequence of six tones. In Lange's study both sequences of tones had the same duration. Equating duration of both sequences was crucial for the dual-task condition of our study, where, in addition to performing a RT task, participants had to perform simultaneously a WM task. Once the memory retention interval was the same for both rhythm conditions, we could measure the effect of rhythm by comparing the two conditions that are balanced in terms of memory demands. In Experiment 1, the WM task required to count and remember how many times a colored fixation point (from three different colors, changing on a trial-by-trial basis) was presented along a block of trials. In Experiment 2, we aimed to replicate Experiment 1 by using a WM task based on Sternberg's memory scanning paradigm, in which participants had to remember a new a sequence of six letters in every trial. If performance on the primary task was affected by the secondary task, it could be assumed that both tasks competed for common limited endogenous resources (cf. Posner and Snyder, 1975; Logan, 1978). If temporal preparation driven by rhythms was independent of resources of executive control, its effect on RT would not be affected by performing concurrently the WM task.

EXPERIMENT 1

METHOD

Participants

Thirty-three undergraduate students (31 females; age range: 19–43 years old; mean age: 22.66 years old) from the Faculty of Psychology of the Universidad de Granada took part in Experiment 1 in exchange of course credits. Participants were randomly assigned to two groups according to task conditions, 16 participants completed the single-task condition and 17 the dual-task condition. In the last condition, one participant was eliminated due to poor performance (22% of correct responses) in the working memory task.

Apparatus and stimuli

Experiment 1 was run on an Intel Core 2 Duo connected to a 17″ LCD monitor. The E-prime software (Schneider et al., 2002) was used for stimulus presentation and to record participants' responses. The viewing distance was approximately 60 cm. Both single-task and dual-task conditions shared the same auditory and visual stimuli appearing in the center of the screen. In both conditions, the temporal preparation task consisted of a sequence of six tones with duration of 250 ms each and a frequency of 700 Hz. This sequence could be temporally regular or irregular. In the regular sequence, the interval between tones was 550 ms. In the irregular sequence the duration of each interval could be either 150, 350, 550, 750, or 950 ms. The order of these five intervals was randomized across trials. Both sequences included the same number of tones (six tones) and had identical duration, therefore, the only difference concerned temporal regularity or irregularity. The target tone was a 100-ms sound of 400 Hz (**Figure 1**).

At the beginning of the trial, a plus sign (1.5° × 1.5°) appeared either in red, blue, or green, chosen at random for each trial with the same probability of appearance. All visual stimuli were presented on a black background in the center of the screen.

Procedure and task

Both verbal and written instructions were given to participants, who had to press the b key as fast as possible when the target tone was presented. Moreover, they were informed that, before the target, a sequence of sounds forming a rhythm would be presented, which was irrelevant for the task and should therefore be ignored.

Each task condition consisted of one practice block and eight experimental blocks composed of 16 trials each. In both task conditions, each trial began with the presentation of a black screen for 50 ms. Next, a plus sign filled with one of the three colors (red, blue, and green) was randomly generated and remained present during the trial. Then, 500 ms after the plus sign's appearance, a regular or irregular rhythm was presented at random. The rhythm was followed by the target tone that appeared after a foreperiod of variable duration (800, 1100, and 1400 ms) that was generated at random for each trial. Each foreperiod had a different probability of occurrence based on a non-aging distribution. It consisted of increasing the frequency of the shorter foreperiod such that the conditional probability for target appearance remained constant through the trial. The target tone appeared at the 800-ms foreperiod, in 50% of the trials, at the 1100 ms foreperiod, in 25% of the trials and at the 1400-ms foreperiod, in 12.5% of the trials. In the remaining 12.5% of trials, the target tone was not presented (catch trials). Participants had a maximum of 1150 ms to respond and in

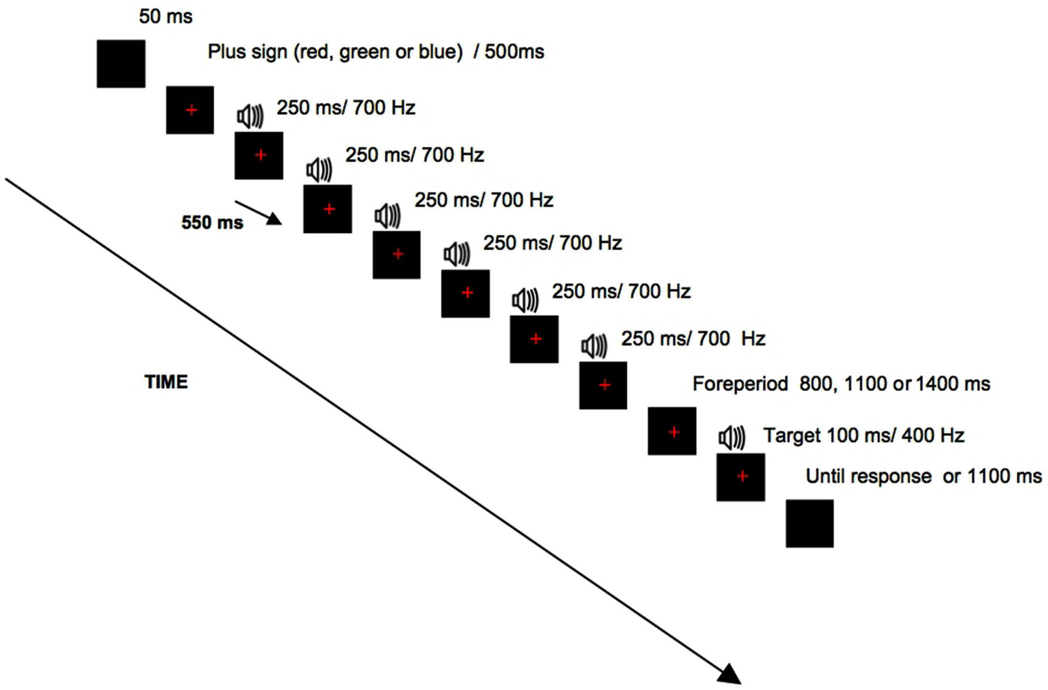

FIGURE 1 | Schematic representation of events in a trial in the regular rhythm. In the irregular rhythm the duration of each interval varied and could be either 150, 350, 550, 750, or 950 ms.

case of responding before target onset, a message provided visual feedback on anticipatory error.

In the dual-task condition, the procedure was similar to the single-task condition, except for that participants should perform simultaneously a WM task. The WM task consisted of remembering how many times each color appeared during a block of trials. At the end of each block, participants should type how many times a certain color (e.g., "*green*") had been presented. Each color was selected at random and with the same probability for the memory test. When participants responded, a message provided feedback about memory accuracy. The word "correct" or "incorrect" filled in green and red color respectively, was presented for 1500 ms.

Design and data analysis

The Experiment 1 constituted a 2 × 3 × 2 design with independent variables of Rhythm (regular and irregular) and Foreperiod (800, 1100, and 1400 ms) as within participants factor and Task (single and double) as a between participants factor.

Practice trials, premature responses (i.e., participants responded before the target appeared), trials with RT below 150 ms and above 1200 ms (0.14% of trials) were eliminated from the analyses. Participants' mean RTs were analyzed by a repeated-measures analysis of variance (ANOVA).

Results

In the WM task, participants' mean accuracy to the color memory test was 89% (7% SD). In the dual-task condition, RT was analyzed only from correct responses in the memory test, in order to assure that participants were actually engaged in the dual-task condition. The RT from responses in the memory test were not included in analyses.

Mean RTs included in the analyses are detailed for each experimental condition in **Table 1**.

The ANOVA showed a significant main effect of Task, $F(1, 30) = 7.85$, $p < 0.01$, with faster RTs in the single-task condition (347 ms) than in the dual-task condition (422 ms). The main effect of Rhythm was also significant, $F(1, 30) = 51.50$, $p < 0.001$, with faster RTs in the regular rhythm (375 ms) than in the irregular rhythm (395 ms). The most relevant finding was that the effect of rhythm did not rely on task condition, since the interaction between Rhythm and Task was not significant, $F < 1$ (see **Figure 2**). Specifically, the effect of Rhythm was significant in both the single-task, $F(1, 30) = 20.88$, $p < 0.001$, η_p^2 : 0.51, and the dual-task condition, $F(1, 30) = 22.80$, $p < 0.001$, η_p^2 : 0.79.

The main effect of Foreperiod, $F(2, 60) = 7.60$, $p < 0.01$, showed faster RTs in the 1100-ms interval (376 ms) than in the 800-ms (393 ms) and 1400-ms (385 ms) intervals. Planned comparisons indicated a significant difference between the 800 and 1100 ms intervals, $F(1, 30) = 33.67$, $p < 0.001$, whereas the difference between 1100 and 1400 ms intervals was marginally significant, $F(1, 30) = 3.61$, $p = 0.06$. The difference between the 800 and 1400 ms intervals was not significant, $F(1, 30) = 2.50$, $p = 0.12$.

The interaction between Rhythm and Foreperiod was significant, $F(2, 60) = 13.59$, $p < 0.001$ (see **Figure 3**). Planned comparisons between regular and irregular rhythms at each interval, revealed a significant effect of Rhythm in both the 800-ms interval, $F(1, 30) = 74.40$, $p < 0.001$, and the 1100-ms interval $F(1, 30) = 26.12$, $p < 0.001$, but not in the 1400-ms interval, $F < 1$.

Table 1 | Mean RTs for each Foreperiod (800, 1100, 1400 ms), Rhythm (regular, irregular), and Task condition (single-task, dual-task).

	Regular rhythm			**Irregular rhythm**		
	800	**1100**	**1400**	**800**	**1100**	**1400**
Single-task	332 (15)	324 (15)	354 (15)	371 (14)	348 (16)	355 (15)
Dual-task	422 (25)	402 (22)	416 (22)	450 (22)	431 (25)	416 (24)

Values in parentheses are standard error of the mean.

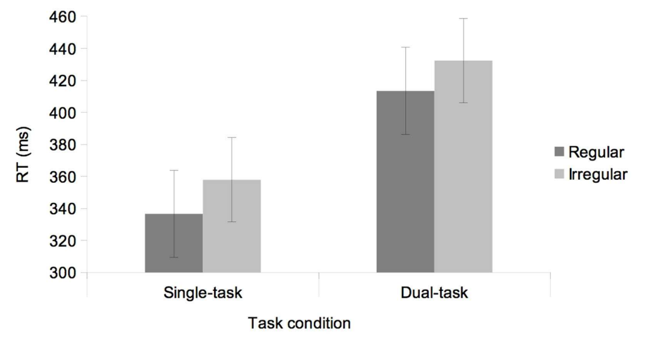

FIGURE 2 | Mean RTs as a function of Rhythm (regular, irregular) and Task condition (single, dual). Error bars represent the standard error of the mean.

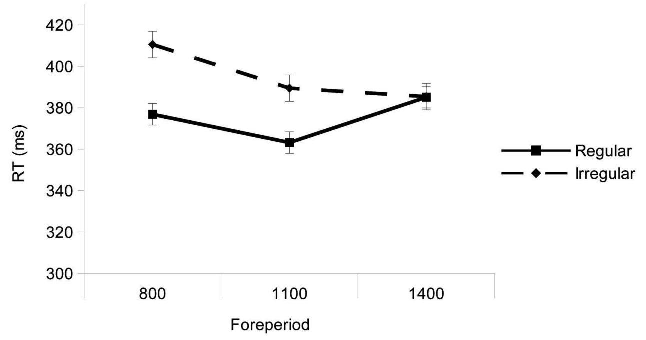

FIGURE 3 | Mean RTs as a function of the Rhythm (regular, irregular) and Foreperiod (800, 1100, and 1400 ms). Error bars represent the standard error of the mean.

Finally, the interaction between Foreperiod and Task also showed a significant result, $F(2, 60) = 3.59$, $p = 0.03$. Further comparisons revealed that RT performance between the two tasks was significant at all foreperiods, $F(1, 30) = 9.35$, $p < 0.01$, $F(1, 30) = 8.39$, $p < 0.01$, and $F(1, 30) = 5.25$, $p = 0.02$, for the 800, 1100, and 1400 ms foreperiods, respectively. Moreover, in the single-task condition, responses were faster in the 1100 interval (335 ms) than both in the 800-ms interval (351 ms), $F(1, 30) = 13.63$, $p < 0.001$, and in the 1400-ms interval (354 ms), $F(1, 30) = 7.58$, $p < 0.01$. The difference between the 800 and

1400 ms intervals was not significant, $F < 1$. In the dual-task condition, planned comparisons showed significant differences between the 800 (435 ms) and 1100 ms (416 ms), $F(1, 30) = 20.38$, $p < 0.001$, and between the 800 and 1400 ms (416 ms) intervals, $F(1, 30) = 6.69$, $p < 0.01$ however, the difference between the 1100 and 1400 ms intervals was not significant, $F < 1$.

DISCUSSION

Experiment 1 confirmed that participants could use rhythms for temporal preparation even though they simultaneously performed a WM task. The effect of rhythm did not depend on WM load, according to a non-significant interaction between Load and Rhythm. Moreover, the main effect of Task showed faster RT in the single-task condition relative to the dual-task condition showing therefore an effective manipulation of the memory task.

In the Rhythm by Foreperiod interaction found in the Experiment 1, temporal preparation guided by rhythms was selective such that RTs were faster in the regular rhythm than in the irregular rhythm condition in the 1100-ms interval, that matched (two steps of) the regular sequence, and in the 800-ms interval (see next paragraph for Discussion on this finding). However, no rhythm effect was found in the 1400-ms interval. The current data diverged from the findings by Lange (2010), where the interaction between rhythm condition and foreperiod showed that participants' responses were faster at the 1400-ms interval, probably due to the variable foreperiod effect (i.e., faster RTs at longer intervals) in her experiment. For this reason, we used a non-aging distribution with catch trials where *a priori* probability of occurrence was larger in the shortest foreperiod, thus holding the same conditional probability of target onset throughout the trial that is, this manipulation increased the uncertainty on the moment of target onset, preventing the foreperiod effect (cf. Sanabria et al., 2011).

The rhythm effect was not restricted to the 1100-ms interval, but it was also observed at the 800-ms interval. This result replicates similar findings of previous studies (Griffin et al., 2001; Sanabria et al., 2011), in which a temporal preparation effect was found at the interval shorter than the inter-onset intervals of the sequence. These results have been interpreted as an anticipatory effect (Griffin et al., 2001), that is, an efficient strategy that would consist of preparing for around the shortest foreperiod and then extending preparation to the following foreperiod. However, this result could be interpreted in two different ways: on the one hand, the irregular rhythm would impair temporal preparation at the 800-ms interval where larger RTs were observed. On the other hand, since the mean duration in both sequences was the same, it would improve temporal preparation at the 1100-ms interval that matched the temporal pattern (i.e., two steps of the mean duration), in both sequences. Future research is required to reveal whether the rhythm effect in the irregular sequence was produced by improvement in the temporal preparation at the 1100-ms interval or impairment at the 800-ms interval.

The Foreperiod by Task interaction reached statistical significance, which revealed significant differences in RT between the 1100-ms foreperiod with respect to the other two foreperiods only in single-task condition. In the dual-task condition, participants responded faster in the 1100-ms foreperiod than in the 800-ms foreperiod, but not faster than in the 1400-ms foreperiod. It would appear then that the response enhancement at the foreperiod matching (two steps) of the rhythm was somehow reduced in the dual-task condition. Capizzi et al. (2012) reported an incremented foreperiod effect in the dual-task condition with respect to the single-task condition. Taken together, Capizzi et al.'s (2012) results and the present findings support the notion that the foreperiod effect results from the action of endogenous temporal preparation, since, in contrast to the rhythm effect, it was affected by the concurrent working memory task. In any case, the foreperiod by task interaction was secondary for the main purpose of our study, and at present, data are not conclusive regarding the nature of the mechanisms involved in the foreperiod effect (Los and Van den Heuvel, 2001; Vallesi and Shallice, 2007; Capizzi et al., 2012).

Although the findings in Experiment 1 suggest the involvement of bottom-up processing (i.e., in opposition to top-down executive control processing) in the temporal preparation driven by rhythms, it is possible that our load manipulation was not optimal to produce strong interference. In the current WM task, memory load was not constant along the block, so that at the beginning the number of colors to be remembered was lower than at the end of the block. Therefore, it would allow paying attention to the rhythms providing an optimal attentional preparation in time. Thus, we designed a new task in Experiment 2 in which the demands of the WM task were the same during the whole block. Specifically, we followed a procedure based on the Sternberg's memory scanning paradigm (Sternberg, 1966). This task consisted of presenting a sequence of six consonant letters (memory list) that participants had to remember during the trial. At the end of the trial, a letter selected at random (probe) was presented and participants had to respond whether that letter was present or absent from the initial memory list. This manipulation involved two task conditions with respect to the memory load: the Low load condition, where the memory list was formed by the same letter, and the High load condition in which the memory list was formed by six different letters. In the case of Experiment 2 the task load conditions were manipulated within participants, thus increasing statistical power to study our main effect to interest.

The RT task consisted of regular and irregular rhythms identical to Experiment 1, but a control condition was further included in which no rhythm was presented (instead, the trial was silent and the only stimulus presented to the participant was the initial fixation point and the last target tone). The idea was to test whether regular and irregular rhythms produced benefits or costs on the temporal preparation based on rhythms.

First, we expected to find that participants could prepare in time by mean of rhythms, showing enhanced RTs in the regular rhythms in comparison to irregular and no rhythm conditions. Moreover, if the secondary Sternberg task did not affect the ability for temporal preparation, this would further suggest that temporal preparation guided by rhythms did not require controlled resources for the generation of endogenous temporal expectancies, specifically meaning that it required other resources than those for WM executive control.

EXPERIMENT 2

METHOD

Participants

Eleven undergraduate students (10 females; age range: 18–24 years old; mean age: 20.64 years old) participated voluntarily and in exchange of course credits in Experiment 2.

Apparatus and stimuli

Experiment 2 comprised the same stimuli as Experiment 1, except for following differences. The fixation point was presented for 500 ms. In addition to the regular and irregular rhythms, a control condition was included in which no sequence of tones was presented. The duration prior to the foreperiod was identical to those of the regular and irregular rhythms (i.e., 2750 ms). The presentation of each condition (regular, irregular, and no rhythm) was equally likely and randomized across trials. The foreperiod presented a constant duration across the trials, 1100 ms long. Given that in Experiment 1 we showed evidence of the selective and enhanced response in the interval that matched the regular sequence (1100 ms), in Experiment 2 we decided to use only one foreperiod for the sake of simplicity.

The stimuli of the memory task consisted of a set of six letters generated at random among the consonants of the alphabet. This set could contain either the same (e.g., "*ssssss*" – Low load condition) or different letters (e.g., "*nspdmc*" – High load condition). Both Task load conditions were presented at random across trials and with the same probability of occurrence.

Procedure and task

As in Experiment 1, participants had to respond to the target tone pressing the space and to ignore the rhythms. After their response to the target tone, a letter was presented on the screen. They were instructed to press the "a" key if that letter was included in the set presented at the beginning of the trial or, on the contrary, press the "z" key, if the letter was not present in the previous set.

The task consisted of one practice block and seven experimental blocks composed of 24 trials each. At the beginning of each trial, the fixation point was presented for 500 ms. Then, the set of six digits appeared for 3000 ms preceding the presentation of the sequence of six tones that could be either the regular, irregular, or no rhythm condition. Next, the target tone was presented after the foreperiod of 1100 ms. When participants responded to the target, the letter for the memory task was displayed on the screen. The inter-trial interval was set to 1100 ms.

Results

An ANOVA was conducted on participants' mean RT with the independent variables of Task load (High and Low) and Rhythm (regular, irregular, no rhythm) as within participants factors. Practice trials, premature responses, and trials with RT below 150 and above 1200 (2.66% of trials) were discarded from analyses. Participants' mean accuracy to the memory test was 89% (6% SD). As in Experiment 1, the analyses only included correct responses in the memory test.

The ANOVA showed a statistically significant main effect of Load, $F(1, 10) = 11.27$, $p < 0.001$, showing slower RTs in the High load condition (389 ms) than the Low load condition (365 ms). The main effect of Rhythm was also significant, $F(2, 20) = 27.21$, $p < 0.001$, indicating faster RTs after the regular rhythm (338 ms), relative to the irregular rhythm (378 ms) and no rhythm conditions (415 ms; see **Table 2**). Planned comparisons showed that the difference between regular and irregular rhythms was significant, $F(1, 10) = 17.45$, $p < 0.01$. Both the regular vs. no rhythm and irregular vs. no rhythm differences reached the statistical significance, $F(1, 10) = 82.94$, $p < 0.001$ and $F(1, 10) = 8.29$, $p < 0.02$, respectively.

Most important, the interaction between Load and Rhythm was not statistically significant, $F < 1$ (see **Figure 4**). In any case, we further analyzed this interaction using planned comparisons. These analyses showed that the regular vs. irregular difference was significant in both Load conditions, $F(1, 10) = 10.77$, $p < 0.01$, $\eta_p^2 = 0.52$ in the high load condition and $F(1, 10) = 6.33$, $p < 0.03$, $\eta_p^2 = 0.39$ in low load condition. The difference between regular and no rhythm was significant in the high load condition, $F(1, 10) = 28.23$, $p < 0.001$, and in the low load condition, $F(1, 10) = 37.66$, $p < 0.001$. Finally, the irregular and no rhythm difference was significant in the low load condition, $F(1, 10) = 12.96$, $p < 0.01$ but in the high load condition it did not reach the statistical significance, $F(1, 10) = 1.81$, $p = 0.21$.

DISCUSSION

The finding of a main effect of Rhythm confirmed that participants could temporally prepare attention by means of regular rhythms. Such rhythm effect was found of a similar magnitude in both the low and high memory load condition confirming that temporal preparation was preserved in the dual-task condition.

Table 2 | Mean RTs for each Load (high, low) and Rhythm (regular, no, irregular).

	Regular rhythm	No rhythm	Irregular rhythm
High load condition	346 (29)	422 (28)	397 (22)
Low load condition	329 (21)	407 (26)	397 (24)

Values in parentheses are standard error of the mean.

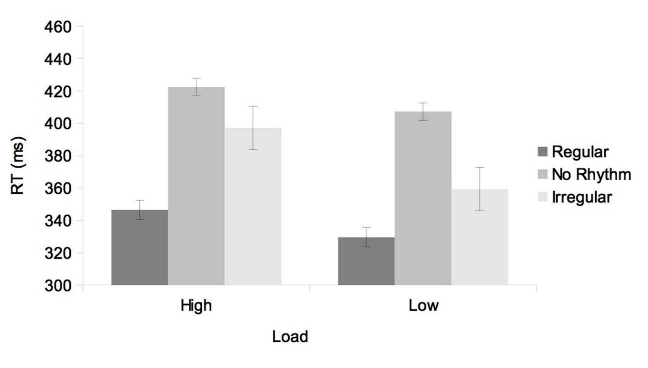

FIGURE 4 | Mean RTs as a function of Load (high, low) and Rhythm (regular, no rhythm, and irregular) in Experiment 2. Error bars represent the standard error of the mean.

Importantly, contrary to Experiment 1, the memory load in Experiment 2 was manipulated in a trial-by-trial manner. The inclusion of a no rhythm condition revealed that even the presence of an irregular rhythm resulted in a benefit in terms of RT performance. It would appear then that the mere presence of the auditory sequence, either regular or irregular, served as a temporal cue for the upcoming target, compared to a condition in which no stimulation was presented prior to the target onset. Crucially, RTs were significantly faster in the regular than in the irregular rhythm condition in both memory load conditions.

GENERAL DISCUSSION

It is currently unclear whether temporal preparation guided by rhythms involves exogenous bottom-up (e.g., Rohenkohl et al., 2011; Sanabria et al., 2011; Triviño et al., 2011), endogenous top-down mechanisms (e.g., Schwartze et al., 2011) or both. The aim of both experiments was to investigate this question by using the dual-task methodology. Assuming that temporal preparation induced by rhythms did not require controlled processing for the building up of endogenous temporal expectancies, the rhythm effect in a simple RT task would not be affected by interference from the WM task. The results in Experiments 1 and 2 showed that participants could prepare in time by means of regular rhythms, resulting in faster RTs in comparison to the irregular rhythm condition (and no rhythm condition in Experiment 2). More relevant, the rhythm effect was present in the high load and low load conditions of both experiments.

An important question to take into account in both experiments concerns to whether the two concurrent tasks involved similar or different sensory modalities. According to the Multiple Resources model (Wickens, 2008), the maximum interference occurs when the two tasks involve stimulus processing within the same sensory modality, as was the case of Capizzi et al. (2012), where both temporal preparation and WM tasks implied visual processing only. Instead, it could be argued that temporal preparation in Experiment 1 of the present study was achieved because the rhythms and the to-be-remember color stimuli did not share the same modality. One could even argue that this was the case in Experiment 2, since the letters in the memory task were presented visually. However, previous research has shown that visual stimuli are kept in short-term memory into a phonological store and that this information is refreshed by subvocal articulation through a process of rehearsal (see Baddeley, 1992, for a discussion).

A recent fRMI study (Habeck et al., 2012) has investigated, by means of the Delayed-Item-Recognition task, the neural substrates involved in non-verbal and verbal visual stimuli. In order to identify the neural regions involved in the non-verbal visual WM, these authors carried out a task consistent of the presentation of a list with one, two, or three abstract line drawings during 3 s and then, a memory test was presented in which participants had to indicate whether the probe stimulus was previously presented. Similarly, another task was performed to identify the neural regions underlying the verbal visual WM, in which letters were used instead of lines drawings. The results showed that in both tasks, verbal and non-verbal, similar frontoparietal brain regions including Broca's area (i.e., the left inferior frontal gyrus) were active. Habeck et al. (2012) suggested that this area would be involved in articulatory rehearsal of verbalizable information regardless of sensorial modality of the to-be-remembered stimuli. Crottaz-Herbette et al. (2004) also reported similar frontal activation for auditory and visual verbal (non-spatial) working memory tasks, suggesting a common neural substrate for working memory rehearsal irrespective of the modality of presentation of the WM stimuli.

In light of Habeck et al.'s (2012) and Crottaz-Herbette et al.'s (2004) results we could have expected a similar outcome in the present study whatever the modality of the stimuli in the WM. Interestingly, Crottaz-Herbette et al. (2004) also reported deactivation (with respect to a control non-WM condition) of the superior and middle temporal auditory cortex during the visual WM task and deactivation of the occipital cortex during the auditory WM task (cf. Laurienti et al., 2002). Therefore, one would have expected a reduced effect of the rhythm cue in the double task condition in Experiment 1 and in the high load condition in Experiment 2 with respect to the single-task and low load conditions. In contrast, no effects of the concurrent WM over temporal preparation driven by rhythms was found in either experiment. In sum, it would appear then that our main result could not be accounted for solely by a difference in the sensory modality of the stimuli in the WM with respect to the simple RT task.

Our results suggest that the simple auditory RT task and the WM task, both in Experiments 1 and 2, did not compete for the same processing resources. This confirms our main hypothesis that performance in the concurrent WM task would interfere performance in the simple auditory RT task if temporal preparation driven by rhythms would rely on executive processing, which it did not seem to be the case. Note, though, that a concurrent auditory perceptual task could have reduced our main auditory RT effect (cf. Santangelo et al., 2008). However, this would not contradict our main conclusion that is based on top-down executive (WM) processing effects on temporal preparation driven by rhythms.

We therefore argue that temporal preparation driven by rhythms in our study did not entail the building up of endogenous top-down temporal expectancies, in sharp contrast to temporal preparation driven by symbolic cues (cf. Capizzi et al., 2012). Our results adds to the extant literature showing that temporal preparation guided by rhythms is produced in a bottom-up, involuntary way (Rohenkohl et al., 2011; Sanabria et al., 2011; Experiment 3; Triviño et al., 2011) and that is not prone to interference by endogenous controlled processes involved in WM tasks.

The present results can be interpreted according to the dynamic attention model of Jones and colleagues (Barnes and Jones, 2000; Jones et al., 2002), where attention can be exogenously captured by rhythms and directed to appropriate moments in time. Specifically, this model assumes that the temporal pattern of rhythms produces automatically an attentional synchrony, which would enhance responses to stimuli presented at the optimum point in time.

In contrast, Schwartze et al. (2011) showed that stimulus-driven synchronization of attention would rely on top-down attention mechanisms. It is interesting to note that the pre-attentive condition in the Schwartze et al.'s (2011) study might be analogous to our dual-task condition. In effect, in both experiments participants had to concentrate in a secondary task while simultaneously listening to a rhythm (regular or irregular) that should be ignored, although in our memory dual-task condition participants were asked to respond to the target onset. In the pre-attentive condition of the Schwartze et al.'s (2011) study, the regular rhythm did not influence automatic auditory processing as revealed by the MMN potential. In contrast, the regular rhythm only influenced potentials related to attentional processing (P3b potential), and this modulation was selective to the attentive condition. These findings suggest that rhythmic, stimulus-driven, synchronization was produced by mechanisms dependent on top-down attentional mechanisms. Interestingly, behavioral performance in the attentive condition in Schwartze et al.'s (2011) study was not significantly different when comparing the regular and irregular rhythm conditions. However, in our research, we observed a rhythm RT effect (i.e., faster for the regular rhythm than for the irregular rhythm) under dual-task conditions, showing that participants could prepare in time by means of regular rhythms even though they were instructed to ignore the sequence of sounds and to attend to the WM task. Future ERP research would be interesting to clarify these apparently contradictory results, investigating how regular rhythms modulate neural processing in a dual-task condition similar to that of Experiment 2.

In sum, our study showed that temporal preparation driven by rhythms resisted the WM task interference, since participants could prepare in time while they simultaneously performed a secondary task. Thus, our study supports that temporal preparation induced by rhythms, in contrast to temporal orienting, involves stimulus-driven attentional processing in the sense that it does not require resources of executive control.

HIGHLIGHTS

(1) We investigated whether temporal preparation induced by rhythms relies on automatic mechanisms by using dual-task methodology. (2) Regular rhythms improved RTs to targets appearing at the moment in time matching the rhythmic pace. (3) This behavioral improvement resisted interference when performing concurrently a working memory task. (4) It is concluded that temporal preparation guided by rhythms involves automatic mechanisms.

ACKNOWLEDGMENTS

This research was supported by Spanish grants from the Ramón y Cajal and Plan Nacional I+D+i programmes, (RYC-2007-00296, PSI2010-15399; Ministerio de Innovación y Ciencia) to Angel Correa, grants SEJ2007-63645 from the Junta de Andalucía and PSI2010-19655 from the Plan Nacional I+D+i (Ministerio de Innovación y Ciencia) to Daniel Sanabria, and the CSD2008-00048 CONSOLIDER INGENIO (Dirección General de Investigación) to Angel Correa and Daniel Sanabria. We are grateful to Diana Cutanda for helping with data collection in Experiment 2, Dr. Sonja Kotz and two reviewers for their valuable comments to improve this manuscript.

REFERENCES

Baddeley, A. (1992). Working memory. *Science* 255, 556–559.

Barnes, R., and Jones, M. R. (2000). Expectancy, attention and time. *Cogn. Psychol.* 41, 254–311.

Capizzi, M., Sanabria, D., and Correa, A. (2012). Dissociating controlled from automatic processing in temporal preparation. *Cognition* 123, 293–302.

Correa, A., Lupiáñez, J., Milliken, B., and Tudela, P. (2004). Endogenous temporal orienting of attention in detection and discrimination tasks. *Percept. Psychophys.* 66, 264–278.

Coull, J. T., and Nobre, A. C. (1998). Where and when to pay attention: the neural systems for directing attention to spatial locations and to time intervals as revealed by both PET and fMRI. *J. Neurosci.* 18, 7426–7435.

Crottaz-Herbette, S., Anagnoson, R. T., and Menon, V. (2004). Modality effects in verbal working memory: differential prefrontal and parietal responses to auditory and visual stimuli. *Neuroimage* 21, 340–351.

Griffin, I. C., Miniussi, C., and Nobre, A. C. (2001). Orienting attention in time. *Front. Biosci.* 6, 660–671.

Habeck, C., Rakitin, B., Steffener, J., and Stern, Y. (2012). Contrasting visual working memory for verbal and non-verbal material with multivariate analysis of fMRI. *Brain Res.* 1467, 27–41.

Jones, M., Moynihan, H., MacKenzie, N., and Puente, J. (2002). Temporal aspects for stimulus-driven attending in dynamic arrays. *Psychol. Sci.* 13, 313–319.

Lange, K. (2010). Can a regular context induce temporal orienting to a target sound? *Int. J. Psychophysiol.* 78, 231–238.

Laurienti, P. J., Burdette, J. H., Wallace, M. T., Yen, Y.-F., Field, A. S., and Stein, B. E. (2002). Deactivation of sensory-specific cortex by cross-modal stimuli. *J. Cogn. Neurosci.* 14, 420–429.

Logan, G. D. (1978). Attention in character-classification tasks: evidence for the automaticity of component stages. *J. Exp. Psychol. Gen.* 107, 32–63.

Los, S. A. (1996). On the origin of mixing costs: exploring information processing in pure, and mixed blocks of trials. *Acta Psychol. (Amst.)* 94, 145–188.

Los, S. A., and Van den Heuvel, C. E. (2001). Intentional and unintentional contributions to nonspecific preparation during reaction time foreperiods. *J. Exp. Psychol. Hum. Percept. Perform.* 27, 370–386.

Posner, M. I., and Snyder, C. R. R. (1975). "Attention and cognitive control," in *Information Processing and Cognition: The Loyola Symposium*, ed. R. Solso (Hillsdale, NJ: Lawrence Erlbaum), 55–85.

Posner, M. I., Snyder, C. R. R., and Davidson, B. J. (1980). Attention and the detection of signals. *J. Exp. Psychol. Gen.* 109, 160–174.

Rohenkohl, G., Coull, J. T., and Nobre, A. C. (2011). Behavioural dissociation between exogenous and endogenous temporal orienting of attention. *PLoS ONE* 6, e14620. doi:10.1371/journal.pone.0014620

Sanabria, D., Capizzi, M., and Correa, A. (2011). Rhythms that speed you up. *J. Exp. Psychol. Hum. Percept. Perform.* 37, 236–244.

Santangelo, V., Finoia, P., Raffone, A., Olivetti Belardinelli, M., and Spence, C. (2008). Perceptual load affects exogenous spatial orienting while working memory load does not. *Exp. Brain Res.* 184, 371–382.

Schneider, W., Eschman, A., and Zuccolotto, A. (2002). *E-Prime User's Guide*. Pittsburgh: Psychology Software Tools Inc.

Schwartze, M., Rothermich, K., Schmidt-Kassow, M., and Kotz, S. A. (2011). Temporal regularity effects on pre-attentive and attentive processing of deviance. *Biol. Psychol.* 87, 146–151.

Sternberg, S. (1966). High-speed scanning in human memory. *Science* 153, 652–654.

Triviño, M., Arnedo, M., Lupiáñez, J., Chirivella, J., and Correa, A. (2011). Rhythms can overcome temporal orienting deficit alter right frontal damage. *Neuropsychologia* 49, 3917–3930.

Vallesi, A., and Shallice, T. (2007). Developmental dissociations of preparation overtime: deconstructing the variable foreperiod phenomena. *J. Exp. Psychol. Hum. Percept. Perform.* 33, 1377–1388.

Vallesi, A., Shallice, T., and Walsh, V. (2007). Role of the prefrontal cortex in the foreperiod effect: TMS evidence for dual mechanisms in temporal preparation. *Cereb. Cortex* 17, 466–474.

Wickens, C. D. (2008). Multiples resources and mental workload. *Hum. Factors* 50, 449–455.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 30 March 2012; accepted: 06 August 2012; published online: 28 August 2012.

Citation: de la Rosa MD, Sanabria D, Capizzi M and Correa A (2012) Temporal preparation driven by rhythms is resistant to working memory interference. Front. Psychology ***3****:308. doi:10.3389/fpsyg.2012.00308*

This article was submitted to Frontiers in Cognition, a specialty of Frontiers in Psychology.