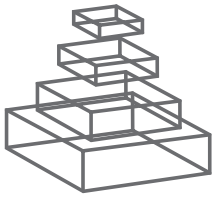
frontiers
RESEARCH TOPICS

DYNAMICS OF DECISION MAKING: FROM EVIDENCE TO PREFERENCE AND BELIEF

Topic Editors
Marius Usher, Konstantinos Tsetsos,
Erica Yu and David A. Lagnado

frontiers in PSYCHOLOGY

frontiers in NEUROSCIENCE



frontiers

Cover image provided by Ibbl sarl, Lausanne CH

ISSN 1664-8714
ISBN 978-2-88919-270-0
DOI 10.3389/978-2-88919-270-0

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

DYNAMICS OF DECISION MAKING: FROM EVIDENCE TO PREFERENCE AND BELIEF

Topic Editors:
Marius Usher, Tel-Aviv University, Israel
Konstantinos Tsetsos, Oxford University, United Kingdom
Erica Yu, University of Maryland, USA
David A. Lagnado, University College London, United Kingdom

At the core of the many debates throughout cognitive science concerning how decisions are made are the processes governing the time course of preference formation and decision. From perceptual choices, such as whether the signal on a radar screen indicates an enemy missile or a spot on a CT scan indicates a tumor, to cognitive value-based decisions, such as selecting an agreeable flatmate or deciding the guilt of a defendant, significant and everyday decisions are dynamic over time. Phenomena such as decoy effects, preference reversals and order effects are still puzzling researchers. For example, in a legal context, jurors receive discrete pieces of evidence in sequence, and must integrate these pieces together to reach a singular verdict. From a standard Bayesian viewpoint the order in which people receive the evidence should not influence their final decision, and yet order effects seem a robust empirical phenomena in many decision contexts. Current research on how decisions unfold, especially in a dynamic environment, is advancing our theoretical understanding of decision making.

This Research Topic aims to review and further explore the time course of a decision - from how prior beliefs are formed to how those beliefs are used and updated over time, towards the formation of preferences and choices and post-decision processes and effects. Research literatures encompassing varied approaches to the time-scale of decisions will be brought into scope:

a) Speeded decisions (and post-decision processes) that require the accumulation of noisy and possibly non-stationary perceptual evidence (e.g., randomly moving dots stimuli), within a few seconds, with or without temporal uncertainty.
b) Temporally-extended, value-based decisions that integrate feedback values (e.g., gambling machines) and internally-generated decision criteria (e.g., when one switches attention, selectively, between the various aspects of several choice alternatives).
c) Temporally extended, belief-based decisions that build on the integration of evidence, which interacts with the decision maker's belief system, towards the updating of the beliefs and the formation of judgments and preferences (as in the legal context).

Research that emphasizes theoretical concerns (including optimality analysis) and mechanisms underlying the decision process, both neural and cognitive, is presented, as well as research that combines experimental and computational levels of analysis.

Table of Contents

frontiers in
PSYCHOLOGY

EDITORIAL
published: 18 October 2013
doi: 10.3389/fpsyg.2013.00758

Dynamics of decision-making: from evidence accumulation to preference and belief

***Marius Usher*[1]*, *Konstantinos Tsetsos*[2], *Erica C. Yu*[3] *and David A. Lagnado*[4]**

[1] *Department of Psychology and Sagol School of Neuroscience, Tel-Aviv University, Ramat-Aviv, Israel*
[2] *Department of Experimental Psychology, University of Oxford, Oxford, UK*
[3] *Department of Psychology, University of Maryland, Maryland, MD, USA*
[4] *Department of Cognitive, Perceptual and Brain Science, University College London, London, UK*
**Correspondence: marius@post.tau.ac.il*

Edited by:
Eddy J. Davelaar, Birkbeck College, UK

Keywords: decision making, belief, evidence accumulation, data-generating process, problem solving

Decision-making is a dynamic process that begins with the accumulation of evidence and ends with the adjustment of belief. Each step is itself subject to a number of dynamic processes, such as planning, information search and evaluation. Furthermore, choice behavior reveals a number of challenging patterns, such as order effects and contextual preference reversal. Research in this field has converged toward a standard computational framework for the process of evidence integration and belief updating, based on sequential sampling models, which under some conditions are equivalent to normative Bayesian theory (Gold and Shadlen, 2007). A variety of models have been developed within the sequential sampling framework that can account for accuracy, response-time distributional data, and the speed-accuracy tradeoff (Busemeyer and Townsend, 1993; Usher and Mcclelland, 2001; Brown and Heathcote, 2008; Ratcliff and McKoon, 2008). Yet there are differences between these models with regard to the mechanism of decision-termination, the optimality of the decision and the temporal weighting of the evidence. There is also a need to extend this framework to preference type of decisions (where the criteria are up to the judge) and to enrich it so as to include control processes (such as exploration/exploitation), information search, and adaptation to the environment, thereby allowing it to capture richer decision problems; for example, when alternatives are not pre-defined, or when the decision-maker is not just accumulating evidence but also adapting beliefs about the data-generating process.

This Research Topic presents new work that investigates the dynamical and mathematical properties of evidence integration and its neural mechanisms and extends this framework to more complex decisions, such as those that occur during risky choice, preference formation, and belief updating. We hope these articles will encourage researchers to explore the computational and normative aspects of the decision process and the observed deviations. We briefly review here the contributions in this collection, starting from simple perceptual decisions in which the information flow is externally controlled to more complex decisions, which allow the observer to control the information flow and other learning strategies, and following on with preference formation.

FAST PERCEPTUAL DECISIONS

The first group of seven articles examines issues that arise in fast perceptual decisions that only allow the subject to control the weighting of the incoming evidence and the termination rule. Nevertheless, the integration time-scale, the temporal weights, and evidence termination can vary and this strongly affects the decision performance (how close people are to optimality) and the fit with the data. Some of these papers also examine the neural mechanisms that implement the decisions. In a mathematically-oriented paper Heathcote and Love (2012) examine a variant of a race model (the linear ballistic accumulator; LBA), which, under certain assumptions about the underlying distributions of starting point and drift-rate variability of evidence accumulation, allows for closed analytical formulas for the full response-time distribution in a lexical decision task and obtains a goodness of fit almost as good as that of the standard LBA model. In another formal paper, van Ravenzwaaij et al. (2012) examine, within the standard drift-diffusion model, the optimality of evidence accumulation strategies in decision situations with unequal frequency of stimuli types. They converge on the result that a bias in the decision starting point is optimal, in both fixed and variable difficulty conditions, though it appears that observers do not fully follow this strategy (but see Moran and Usher, in preparation). In another paper that examines sequential effects and decision biases in binary choice tasks, Goldfarb et al. (2012) present a simple extension of the standard decision model, which assumes changes in starting point depend on stimulus repetitions and alternations, combined with a response criteria increase following errors. This model accounts for a rich data of sequential dependence in response time and accuracy. In a paper contributed by Tsetsos et al. (2012), the aim was to contrast the standard drift-diffusion algorithm, which assumes that the evidence is given temporally-uniform decision weights, and the leaky competing accumulator model (LCA), which predicts a variety of temporal weighting patterns, including (for some model parameters) a specific interaction between stimulus duration and temporal weighting. While the LCA-predicted interaction was confirmed in some of the observers (who performed multiple sessions with the moving dots displays), future work will be needed to further characterize how temporal weighting of evidence depends on task characteristics and individual differences. The issue of temporal weighting and its dependence on characteristics of evidence accumulation and type of decision-boundary is further discussed in a review paper by Zhang (2012), who also examines how these characteristics affect decision optimality. Lastly, two papers discuss the neural mechanisms of perceptual decisions.

Simen (2012) examines a two-layer neural model that includes accumulators and bistable cell-assemblies that can implement the decision-boundary—which is assumed without much discussion in the standard approach—and discusses the difficulties of mapping those processing units to the neural recordings observed in brain data. van Vugt et al. (2012) use a model-driven approach to reveal the EEG correlates of evidence accumulation for a motion discrimination task. The authors use a novel computational technique to show that the time-course of the EEG activity demonstrated a non-linear profile—a finding that may arbitrate the dispute between linear (e.g., Brown and Heathcote, 2008) and non-linear (e.g., Usher and Mcclelland, 2001) models of evidence integration. Moreover, this paper indicates the possibility of identifying individual differences in evidence integration (e.g., speed-accuracy trade-off) from the EEG signal, offering a useful tool for characterizing the computational properties of the decision mechanism.

ADAPTIVE DECISION MAKING

The second group of six articles examines decisions that extend over a longer time-frame and which allow the subject to control the evidence accumulation process, and to form and update beliefs about the state of the environment. The study by Knox et al. (2011) of exploration and exploitation suggests that human decision makers learn from interaction with their environment in a reflective manner (without requiring direct observation of changes in the environment) but yet do not plan optimally because they do not consider the long-term information value of actions. The contribution by Osman and Speekenbrink (2012) extends this inquiry by studying how knowledge about the values of actions can be affected by tasks of prediction (outcome estimation) and control (interventions to achieve an outcome). They demonstrate a distinction between prediction and control whereby controllers were able to transfer their knowledge to tests of prediction but not vice versa. In this way, the concept of control is similar to that of planning for a goal rather than for adapting to an environment as in Knox et al. (2011) but, in both of these papers, decision makers cycle from evidence accumulation, to action, to feedback, and back again (cf. model-based learning; Sutton and Barto, 1998). Yu and Lagnado (2012) use this framework in a slot machine paradigm to show that, while participants over time came to understand the observed environment (slot machine payouts) accurately, their understanding of the underlying structure of the environment was flawed. Beliefs about structure and causality were more strongly influenced by initial beliefs than by experience. Also studying decisions from experience, Dutt and Gonzalez (2012) explore the role of inertia, or the tendency to repeat one's final decision, irrespective of its outcome. They show both the advantages and disadvantages of incorporating inertia into an instance-based learning model of repeated binary choice. In contrast to this focus on inertia, Lange et al. (2012) demonstrate how decision makers adapt to the environment, using a new model that combines the HyGene model (Thomas et al., 2008) with the context-activation model (Davelaar et al., 2005). Across two experiments that manipulate serial order, consistency of newly-acquired evidence with previously-generated hypotheses, and elicitation timing, the authors show that not all data have an equal impact on hypothesis generation processes: newly-acquired data can cause inconsistent hypotheses to be purged from working memory. The authors propose that whether this results in a recency or primacy effect is likely to depend on the richness of the information and its rate of presentation.

PREFERENCE-BASED DECISIONS

The next group of three articles addresses preference formation, in situations (risky choice and multi-attribute decisions) that do not set up an objective/normative criterion, but rather leaves this to the subject's control. Fiedler and Glockner (2012) monitor how people choose between lotteries using eye-tracking to distinguish between competing models of risky choice. The results disconfirmed Take-the-Best, or lexicographic heuristics, in favor of compensatory models that assume observers integrate outcomes with attentional weights determined by outcome probability. In particular, people gather more information within (rather than between) lotteries and they tend to gather more information (toward the end of the decision) from the chosen alternative, indicating top–down feedback from alternative to processing representations. Also using eye-tracking, Krajbich et al. (2012) propose a formalization of the influence of visual fixations on the dynamics of preference formation. The authors build on the attentional diffusion model (aDDM), which modulates the rate of evidence-accumulation depending on the position of visual fixation, to explain the responses and reaction times of human subjects during purchasing (accept/reject) decisions. The study demonstrates how small attentional fluctuations during the deliberation period can influence the decision outcome. This approach is closely related to theoretical models of multi-attribute choice [e.g., decision field theory, Roe et al. (2001); and value-based LCA, Usher and Mcclelland (2004)], in which attentional switching to different choice aspects drives preference formation. This class of models is extended in Wollschlager and Diederich (2012), which presents a novel model of contextual preference reversal (attraction, similarity, and compromise effects) for multi-alternative, multi-attribute choice: the 2N-ary Choice Tree model. The model offers closed-form expressions for choice probabilities and response time distributions and, contrary to previous theories, explains reversal effects by assuming that attentional weights depend on the alternatives in the choice-set [cf. a recent study, which appeared after this Research Topic and provides an explicit mechanism for how the alternatives affect weights to the choice attributes: Bhatia (2013)].

NOVEL OR INTEGRATIVE APPROACHES

Finally, two papers aim to provide novel or integrative frameworks for understanding dynamical decision making. Trueblood and Busemeyer (2012) present a decision model based on principles of quantum theory—a radical shift from the standard framework—which provides a novel account of order effects in belief updating and inference. This paper provides an introduction to the elegant principles of quantum probability. This theory is of great potential, although stronger data might be needed to persuade the skeptical readers (e.g., showing cyclic changes in order effects). In the last paper Fox et al. (2013) present an overarching framework for the entire decision making cycle, from

the framing of a decision to establishing preferences and making commitments. They extend the standard model to more ecological and dynamic situations, in which the alternatives are not predefined and the agent faces a variety of constraints and conflicts. The theory situates dynamical decision making with respect to other high-level cognitive capabilities such as problem solving, planning and collaborative decision-making.

CONCLUSIONS AND FUTURE WORK

We believe that this collection has revealed a number of important aspects of the nature of decision processes. More importantly, we hope that it will stimulate readers to keep probing these processes. Various key questions are still unresolved. How close are people to optimality when making decisions? Why does this vary so much between the cases of evidence and preference? Is the Bayesian framework a general one for all types of decisions (can one extend it to more complex cases that allow the subject control over the information flow and the decision criteria?). What are the neural mechanisms, and the nature of individual differences? Future research into these topics, should surely keep us stimulated for the near future.

ACKNOWLEDGMENTS

We want to thank Eddy J. Davelaar for very helpful Editorial suggestions on this paper. Marius Usher is supported by the Israeli Science Foundation (Grant 743/12).

REFERENCES

Bhatia, S. (2013). Associations and the accumulation of preference. *Psychol. Rev.* 120, 522–543. doi: 10.1037/a0032457

Brown, S. D., and Heathcote, A. (2008). The simplest complete model of choice response time: linear ballistic accumulation. *Cogn. Psychol.* 57, 153–178. doi: 10.1016/j.cogpsych.2007.12.002

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. doi: 10.1037/0033-295X.100.3.432

Davelaar, E. J., Goshen-Gottstein, Y., Ashkenazi, A., Haarmann, H. J., and Usher, M. (2005). The demise of short-term memory revisited: empirical and computational investigations of recency effects. *Psychol. Rev.* 112, 3–42. doi: 10.1037/0033-295X.112.1.3

Dutt, V., and Gonzalez, C. (2012). The role of inertia in modeling decisions from experience with instance-based learning. *Front. Psychol.* 3:177. doi: 10.3389/fpsyg.2012.00177

Fiedler, S., and Glockner, A. (2012). The dynamics of decision making in risky choice: an eye-tracking analysis. *Front. Psychol.* 3:335. doi: 10.3389/fpsyg.2012.00335

Fox, J., Cooper, R. P., and Glasspool, D. W. (2013). A canonical theory of dynamic decision-making. *Front. Psychol.* 4:150. doi: 10.3389/fpsyg.2013.00150

Gold, J. I., and Shadlen, M. N. (2007). The neural basis of decision making. *Annu. Rev. Neurosci.* 30, 535–574. doi: 10.1146/annurev.neuro.29.051605.113038

Goldfarb, S., Wong-Lin, K., Schwemmer, M., Leonard, N. E., and Holmes, P. (2012). Can post-error dynamics explain sequential reaction time patterns? *Front. Psychol.* 3:213. doi: 10.3389/fpsyg.2012.00213

Heathcote, A., and Love, J. (2012). Linear deterministic accumulator models of simple choice. *Front. Psychol.* 3:292. doi: 10.3389/fpsyg.2012.00292

Knox, W. B., Otto, A. R., Stone, P., and Love, B. C. (2011). The nature of belief-directed exploratory choice in human decision-making. *Front. Psychol.* 2:398. doi: 10.3389/fpsyg.2011.00398

Krajbich, I., Lu, D., Camerer, C., and Rangel, A. (2012). The attentional drift-diffusion model extends to simple purchasing decisions. *Front. Psychol.* 3:193. doi: 10.3389/fpsyg.2012.00193

Lange, N. D., Thomas, R. P., and Davelaar, E. J. (2012). Temporal dynamics of hypothesis generation: the influences of data serial order, data consistency, and elicitation timing. *Front. Psychol.* 3:215. doi: 10.3389/fpsyg.2012.00215

Osman, M., and Speekenbrink, M. (2012). Prediction and control in a dynamic environment. *Front. Psychol.* 3:68. doi: 10.3389/fpsyg.2012.00068

Ratcliff, R., and McKoon, G. (2008). The diffusion decision model: theory and data for two-choice decision tasks. *Neural Comput.* 20, 873–922. doi: 10.1162/neco.2008.12-06-420

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. doi: 10.1037/0033-295X.108.2.370

Simen, P. (2012). Evidence accumulator or decision threshold - which cortical mechanism are we observing? *Front. Psychol.* 3:183. doi: 10.3389/fpsyg.2012.00183

Sutton, R. S., and Barto, A. G. (1998). *Introduction to Reinforcement Learning*. Cambridge: MIT Press.

Thomas, R. P., Dougherty, M. R., Sprenger, A. M., and Harbison, J. I. (2008). Diagnostic hypothesis generation and human judgment. *Psychol. Rev.* 115, 155–185. doi: 10.1037/0033-295X.115.1.155

Trueblood, J. S., and Busemeyer, J. R. (2012). A quantum probability model of causal reasoning. *Front. Psychol.* 3:138. doi: 10.3389/fpsyg.2012.00138

Tsetsos, K., Gao, J., Mcclelland, J. L., and Usher, M. (2012). Using time-varying evidence to test models of decision dynamics: bounded diffusion vs. the leaky competing accumulator model. *Front. Neurosci.* 6:79. doi: 10.3389/fnins.2012.00079

Usher, M., and Mcclelland, J. L. (2001). The time course of perceptual choice: the leaky, competing accumulator model. *Psychol. Rev.* 108, 550–592. doi: 10.1037/0033-295X.108.3.550

Usher, M., and Mcclelland, J. L. (2004). Loss aversion and inhibition in dynamical models of multialternative choice. *Psychol. Rev.* 111, 757–769. doi: 10.1037/0033-295X.111.3.757

van Ravenzwaaij, D., Mulder, M. J., Tuerlinckx, F., and Wagenmakers, E. J. (2012). Do the dynamics of prior information depend on task context? An analysis of optimal performance and an empirical test. *Front. Psychol.* 3:132. doi: 10.3389/fpsyg.2012.00132

van Vugt, M. K., Simen, P., Nystrom, L. E., Holmes, P., and Cohen, J. D. (2012). EEG oscillations reveal neural correlates of evidence accumulation. *Front. Neurosci.* 6:106. doi: 10.3389/fnins.2012.00106

Wollschlager, L. M., and Diederich, A. (2012). The 2N-ary choice tree model for N-alternative preferential choice. *Front. Psychol.* 3:189. doi: 10.3389/fpsyg.2012.00189

Yu, E. C., and Lagnado, D. A. (2012). The influence of initial beliefs on judgments of probability. *Front. Psychol.* 3:381. doi: 10.3389/fpsyg.2012.00381

Zhang, J. (2012). The effects of evidence bounds on decision-making: theoretical and empirical developments. *Front. Psychol.* 3:263. doi: 10.3389/fpsyg.2012.00263

Received: 15 August 2013; accepted: 27 September 2013; published online: 18 October 2013.

*Citation: Usher M, Tsetsos K, Yu EC and Lagnado DA (2013) Dynamics of decision-making: from evidence accumulation to preference and belief. Front. Psychol. **4**:758. doi: 10.3389/fpsyg.2013.00758*

This article was submitted to Cognitive Science, a section of the journal Frontiers in Psychology.

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 23 August 2012
doi: 10.3389/fpsyg.2012.00292

Linear deterministic accumulator models of simple choice

Andrew Heathcote* and Jonathon Love

School of Psychology, The University of Newcastle, Callaghan, NSW, Australia

Edited by:
Marius Usher, Tel Aviv University, Israel

Reviewed by:
Patrick Simen, Oberlin College, USA
KongFatt Wong-Lin, University of Ulster, Northern Ireland
Rani Moran, Tel Aviv University, Israel

****Correspondence:***
Andrew Heathcote, Psychology Building, The University of Newcastle, University Avenue, Callaghan, NSW 2308, Australia.
e-mail: andrew.heathcote@newcastle.edu.au

We examine theories of simple choice as a race among evidence accumulation processes. We focus on the class of deterministic race models, which assume that the effects of fluctuations in the parameters of the accumulation processes between-choice trials (*between-choice noise*) dominate the effects of fluctuations occurring while making a choice (*within-choice noise*) in behavioral data (i.e., response times and choices). The latter deterministic approximation, when combined with the assumption that accumulation is linear, leads to a class of models that can be readily applied to simple-choice behavior because they are computationally tractable. We develop a new and mathematically simple exemplar within the class of linear deterministic models, the Lognormal race (LNR). We then examine how the LNR, and another widely applied linear deterministic model, Brown and Heathcote's (2008) LBA, account for a range of benchmark simple-choice effects in lexical-decision task data reported by Wagenmakers et al. (2008). Our results indicate that the LNR provides an accurate description of this data. Although the LBA model provides a slightly better account, both models support similar psychological conclusions.

Keywords: evidence accumulation, mathematical modeling, response time, linear ballistic accumulator, lexical-decision task

INTRODUCTION

Humans and other organisms often have to respond to stimuli under time pressure that requires them to make choices in a few seconds or less. In contrast to complex choices requiring an extended period of deliberation during which a long series of cognitive operations are completed, rapid choices are usually assumed to have a simple cognitive architecture consisting of three stages: stimulus encoding, response selection, and response execution. Response selection in simple choice is almost universally modeled by evidence accumulation, that is, by a process that accumulates evidence until the amount favoring one of the choices is sufficient to exceed an evidence boundary. Evidence accumulation has the disadvantage that it becomes increasingly time consuming when the evidence boundary is high. However, increasing response caution by increasing the boundary is also assumed to have utility because it ameliorates the effects of various types of noise that can cause choice errors. This assumption, and the task of providing a quantitative account of the relationship between speed and accuracy (*speed-accuracy trade-off*), has been pivotal for the development of models of simple choice.

Since the earliest proposals (e.g., Stone, 1960), it has usually been assumed that fluctuations in evidence occurring during the accumulation process (i.e., *within-choice noise*) are the dominant cause of both choice errors and of variations in response time (RT) from choice-to-choice. However, it soon became evident that within-choice noise is not by itself sufficient to enable models to provide a comprehensive account of choice behavior. In seminal work, Laming (1968) and Ratcliff (1978) demonstrated that accounting for not only the frequency with which different choices are made but also the distribution of RT for every type of choice requires the addition of effects due to choice-to-choice fluctuations (i.e., *between-choice noises*).

In a departure from the usual assumption, Brown and Heathcote (2005a) asked whether between-choice noises alone could provide a comprehensive account of simple-choice behavior. Although acknowledging that a range of extrinsic (e.g., stimulus) and intrinsic (e.g., neural) factors can cause within-choice noise, they proposed that the attendant behavioral effects might sometimes be small enough to neglect. In support of this simplifying approximation, which they described as "ballistic," they demonstrated that a model with no within-choice noise could provide a detailed account of a broad range of benchmark simple-choice behaviors. Their model, the ballistic accumulator (BA), was further simplified by Brown and Heathcote (2008) into the linear ballistic accumulator (LBA). The LBA was shown to provide an account of benchmark phenomena on par with the BA while gaining considerably in ease of application because of greater mathematical and computational tractability.

Here we extend Brown and Heathcote's (2005a, 2008) line of argument by developing an even more mathematically tractable evidence accumulation model that shares with the LBA the assumptions that accumulation is linear and deterministic. We first set the context for this development by reviewing the roles of different types of noise in evidence accumulation models and by defining a framework within which the LBA, and our new proposal, the Lognormal race (LNR), are special cases.

Next we motivate the LNR model's Lognormal distribution assumption, derive mathematical results, and show that the LNR model, because of its simplicity, is required to explain speed-accuracy trade-offs in an unconventional way, via changes in evidence accumulation. Finally, we test and compare the LBA and LNR models by fitting them to behavioral data from a lexical-decision task (i.e., classifying a letter string as either a word or a non-word) reported by Wagenmaker et al.'s (2008, Experiment 1).

We focused on Wagenmaker et al.'s (2008) experiment because it produced a very large speed-accuracy trade-off using instructions that emphasized either response speed or response accuracy. Fitting methods developed by Donkin et al. (2011a) enabled us to systematically explore a large variety of LBA and LNR model parameterizations that instantiate different ways to quantitatively explain the speed-accuracy trade-off. We show that LNR model is able to provide an accurate description of the frequency of each choice and its associated RT distribution. We also show that the LBA model is only able to provide the same accurate description if it explains a large part of the observed speed-accuracy trade-off in the same unconventional way as the LNR.

SOURCES OF NOISE IN EVIDENCE ACCUMULATION MODELS

Early evidence accumulation models – random walks and their continuous analog, a diffusion process (e.g., Stone, 1960) – assumed only within-choice noise. However, such simple models are inadequate because they predict correct and error choices have identical RT distributions, whereas empirically correct and error RT differ in regular and often replicated ways. For example, when decision accuracy is stressed errors are slower than correct responses, but when decision speed is stressed this difference decreases and can even reverse (e.g., Ratcliff and Rouder, 1998). These limitations can be remedied by the addition of two sources of between-choice noise.

First, Laming (1968) showed that variability in the starting point of a random walk process causes fast errors. As the starting point determines the amount of evidence required for each choice, there is an attendant between-choice fluctuation in response bias. Second, Ratcliff (1978) accounted for the more commonly occurring slow errors in a diffusion model by allowing the mean rate of evidence accumulation to differ between trials. Between-choice noise in the mean rate of evidence accumulation also allows these models to escape a prediction that is clearly false for many choice tasks; that perfect accuracy can be achieved by a sufficient increase in the amount of evidence required to make a choice.

What are the causes of these types of between-choice noise? In Ratcliff's (1978) application – episodic recognition memory – mean rate variation could plausibly be attributed to substantial differences in memorability between test items (words), as responses to different words were aggregated within experimental conditions. Subsequent research has shown that mean rate variation is also required to fit behavioral data from paradigms using homogenous test items within each experimental condition. This suggests there may be other causes of mean rate variations not related to item effects, such as choice-to-choice fluctuations in attention and arousal. Shadlen and Newsome (1998) provide a potential neural cause; they showed that correlations among the firing of neurons coding the same stimulus also cause choice-to-choice variations in mean spike-rates.

Sequential effects are the most commonly proposed cause of between-choice noise in the starting points of evidence accumulation (i.e., in the amount of evidence required for each choice). Simple-choice paradigms typically require participants to make a series of closely spaced decisions, so residual effects from previous decisions have been proposed as a source of start-point noise in cognitive (e.g., Brown et al., 2008) and neurophysiological (e.g., Gao et al., 2009) process models. Van Maanen et al. (2011) recently reported evidence that model-based estimates of choice-to-choice fluctuations in the amount of evidence required for a response are correlated with changes in hemodynamic responses in areas associated with response caution, the pre-supplementary motor area and anterior cingulate (see also Huettel et al., 2002).

The most widely and successfully applied evidence accumulation model, the Ratcliff diffusion model (RDM, see Ratcliff and McKoon, 2008, for a summary), owes its ability to provide a comprehensive account of decision behavior to the inclusion of between-choice noise in both start points and mean rates. More recently, a third type of between-choice noise, in the time to complete encoding and response production processes (denoted *Ter*). *Ter* noise was required by Ratcliff et al. (2004) to be able to enforce the assumption that word frequency selectively influences the rate of evidence accumulation in a lexical-decision task, as otherwise they could not account for systematic effects of word frequency on fast responses.

Race models constitute a second important class of evidence accumulation models (see Marley and Colonius, 1992, for an overview). In a race model each choice is represented by a separate accumulator, with the choice made corresponding to the first accumulator to hit its evidence boundary, and RT to the time required to do so. Like the diffusion model, early race models (e.g., Vickers, 1979) assumed a dominant role for within-choice noise and linear accumulation (i.e., equal weighting of samples taken earlier vs. later in the accumulation process). These simple assumptions were shown to be problematic because they predicted that the distribution of RT became more symmetric as overall RT slowed. Empirically, RT distributions show strong positive skew for all but the simplest and most rapid decisions (Luce, 1986).

This problem with race models was overcome by non-linear accumulation in Usher and McClelland's (2001) Leaky Competitive Accumulator (LCA) model. Their model's non-linear accumulation mechanisms were inspired by the fact that single-cell neural dynamics are commonly found to be "leaky" (i.e., firing rates return to baseline in the absence of input) and competitive (i.e., increased firing in one neuron can suppress firing in another). The interplay between these two types of non-linearity (i.e., either one or the other dominating) can result either in late-arriving evidence being more influential on the eventual choice (due to leakage) or in early arriving evidence being more influential (due to competition). Non-linear accumulation has also been proposed in a generalization of the class of diffusion models, an Ornstein–Uhlenbeck process, as part of Busemeyer and Townsend's (1993) Decision Field Theory (DFT). Ratcliff and Smith (2004) claimed that this generalization was not required based on an analysis that estimated the degree of non-linearity from fits to data. However, more recent work by Leite and Ratcliff (2010), using the same type of analysis, found leakage to be necessary in non-competitive race models dominated by within-choice noise.

Brown and Heathcote's (2005a, 2008) BA and LBA are race models that have their roots in models from both cognitive psychology and decision neuroscience. The BA is a race model identical in architecture and deterministic non-linear dynamics

to Usher and McClelland's (2001) LCA, but with only between-choice (start point and rate) noise. The LBA model removes two further components of the LCA, leakage in accumulation, and competition between accumulators. As a result accumulation is linear in the LBA and the level of evidence in one accumulator is independent of the level in other accumulators until one hits its boundary. At that time all other accumulators are inhibited, so that only one response is made. The overall architecture of the LBA is identical to Logan and Cowan's (1984) horse-race model of the stop-signal paradigm. The LBA model's assumption that inhibition plays a role after, rather than during, accumulation is largely consistent with Boucher et al.'s (2007) neurobiological findings in the stop-signal task.

The LBA's assumption that accumulation is linear and deterministic, and that the rate characterizing this linear accumulation varies among choices according to a normal distribution, are shared with Carpenter's (1981) LATER model. The LATER model has been widely applied to behavioral and neuroscience studies where responding is via eye-movements and the focus is on modeling (i.e., non-choice) RT. Ratcliff (2001) pointed out that LATER is unable to account for effects related to error responses in choice paradigms. The LBA differs from LATER in assuming that the distance from the starting point of accumulation to the boundary varies between trials according to a uniform distribution. This added assumption allows the LBA to account for error related phenomena, such as systematic differences between correct and error responses and speed-accuracy trade-offs.

As pointed out by Ratcliff (2001), within-choice noise models account for speed-accuracy trade-off because accumulation integrates out moment-to-moment fluctuations in evidence. In the LBA speed-accuracy trade-off can occur because accumulation integrates out response bias due to between-choice start-point noise. In both model classes between-choice rate noise also serves to limit the accuracy that can be achieved by increasing response caution. This provides one explanation of the observation that even very slow decisions can be inaccurate in some tasks.

It is important to clarify the meaning of the term "ballistic" as employed by Brown and Heathcote (2005a, 2008), as it is non-standard. It does not imply that, like a projectile fired from a cannon, the trajectory of evidence accumulation is entirely determined when initiated. Rather, it indicates a lack of within-choice noise. To illustrate this point, consider Brown and Heathcote's (2005b) experiment using stimuli that briefly (e.g., for 90 ms) favored one choice, then switched to favor an alternative choice. They conceived of the rate of evidence accumulation as being able to change as a function of the stimulus change during a trial. Interference paradigms, such as the flanker task (e.g., Gratton et al., 1988), provide another case where it is likely that the input to the accumulation process is non-stationary (i.e., changes over time). Naturally, sensitivity to stimulus change is limited by the low-pass filtering imposed by sensory processes, and such sensitivity will also vary depending on attention-mediated selection of task relevant vs. irrelevant features. An example is provided by the global vs. local motion classification task with random-dot kinematogram stimuli used by Ho et al. (2009); they reported a successful application of a stationary-rate LBA model to global motion choices based on their rapidly time-varying stimuli.

THE LINEAR DETERMINISTIC ACCUMULATION FRAMEWORK

In this section we articulate a general framework for linear deterministic accumulation models. Within this framework particular models differ in the assumptions they make about the distributions followed by each type of between-choice noise. We begin by outlining the general framework, using the LBA model as an illustration, and then we develop a new model that makes different distributional assumptions, the LNR.

The LBA assumes a uniform distribution of start-point noise and normal distribution for rate noise. These assumptions were made both as a matter of convention (e.g., the same assumptions are made by the RDM) and mathematical convenience. Mathematically, they enable computationally tractable solutions for the density and cumulative density functions describing the distribution of times at which the evidence total first hits the boundary of a single accumulator. These two functions can then be easily combined to determine the likelihood of any given response at any given time from a set of one or more potential responses (i.e., for a race amongst any number of accumulators). A likelihood, which is not easily computed for alternative models such as the LCA and RDM, enables efficient model estimation, and so has facilitated applications of the LBA (see Donkin et al., 2011a, for a tutorial).

Equation 1 characterizes the time, T, for a single evidence total to accumulate to a boundary without the specific commitments to distributional assumptions made by Brown and Heathcote's (2008) LBA.

$$T = \frac{D}{V} \tag{1}$$

The numerator of the ratio in (1), $D \geq 0$, indicates the distance between the starting point of evidence accumulation and the boundary, and (1) assumes that T is undefined if $V \leq 0$. For the LBA, $D \sim B + U(0, A)$, where "$\sim$" means "is distributed as," $U(0, A)$ indicates the uniform start-point distribution on the interval from 0 to $A (A \geq 0)$, and $B (B \geq 0)$ is the distance from the upper bound of the start-point distribution to the evidence boundary. The denominator, V, is the rate (velocity) of evidence accumulation. The LBA rate distribution is normal with a mean of v and standard deviation sv: $V \sim N(v, sv)$.

A LEXICAL-DECISION TASK EXAMPLE

Figure 1 illustrates an LBA model of a lexical-decision task, in which participants have to decide if a string of letters makes up a word. **Figure 1** illustrates a trial in which the stimulus is a word, and so the rate distribution for the true (i.e., word) accumulator has a higher mean than the rate distribution for the false (i.e., non-word) accumulator. In this illustration the sampled rates (indicated by the slope of the dotted line) follow the same order as the mean rates, but a choice error (i.e., a non-word response) is made because the non-word accumulator hit its boundary first. The error occurs because the non-word accumulator starts with a higher level of evidence.

One way in which a speed-accuracy trade-off can be explained is illustrated by considering what would happen if the boundary were sufficiently increased in **Figure 1**; the higher rate of the word accumulator would eventually overcome the initial response bias in

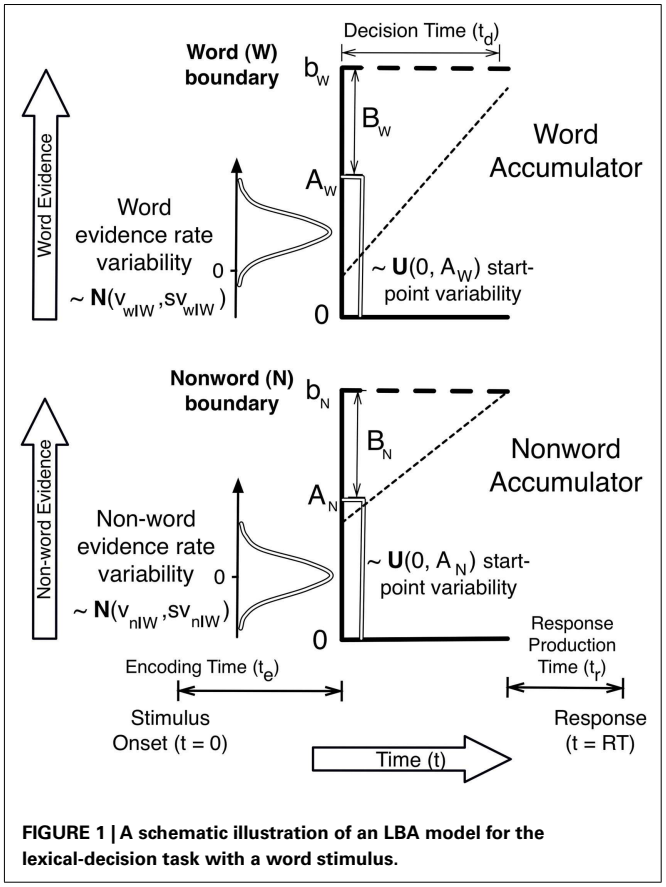


FIGURE 1 | A schematic illustration of an LBA model for the lexical-decision task with a word stimulus.

favor of the non-word accumulator and an accurate word response would be made. Past applications of the BA and LBA (Brown and Heathcote, 2005a, 2008) have assumed that this response caution based mechanism explains speed-accuracy trade-off caused by speed vs. accuracy emphasis instructions. However, it might also be explained by a change in the upper boundary of start-point noise distribution, *A*. If *A* decreased under accuracy emphasis responding would slow, as the average distance from start-point to boundary would increase, and become more accurate, as bias favoring the false accumulator would become weaker on average. The same is true of other accumulator models, such as LCA and DFT.

The overlap of the two rate distributions in **Figure 1** illustrates why responding may not always be entirely accurate even with a very high boundary. On some trials a higher rate will be sampled for the incorrect (non-word) accumulator. In this case, if a quick correct response is not caused by response bias, an error will occur no matter how high the boundary is set. Although they have not conventionally done so, evidence accumulation models might also explain the effects of speed vs. accuracy emphasis through changes in mean rates and the level of rate noise (e.g., between-choice noise in the LBA and both within-choice and between-choice noise in the RDM, LCA, and DFT models). For example, if accuracy emphasis caused *v* and *sv* parameters to decrease equally for all accumulators, RT would slow (as it would take longer to hit an evidence boundary) and errors would decrease (due to a reduction in the overlap of the rate distributions).

Given that Wagenmakers et al. (2008) manipulated speed vs. accuracy emphasis between trial blocks participants may have had time to make global changes in factors like attention and arousal that might plausibly affect accumulation rates (Kleinsorge, 2001). Hence, rather than imposing one particular way of explaining the effects of emphasis, we took a more exploratory approach in fitting in Wagenmakers et al.'s data. That is we fit all of the different possible ways of explaining the effect of emphasis by allowing appropriate variation in the LBA's *B*, *A*, *v*, and *sv* parameters.

MULTIPLE AND CONTINGENT CHOICE

The two-choice case illustrated in **Figure 1** can be generalized to choice between any numbers of alternatives, where one accumulator corresponds to each alternative. Suppose the densities of T at time t for each of $i = 1 \ldots N$ accumulators are denoted by $f_i(t)$ and the survivor functions (i.e., the complements of the cumulative densities) by $S_i(t)$. If the corresponding D_i and V_i are assumed independent between accumulators, as is the case for the LBA, the likelihood of response i at time t, where Π denotes a repeated product, is:

$$T_d\,(i,t) = f_i\,(t) \prod\nolimits_{j \neq i} S_j\,(t) \tag{2}$$

The expression for decision time (T_d) in (2) describes what is sometimes called a defective density, a curve that integrates to a value less than or equal to one, where that value corresponds to the probability that response i wins the race.

Even more general, yet still computationally tractable, models can be derived based on race equations such as (2). For example, Eidels et al. (2010) describe a model in which binary responses are made contingent on logical relations between two stimuli as computed by four linear deterministic accumulators. This illustrates the considerable power and generality afforded by independent race models. As we discuss below, the LNR model extends the power of this approach by providing a tractable approach to including correlations between the inputs to accumulators.

RESIDUAL TIME

It is necessary to make one further addition to the framework to describe observed behavior, an account of the "residual" component of RT not accounted for by decision time. This non-decision time is conventionally annotated "Ter," an acronym for time (T) for encoding (e) and response (r). As illustrated in **Figure 1**, RT is typically assumed to be the sum of decision time and a residual time that does not vary as a function of either the stimulus or response. This invariance is reasonable in typical rapid choice paradigms where the difficulty of stimulus encoding and response production is fairly homogenous, but it need not always apply. For example, Karayanidis et al. (2009) reported large differences in residual time for fits of a diffusion model in a cued-task-switching paradigm as a function of whether the cue indicated a repeat or switch in task.

A second consideration related to residual time concerns whether it is a constant or whether it is variable. Although some variation in the processes causing residual time is highly likely, a constant will provide a reasonable approximation if that variation is small relative variation associated with decision time.

For example, Smith (1995) reports evidence consistent with the standard deviation of motor production time in button-pressing tasks being of the order of 10 ms, which would mean it accounts for less than 1% of the overall variability in RT in even quite rapid choices. However, between-choice variability in residual time, in particular the sum of a constant and a uniform random deviate, has become standard in recent applications of the RDM (e.g., Ratcliff et al., 2004). We assume a constant residual time in the model tests described later, as that provides a substantial advantage in terms of computational speed for both the LBA and LNR. Although that suited the exploratory aims of these tests it does not indicate a commitment to residual time always being a constant in the general linear deterministic framework.

THE LOGNORMAL RACE MODEL

We make two observations that provide some motivation for the LNR. First, as Brown and Heathcote (2008) note, it is possible for an LBA to fail to respond where no sampled rate is positive, although they found the probability of a non-response to be negligible in fits to data. However, non-responding is not a necessary characteristic of the general class of linear deterministic accumulators. Non-responding does not occur in the LNR model because the logarithm of the rate is assumed to have a normal distribution, and hence the rate for every accumulator has a (necessarily positive) Lognormal distribution. The same would be true of any other linear deterministic model that assumes a positive rate distribution.

Second, it turns out that the independence between accumulators assumed in (2) is not necessary to derive a computationally tractable expression for the LNR likelihood. Although widely made, we argue that the assumption of independence between accumulator inputs and/or between accumulator distances may be questionable in at least some circumstances, such as when evidence for each response alternative is derived from the same stimulus characteristics. In such circumstances, it is natural to assume choice-to-choice variations in stimuli will cause the inputs to different accumulators to be correlated to some degree, even if the input to each accumulator also contains some stimulus independent sources of noise. The Lognormal distributional assumption allows us to avoid the independence assumption without greatly increasing the computational cost of estimation.

The LNR model derives from the work of Ulrich and Miller (1993), who proposed that RT dynamics could be approximated by simple version of a "continuous flow" or "partial outputs" system (Schweikert, 1989; Townsend and Fikes, 1995). They examined a system that is time invariant (autonomous), in the sense that its rate of change does not depend directly on time, and that has no memory, in the sense that its rate of change is independent of its current state (activation). The latter property distinguishes this system from perhaps the most well known partial outputs model, McClelland's (1979) cascade process. Ulrich and Miller derived the prediction of approximately Lognormal RT for a flow with stages characterized by a constant rate of change (i.e., linear accumulators). In the next section we briefly summarize their development and show how it can be used to model simple (i.e., non-choice) RT tasks (e.g., press a button when a light comes on). We then expand the development to account tasks where participants must choose between two or more responses.

SIMPLE RESPONSE TIME

Suppose a flow is made up of S linear accumulator stages, with associated rates v_s, $s = 1..S$, the activation, x, of the terminal stage as a function of time, t, is given by $x_S = tV$, where $V = \Pi_{s=1..S} v_s$. The flow can be approximated as being "lumped," in the sense that transmission occurs instantaneously from the initial to a terminal stage with any delay confined to a time, t_1, between stimulus presentation and the time at which that presentation begins to effect activation in the flow. Without loss of generality it can also be assumed that the sensory input to the first stage is $x_0 = 1$, as any differences in input magnitude can be absorbed in to the rate of the first stage. The terminal stage is a unit that represents a response, in the sense that a response is initiated when the activation of a terminal unit travels a distance D from its initial state.

The time taken to traverse that distance from the time that sensory processing commences has the same form, $T = D/V$, as (1) from the general framework. Ulrich and Miller's (1993) system is sufficient to model simple RT given the specification of the time required for response production, t_2. They assumed stage rates vary from choice-to-choice with distributions, Z_s, that are positive, independent and identically distributed, with finite first and second moments, μ_s and σ_s^2. It follows from the central limit theorem that the rate distribution $V \sim \exp(\Sigma_{s=1..S} Z_s)$ can be approximated by $\mathrm{LN}(\mu_V, \sigma_V^2)$, where LN indicates a Lognormal distribution with a mean, μ_V, and variance, σ_V^2 that equal the sums of the first and second central moments, respectively, of the Z_s.

We define the LNR as *a model made up of one or more racing Lognormal accumulators* (i.e., accumulators for which $T \sim \mathrm{LN}$). Of course, when there is only one accumulator there is not really a race, but the single accumulator model is useful for modeling simple RT. Whether motivated by Ulrich and Miller's (1993) flow argument, or simply made as an *ad hoc* assumption, a Lognormal distribution for the rate of evidence accumulation is very mathematically convenient in the linear deterministic accumulation framework. This is so because both the inverse of a Lognormal variable, and the product of independent Lognormal variables, also have a Lognormal distribution. If one assumes, as did Ulrich and Miller, that the distance D is a constant, it follows that $T \sim \mathrm{LN}(\mu, \sigma^2)$, where $\mu = \ln(D) - \mu_V$ and $\sigma^2 = \sigma_V^2$.Consequently, simple RT is predicted to have a shifted Lognormal distribution, where the shift, which defines the lower bound of the distribution, equals $t_1 + t_2$. Note that if accumulators in which $T \sim \mathrm{LN}$ are embedded in a race architecture, choice RT is also predicted to be Lognormal in the limit of high accuracy. This occurs because when accuracy is high the race has no effect on RT distribution because one accumulator (representing the correct response) always wins.

Consistent with these predictions, the shifted Lognormal has a long history of use to describe simple and choice RT distributions (e.g., Woodworth and Schlosberg, 1954), which it does with an accuracy that has been found to be on par with the most widely used descriptive model, the ExGaussian distribution (Ratcliff and Murdock, 1976). The Lognormal distribution is bounded below by zero ($x > 0$) with density:

$$f(x, \mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln(x)-\mu}{\sigma}\right)^2} \quad (3)$$

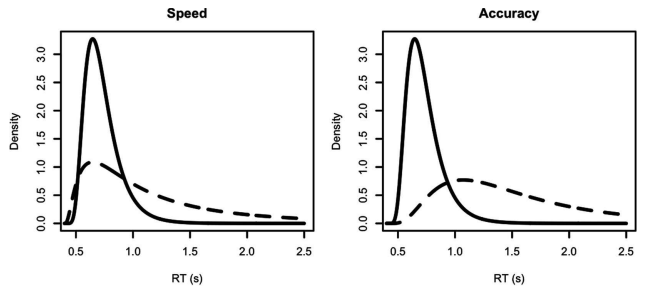


FIGURE 2 | Examples of independent Lognormal distributions, all with shift θ = 0.4 s. Solid lines represent the true (correct response) accumulator distribution, dashed lines the false (error response) accumulator distribution. The left panel represents a speed-emphasis condition, the right panel an accuracy emphasis condition. For the speed condition the (μ, σ^2) parameters for each density are: solid line (−1.2, 0.2), dashed line (−0.5, 0.9). The corresponding mean RTs for correct and error responses are 0.7 and 0.64 s (fast errors) with 75% accuracy. For the accuracy condition the (μ, σ^2) parameters for each density are: solid line (−1, 0.2), dashed line (0, 0.4). The corresponding mean RTs for correct and error responses are 0.78 and 0.84 s (slow errors) with 90% accuracy.

The Lognormal survivor function (S) can be expressed in terms of the standard normal cumulative distribution, Φ, which has rapidly computable approximations:

$$S(x, \mu, \sigma) = 1 - \Phi\left(\frac{1n(x) - \mu}{\sigma}\right) \tag{4}$$

To allow for a lower bound greater than zero a shift parameter, $0 < \theta < \min(x)$, can be added by substituting $(x - \theta)$ for x in (3) and (4). **Figure 2** shows four examples of the shifted Lognormal density.

It is important to observe that this Lognormal form is not uniquely predicted by assuming that distance is a constant. It also follows if distance is a random variable that also has a Lognormal distribution. That is, if $D \sim \exp(Z_D)$, where $Z_D \sim N(\mu_D, \sigma_D^2)$, then $T \sim \mathrm{LN}(\mu, \sigma^2)$, where $\mu = \mu_D - \mu_V$ and $\sigma^2 = \sigma_D^2 + \sigma_V^2$. Similarly, if V is a constant and $D \sim \exp(Z_D)$ $T \sim \mathrm{LN}(\mu, \sigma^2)$, where $\mu = \mu_D - \ln(V)$ and $\sigma^2 = \sigma_D^2$. These observations indicate that it is not possible to determine which of D of V are constant or Lognormal, or whether both are Lognormal, based only on the form of the distribution of T.

In contrast, a Lognormal distribution for T does not follow when D has alternative distributional forms. This is true even when the two cases discussed so far are combined, so that D has a shifted Lognormal form, $D \sim d + \exp(N(\mu_D, \sigma_D^2))$, where d is a constant. Allowing distance to have a shifted Lognormal distribution would provide similar flexibility to the LBA, where the analogous condition is that there is some distance greater than zero between the top of the start-point distribution and the boundary (i.e., $B > 0$). This type of flexibility is necessary for the LBA to be able to account for responding under accuracy emphasis. That is, good fits of the LBA can be obtained with small values of B under speed emphasis (Brown and Heathcote, 2008), but in most other situations estimates of B are substantially greater than zero. The fits of the LNR model reported below enable us to test whether it fails because it lacks similar flexibility (i.e., it does not allow distance to be both variable and to have a minimum value greater than zero).

CHOICE RESPONSE TIME

As in the LBA, the LNR model of N-choice paradigms assumes a race among $n = 1 \ldots N$ linear accumulators. The winning unit triggers its corresponding response production process and effectively inhibits all other choice units instantaneously, so only one response is made. In this section we explicitly derive the likelihood for a two-choice LNR model and point out the relatively straightforward extension required for the N-choice case. We do so with a quite general characterization of the sets of rate and distance parameters defined over choice units in terms of arbitrary finite variance-covariance matrices. We do, however, assume that distance and rate variation can be approximated as independent, which is plausible given they originate from different sources. We also explicitly develop results for an LNR model in which both distances and rates have Lognormal distributions, $D \sim e^{N(\mu_d, \sigma_d^2)}$ and $V \sim e^{N(\mu_v, \sigma_v^2)}$, an so:

$$T = \frac{D}{V} \sim e^{N(\mu_d - \mu_v, \sigma_d^2 + \sigma_v^2)} = e^{N(\mu, \sigma^2)} \tag{5}$$

Note that in (5) and in the following, results for the case where distance is a constant are obtained by assuming parameters related to variability in distance (e.g., σ_d) are set to zero.

Consider a race between two accumulators, so we now have two-vectors of random variable for distances and rates. Distances might be correlated between the accumulators and rates might be correlated between the accumulators, but distances and rates are assumed not to be correlated within an accumulator, or between accumulators. Given MVN (μ,Σ) denotes a multivariate normal random variable with mean vector μ and variance-covariance matrix Σ:

$$\mathbf{d} \sim e^{\mathrm{MVN}(\mu_d, \Sigma_d)},\ \sum\nolimits_d = \begin{bmatrix} \sigma_{d1}^2 & \sigma_{d1.d2}^2 \\ \sigma_{d1.d2}^2 & \sigma_{d2}^2 \end{bmatrix} \tag{6}$$

$$\mathbf{v} \sim e^{\mathrm{MVN}(\mu_v, \Sigma_v)},\ \sum\nolimits_v = \begin{bmatrix} \sigma_{v1}^2 & \sigma_{v1.v2}^2 \\ \sigma_{v1.v2}^2 & \sigma_{v2}^2 \end{bmatrix} \tag{7}$$

Using the fact that a difference between two independent multivariate normal random variables is also multivariate normal, where the mean vector of the difference is the difference of the mean vectors, and the variance/covariance matrix of the difference is the sum of the variance/covariance matrices, we can write:

$$T \sim e^{\mathrm{MVN}(\mu, \Sigma)},\ \mu = \mu_d - \mu_v,$$
$$\sum = \begin{bmatrix} \sigma_{d1}^2 + \sigma_{v1}^2 & \sigma_{d1.d2}^2 + \sigma_{v1.v2}^2 \\ \sigma_{d1.d2}^2 + \sigma_{v1.v2}^2 & \sigma_{d2}^2 + \sigma_{v2}^2 \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & \sigma_{1.2}^2 \\ \sigma_{1.2}^2 & \sigma_2^2 \end{bmatrix} \tag{8}$$

In order to get the likelihood that, say, accumulator 2 wins the race at time x (i.e., hits its boundary at $T = x$ and accumulator 1 has not yet hit its boundary) we need to multiply the marginal density of accumulator 2 by the conditional survivor function of accumulator 1. The marginal distribution of accumulator 2 is Lognormal, $\exp[N(\mu_2, \sigma_2^2)]$, as is the conditional distribution

of accumulator 1. That is, the distribution of T_1 conditional on $T_2 = x$ is Lognormal. Denoting the correlation $\rho = \sigma_{12}^2/(\sigma_1\sigma_2)$:

$$T_1|\,(T_2 = x) \sim e^N\left(\mu_1 + (\sigma_1/\sigma_2)\,\rho\,(\ln(x) - \mu_2)\,,\ \left(1-\rho^2\right)\sigma_1^2\right) \tag{9}$$

For the case of $N > 2$ accumulators the required conditional is multivariate Lognormal. In the $N = 2$ accumulator case:

$$L_2(x) = f(x,\mu_2,\sigma_2)S\left(x, \mu_1 + \frac{\sigma_1}{\sigma_2}\rho\,(ln(x) - \mu_2)\,,\left(1-\rho^2\right)\sigma_1^2\right) \tag{10}$$

The likelihood that accumulator 1 finishes first at time $T = x$ is obtained by exchanging indices in (10).

CORRECT AND ERROR-RESPONSE SPEED

Consistent with many other studies (e.g., Ratcliff and Rouder, 1998), Wagenmaker et al.'s (2008) speed vs. accuracy emphasis manipulation systematically affected the relative speed of correct and error responses; error responses were slower than correct responses under accuracy emphasis and equal or faster under speed emphasis. In order to understand the LNR and LBA fitting results that we report in the next section, it is useful to illustrate how each model can produce fast and slow errors. This illustration particularly focuses on the variability parameters, which play a more important role in the LNR than have analogous variability parameters in previous applications of the LBA.

The LBA predicts fast errors when the distance B (see **Figure 1**) is smaller. Errors are fast in this case because incorrect responses mostly occur when there is a strong initial bias toward the wrong response. This bias can occur on some trials due to random variation in the starting points of the accumulators. Hence, even if the error accumulator has a slow rate it can quickly achieve its boundary, and so produce a fast response. No similar characterization of fast errors is possible for the LNR model as bias (distance) and rate effects combine additively to determine distribution shape.

However, the LNR model can still produce fast errors, as illustrated by the left panel of **Figure 2**, which simulates results from a speed-emphasis condition. In this panel the variance parameter for the false (error response) accumulator is twice that of the true (correct response) accumulator. Despite having a much slower mean, and consequently only winning 25% of races (i.e., a 25% error rate), its greater variance causes the false accumulator to produce fast responses when it wins the race. This can be seen as the higher density for the false accumulator (dashed line) than the true accumulator (solid line) on the left of the RT distributions. As a result, when errors occur they tend to be faster than correct responses (0.64 vs. 0.7 s in the example). The right-hand panel of **Figure 2** shows that greater false than true accumulator variance does not always produce fast errors. In this example performance is slower overall and more accurate (10% errors) with the mean for error responses being greater than for correct responses (0.84 vs. 0.78 s in the example).

These considerations suggest that the LNR will require different variances for true and false accumulators in order to fit manipulations that affect the relative speed of correct and error responses. Clearly these differences must arise from differences in rate variance, as whether an accumulator represents a correct or error response is determined by the stimulus. We consider how such differences might arise after reporting the results of model fitting in the next section.

Note that fast errors could also occur in the LBA if the *sv* (rate standard deviation) parameter is greater for false than true accumulators. This possibility has not been tested before because previous applications of the LBA have either assumed a fixed value of *sv* (e.g., unity), or that the same estimated value applies for all accumulators and experimental conditions. These assumptions about *sv* were motivated by the fact that an accumulation-related LBA parameter must be fixed in order for the model to be identifiable. However, in a design such as was used by Wagenmakers et al. (2008) identifiability requires only that one parameter value be fixed for one accumulator in one condition (see Donkin et al., 2009b). The next section reports fits of LBA models that only place this minimal constraint on *sv* so that we can examine all potential explanations for fast errors.

MODEL TESTING

We fit the LBA and LNR models to data from Wagenmaker et al.'s (2008) experiment one, where participants made decisions about whether a string of letters constituted a word. These lexical decisions were made about four types of stimuli, non-words (nw) and high-frequency (hf), low-frequency (lf), and very low-frequency (vlf) words. Participants made decisions either under speed or accuracy emphasis instructions in different experimental blocks. Accuracy blocks were preceded by the message "Try to respond accurately" and "ERROR" was displayed after each wrong response. Speed blocks were preceded by the message "Try to respond accurately" and "TOO SLOW" was displayed after each response slower than 0.75 s. We report analyses of data from 17 participants (31,412 data points) in their Experiment 1, including the 15 participants analyzed in Wagenmakers et al. (2008) and two extras (we thank Eric-Jan Wagenmakers for supplying this data).

FITTING METHODS

As we examined a large number of parameterizations of each model that varied widely in the number of estimated parameters, we compared fits both within and between model types using the AIC and BIC model selection criteria (Myung and Pitt, 1997). These criteria measure badness-of-fit using twice minus the log-likelihood, which we will call the deviance (D). A model is selected from amongst a set of models if it has the lowest AIC or BIC. Both criteria add to the deviance a penalty that increases with the number of parameters estimated to obtain a fit. Hence, better fitting models (i.e., with a smaller D) that have a larger numbers of parameters may not be selected if the extra parameters do not produce a sufficiently large improvement in fit (i.e., reduction in D). We consider both AIC and BIC criteria as they have different merits flowing from the underlying quantities that they approximate (see Burnham and Anderson, 2004; Vrieze, 2012), with BIC generally preferring simpler models.

Our LNR models assumed there were no correlation among rates or among distances, so fits were obtained using the correspondingly simplified from of (10) corresponding

to (2), i.e., $F_1(x) = f(x, \mu_1, \sigma_1)S(x, \mu_2, \sigma_2)$ and $F_2(x) = f(x, \mu_2, \sigma_2)S(x, \mu_1, \sigma_1)$. Although we argued correlations are plausible, in this initial exploration we fixed them at zero for three reasons: (1) to make understanding the LNR model easier, (2) to determine if correlation is necessary to accommodate the benchmark effects present in Wagenmaker et al.'s (2008) data, and (3) to make the LNR and LBA models more comparable.

As well as the standard LBA model described by Brown and Heathcote (2008) we also fit two variations that guaranteed a response on every trial. In the first variation rates for all accumulators were guaranteed to be positive on every trial by sampling them from univariate normal distributions truncated below at zero, and in the second at least one sampled rate was guaranteed to be positive on every trial by sampling the rates from a truncated multivariate normal distribution[1]. All of these LBA models produced similar fits and parameter estimates, so we report fits from the second variant for two reasons: (1) it is most directly comparable to the LNR model, which also has positive rates for all accumulators on every trial and (2) it is of interest in itself as an alternative to the LNR for solving the issue of potential non-responding in the original LBA model.

As it is not our focus here we do not provide a direct comparison with fits of the RDM reported by Wagenmaker et al.'s (2008). Note, however, that Donkin et al. (2011b) found similar quality fits to this data for Ratcliff diffusion and LBA models with similar numbers of parameters. In this comparison both parameterizations were strongly constrained based on past findings about these models. In contrast, because past findings are not available for the LNR model, and our aim here is exploratory, our most complex LNR model was very flexibly parameterized. It freely estimated both μ (mean) and σ^2 (variance) parameters over a false vs. true accumulator factor (C), a stimulus type factor (W) with four levels (hf, lf, vlf, and nw), and an instruction emphasis factor (E) with two levels (speed vs. accuracy), so in total there are potentially $2 \times 4 \times 2 = 16$ estimates of each type. Note that C factor represents whether, for a given stimulus, a parameter corresponds to the accumulator for the correct (true) response (i.e., the word accumulator for a word stimulus or the non-word accumulator for a non-word stimulus) or for an error (false) response (i.e., the word accumulator for a non-word stimulus or the non-word accumulator for a word stimulus).

In order to allow similar flexibility for the LBA, and so that we can compare effects on analogous parameters, we report fits of an LBA model with a less constrained parameterization than is conventional. The LBA B and A parameters were allowed to vary with a word vs. non-word response accumulator factor (lR, standing for "latent response"), in order to accommodate response bias. We also allowed these parameters to vary as a function of the speed vs. accuracy emphasis factor (E), as the analogous quantity in the LNR model, the mean and variance of distance, have the same freedom to vary as components of the Lognormal mean and variance parameters. LBA rate mean (v) and variability (sv) were allowed to vary with C, E, and W as the analogous LNR quantities, the rate mean and variance, are components of the LNR mean and variance parameters.

We also allowed residual time to vary with emphasis instructions (E) for both models as this was favored by AIC and BIC model selection in most cases (see Rinkenauer et al., 2004, for evidence supporting an effect of emphasis instructions on response production). We denote residual time as $t0$, as it corresponds to the lower bound of the distribution of RT. We used this notation rather than by the conventional *Ter* notation because, at least in the flow interpretation of the LNR, stimulus-encoding time might be seen as part of the accumulation process, in which case $t0$ should be thought of as the sum of response production time and the "dead-time" between stimulus onset and the first stimulus contingent change in the firing rates of sensory neurons.

In summary, there were 34 estimated LNR parameters (16 each for μ and σ^2 and two residual time parameters) compared to 41 LBA parameters (16 v, 15 sv, 4 B and 4A, and two residual time parameters). Note that we fixed the intercept of sv estimates at one (for the false accumulator in the accuracy condition for high-frequency words) to make the LBA model identifiable, so there are 15 rather than 16 estimated sv parameters. In order to compactly refer to models we useR (R Development Core Team, 2012) linear model notation adapted to our multiple parameter-type setting. For example the notation $B \sim \mathrm{lR}^* E$ indicates estimation of the main effects of lR and E, and their interaction ($\mathrm{lR} \times E$). Similarly $v \sim E^* W^* C$ indicates estimation of the three main effects, three two-way interactions ($E \times C$, $E \times W$, and $W \times C$), and one three-way interaction ($E \times W \times C$). Using "&" to indicate a join between parameterizations for different parameter types, we denote the 34 parameter LNR model as $\mu \sim E^* W^* C$ & $\sigma^2 \sim E^* W^* C$ & $t0 \sim E$ and the 41 parameter LBA model as $B \sim \mathrm{lR}^* E$ & $A \sim \mathrm{lR}^* E$ & $v \sim E^* W^* C$ & $sv \sim E^* W^* C$ & $t0 \sim E$.

The models were fit to each participant's data separately using the exact maximum-likelihood-based methods described in detail in Donkin et al. (2011a)[2]. This method fits a hierarchy of simplified models that are special cases of the most complex ("top") model (i.e., the 34 parameter LNR model and the 41 parameter LBA model). The simplest LNR model in the hierarchy estimated the same parameters of each type for all conditions, except that it allowed different mean rate parameters for true and false accumulators (i.e., model $\mu \sim C$ & $\sigma^2 \sim 1$ & $t0 \sim 1$, where "$\sim$1" indicates an intercept-only estimate). Similarly, the simplest LBA model allowed only mean rate to differ between true and false accumulators ($B \sim 1$ & $A \sim 1$ & $v \sim C$ & $sv \sim 1$ & $t0 \sim 1$).

Best fitting parameters for the simplest models were used as starting points when searching for best fits of models that estimated the effect of one extra factor for one type of parameter. Best fits for these models were used in turn as starting points for fits of

[1]Density and cumulative density functions for the first "positive LBA" are obtained by dividing the corresponding expressions in Brown and Heathcote (2008) by $\Phi(v/sv)$. Heathcote and Hayes (2012) give details of the second, which introduces a dependency between rates for different accumulators.

[2]We limited the influence of outlying observations, which can be problematic for exact maximum-likelihood methods (Ratcliff and Tuerlinckx, 2002), by placing a floor on the likelihood of an observation corresponding to the assumption of a small (10^{-5}) probability that participant's data was contaminated by unbiased guesses with an RT drawn from a uniform distribution over the range from 0.3 to 2.5 s beyond which Wagenmakers et al. (2008) censored their data. We experimented with contamination probabilities up to 1% and obtained similar results.

models that estimated one additional factor, and so on up to the most complex model. For example the simplest LNR model, $\mu \sim C$ & $\sigma^2 \sim 1$ & $t0 \sim 1$, provides a start point for six models: $\mu \sim W^*C$ & $\sigma^2 \sim 1$ & $t0 \sim 1$, $\mu \sim E^*C$ & $\sigma^2 \sim 1$ & $t0 \sim 1$, $\mu \sim C$ & $\sigma^2 \sim C$ & $t0 \sim 1$, $\mu \sim C$ & $\sigma^2 \sim W$ & $t0 \sim 1$, $\mu \sim C$ & $\sigma^2 \sim E$ & $t0 \sim 1$, and $\mu \sim C$ & $\sigma^2 \sim 1$ & $t0 \sim E$. One or more of these (nested) models then provided start points for further (nesting) models with three free factors and so on up to the top model.

The aim of fitting nesting models from immediately nested models is to avoid getting stuck in local minima and so to obtain true maximum-likelihood fits for more complicated models by fitting them from a variety of plausible starting points. We cannot prove that the method always achieves its aim but have found it does so in simulation studies fitting model of similar complexity to those examined here when using similar sample sizes to the data examined here. In contrast, we have found other less thorough methods often fail badly, and although hand tuning can often remedy such problems that is not feasible when exploring large sets of models. For example, here we fit all possible combinations of factors for different parameters, resulting in 64 LNR model and 1024 LBA models per participant (i.e., 18,496 models in total over 17 participants) and this was done for a number of variants (e.g., different types of LBA models and different levels of response contamination).

In total the results reported here were based on 193 LNR fits per participant and 5121 LBA fits per participant (90,338 fits in total). The R (R Development Core Team, 2012) *optim* function was used for each fit with default settings. It was repeatedly applied (starting each time from the previous solution) until the log-likelihood increased by less than 0.1. All model hierarchies were checked to ensure that nested models had an equal or higher deviance (within numerical tolerance) than nesting models, and any exceptions corrected by refitting.

MODEL SELECTION

Table 1 reports badness-of-fit (deviance) and AIC and BIC model selection results relative to the best model according to that statistic (i.e., the model with a zero in the corresponding column of **Table 1**). The table reports model selection results for three hierarchies. The first two are the full hierarchies below the top LBA ($B \sim \mathrm{lR}^*E$ & $A \sim \mathrm{lR}^*E$ & $v \sim E^*W^*C$ & $sv \sim E^*W^*C$ & $t0 \sim E$) and LNR ($\mu \sim E^*W^*C$ & $\sigma^2 \sim E^*W^*C$ & $t0 \sim E$) models. The third hierarchy consists of all models nested within an LBA model that enforces the conventional assumption that speed vs. accuracy emphasis does not influence evidence-rate-related parameters (i.e., v and sv): $B \sim \mathrm{lR}^*E$ & $A \sim \mathrm{lR}^*E$ & $v \sim W^*C$ & $sv \sim W^*C$ & $t0 \sim E$. We call this set of models the "conventional LBA" hierarchy, although it is still somewhat more flexible than most previous LBA applications as it allows changes in the A as well as the B parameter with emphasis.

Table 1 reports results for the most flexible model in each hierarchy its upper section, results for the best (smallest) AIC model in each hierarchy in the middle section, and results for the best (smallest) BIC model in each hierarchy in its lower section. Within each section the first two models are from the full LBA and LNR hierarchies, respectively, and the last model is from the conventional LBA hierarchy. As might be expected, the full LBA model, with the largest number of parameters, has the best fit (lowest deviance). Models from this same LBA hierarchy were also selected by AIC and BIC. Note that this is not because the selected models are more flexible as indexed by number of parameters; the best overall models have the same number parameters, or fewer, than the best models from the other hierarchies. Selection between the conventional LBA and LNR hierarchies is more equivocal, the LNR clearly wins on AIC but just looses on BIC.

The AIC and BIC results from the full LBA hierarchy are consistent in selecting the $v \sim W^*C$ and $sv \sim E^*C$ parameterization for rate-related parameters. That is, there is a selective influence of stimulus type on mean rate and selective influence of emphasis on rate variability, with true and false accumulators differing on both types of parameter. The conventional explanation of emphasis effects in terms of the boundary (B) is also consistently supported. However, an effect of emphasis on start-point noise (A) receives support only from AIC, indicating it has a weaker effect. The same is true of response bias (lR) effects on B, and there is no support for any effect of response bias on A. Support for an effect of emphasis on residual time ($t0$) was also inconsistent.

Model selection for the conventional LBA hierarchy produces results largely consistent with those for the full hierarchy. Stimulus type has a selective influence on mean rate (v), with true and false accumulators differing on both types of rate parameter. Both criteria support an effect of emphasis on the B, A, and $t0$ parameters, consistent with these parameters making up some of the fit provided by the sv parameter in the full hierarchy. Stronger response-bias effects, particularly on the B parameter, are also evident.

Within the LNR hierarchy both AIC and BIC pick the same model, which drops only the effect of stimulus type on variance relative to the full model. These results are largely consistent with effects selections for analogous parameters in the full LBA hierarchy (i.e., neglecting distance effects μ is analogous to $-\ln(v)$ and σ to sv), in that stimulus type selectively influences μ and both μ and σ differ between true and false accumulators. Two differences are that emphasis affects both μ and σ, not just the variability parameter as in the LBA, and it also consistently affects $t0$.

In summary, for the LBA, model selection results indicate that, in contrast to the conventional assumption, the speed vs. accuracy emphasis effect is best explained when rate variability changes with emphasis. Model selection also supports a role for the conventional boundary-based mechanism, but less so for the start-point based mechanism. Further, in contrast to previous applications where only the mean rate has been allowed to differ between correct and false accumulators, there was clear support for a difference in rate variability between accumulators. These results show that the LBA, at least to some degree, uses the same mechanism as the LNR to explain differences in the relative speed of correct and error responses, differences in the variability associated with true and false accumulators. The classes of models differ, at least in terms of the AIC and BIC selected models, in that the LNR mean parameter also has a role in explaining the emphasis effect.

MODEL FIT

We focus on the fit of the model selected by both AIC and BIC from the full LNR hierarchy and by AIC for the LBA hierarchy in order to determine how well these relatively simple models capture the

Table 1 | Model selection statistics for the LBA and LNR top models and the conventional LBA top model (upper section) and the best AIC (middle section) and BIC (bottom section) model in each hierarchy.

Model	NP	Δ*D*	ΔAIC	ΔBIC
$B \sim \text{lR}*E$ & $A \sim \text{lR}*E$ & $v \sim E*W*C$ & $sv \sim E*W*C$ & $t0 \sim E$	41	0	15	210
$\mu \sim E*W*C$ & $\sigma^2 \sim E*W*C$ & $t0 \sim E$	34	19	20	156
$B \sim \text{lR}*E$ & $A \sim \text{lR}*E$ & $v \sim W*C$ & $sv \sim W*C$ & $t0 \sim E$	25	68	51	112
$B \sim \text{lR}*E$ & $A \sim E$ & $v \sim W*C$ & $sv \sim E*C$ & $t0 \sim E$	19	29	0	11
$\mu \sim E*W*C$ & $\sigma^2 \sim E*C$ & $t0 \sim E$	22	36	13	49
$B \sim \text{lR}*E$ & $A \sim \text{lR}*E$ & $v \sim W*C$ & $sv \sim C$ & $t0 \sim E$	19	79	50	61
$B \sim E$ & $A \sim 1$ & $v \sim W*C$ & $sv \sim E*C$ & $t0 \sim 1$	15	59	22	0
$\mu \sim E*W*C$ & $\sigma^2 \sim E*C$ & $t0 \sim E$	22	36	13	49
$B \sim \text{lR}*E$ & $A \sim E$ & $v \sim W*C$ & $sv \sim C$ & $t0 \sim E$	17	83	51	45

NP, number of parameters per participant; ΔD, $D - min(D)$, where D indicates deviance; ΔBIC, $BIC - min(BIC)$; ΔAIC, $AIC - min(AIC)$. D, BIC, and AIC values are summed over participants then divided by number of participants.

pattern of effects on correct RT and error rates (**Figure 3**) and error RT (**Figure 4**). The ability of the models to capture RT distribution is indicated by displaying results for the 10th, 50th, and 90th percentiles[3] of RT distribution. Fits to intermediate percentiles are of similar quality to the percentiles shown, and are omitted for display clarity.

The goodness-of-fit figures average corresponding observed and predicted values over participants. Predicted values corresponding to each observed RT were obtained by simulating 100 times as many values as were observed. For example, if there were 100 correct and 9 error responses, 10,000 correct and 900 error RTs were simulated and predictions obtained from the order statistics of the simulated data (e.g., the fastest of the 90 error RTs was assumed to correspond to the 10th percentile of the simulated sample; in general the ith ordered observation from N was equated with the $i/(N+1)$ quantile of the simulated data). Proportion correct was calculated from the full set of samples required to obtain the predicted RT values. Observed and predicted values were treated in the same way to get the averages in the figures; RT averages were calculated omitting missing values.

Figure 3 shows that the best LNR and AIC-best LBA models accurately capture correct RT distribution. The simpler BIC-best LBA model, whose results are not shown, provided an equally accurate account of correct RT to the AIC-best LBA model. The word frequency manipulation provides a strong test of the simplifying assumption, which we made in fitting both the LNR and LBA models, that residual time is a constant, given the Ratcliff diffusion requires variable residual time to account for the effect of word frequency on fast correct responses (i.e., the 10th percentile). Clearly neither model requires variation in residual time to provide a very accurate account of fast correct responses.

Figure 3 shows that the best LNR model accurately captures error rates, although it slightly under-predicts the accuracy-speed effect for high-frequency words, although overall it captures 94% of the average difference. The AIC-best LBA model does not perform quite as well, capturing 83% of the average difference. Although this might be attributed to the AIC-best LBA model being simpler (19 parameters) than the best LNR model (22 parameters), the more complicated top LBA model (41 parameters) does not do much better (86%). The simpler BIC-best LBA model (15 parameters), which drops the effect of emphasis on A, does considerably worse (67%). We also included in the comparisons in **Figure 3** the top model from the conventional LBA hierarchy. This model is more complicated (25 parameters) than either of the comparison models, and clearly it is able to accurately capture the correct RT distribution. However, it captures only 20% of the average speed-accuracy difference in error rates.

Figure 4 shows that the best LNR model and AIC-best LBA model captures the overall pattern in error RT in the speed-emphasis condition. In the accuracy emphasis condition there are fewer errors, particularly in for high-frequency words and nonwords, so RT estimates are noisy. However, both models appear to systematically underestimate overall variability, resulting either in overestimation of the 10th percentile or underestimation of the 90th percentile. As was the case for correct RT, the simpler BIC-best LBA model, whose results are not shown, provides a similar account of error RT to the AIC-best LBA model.

Figure 5 focuses on the relative speeds of (median) correct and error RT, with confidence intervals omitted to make the pattern of results clearer. Both the best LNR and best LBA models capture the general pattern of slower error than correct responses under accuracy emphasis and faster error than correct responses under speed emphasis.

DISCUSSION

Our fits show that the LNR model is able to provide quite an accurate descriptive account of all of the effects in Wagenmaker et al.'s (2008) first experiment. Recall that the LNR model we fit is simplified in two senses; it assumes no variability in residual time, and it assumes inputs and distances are uncorrelated across accumulators. It is also without the LBA model's freedom to

[3]Models were not fit to these or any other set of percentiles; percentiles are only used to summarize goodness-of-fit. In contrast, the Ratliff diffusion model is usually fit to the 10th, 30th, 50th, 70th, and 90th percentiles, as that requires only calculation of the CDF, which is much quicker to compute for the RDM than the likelihood. We used exact maximum likelihood because this allows us to compute exact AIC and BIC values rather than relying on a multinomial-maximum-likelihood approximation (see Heathcote and Brown, 2004; Speckman and Rouder, 2004) required for percentile methods.

combine trial-to-trial variability in distance with a distance greater than zero. That none of these restrictions caused a bad fit is not, by itself, evidence that these restrictions might not have to be relaxed in other situations. However, it does encourage the wider application of this simple and tractable form of the LNR model.

Figure 6 shows average mean (μ) and variance (σ^2) parameter estimates for the best LNR model (i.e., the model selected by both AIC and BIC). The false accumulator (i.e., the accumulator corresponding to the error response) has a greater mean parameter than the true accumulator. Note that this results in an increase in both the mean and variance of boundary-crossing times for the false accumulator, because the μ parameter affects both. We also found that the false accumulator had a greater variance than the true accumulator, consistent with the example given in **Figure 2**. Again, this results in an increase in both the mean and variance of boundary-crossing times for the false accumulator, because the parameter σ parameter also affects both. These differences in μ and σmight arise, due to differences in the distributions of evidence accumulation rates for true and false accumulators. To provide a concrete example, suppose a template-matching process produces evidence for each response. If poorer matches produce outputs that at not only weaker on average but also more variable than outputs for strong matches the pattern displayed in **Figure 6** could be found. In support

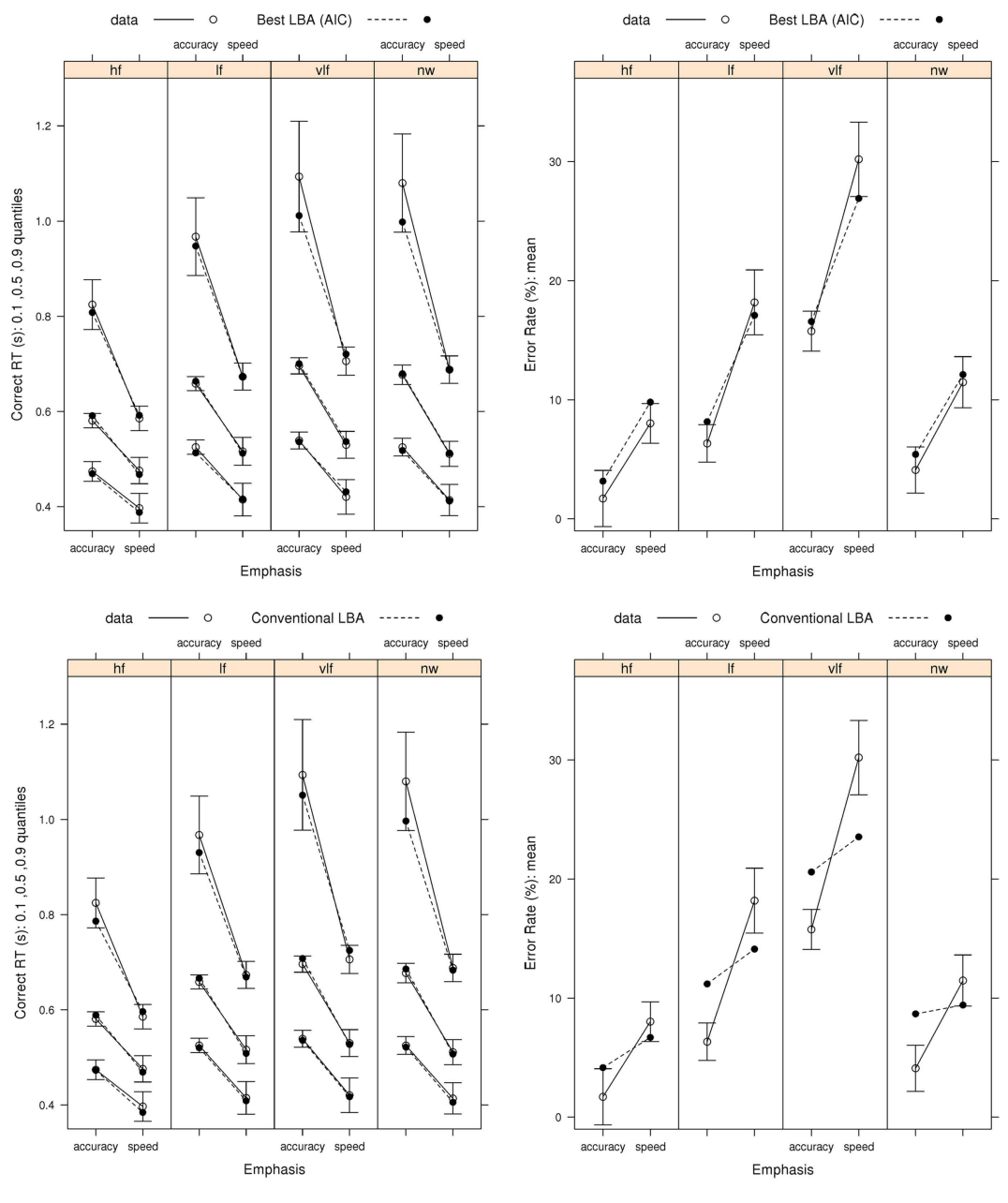


FIGURE 3 | Continued

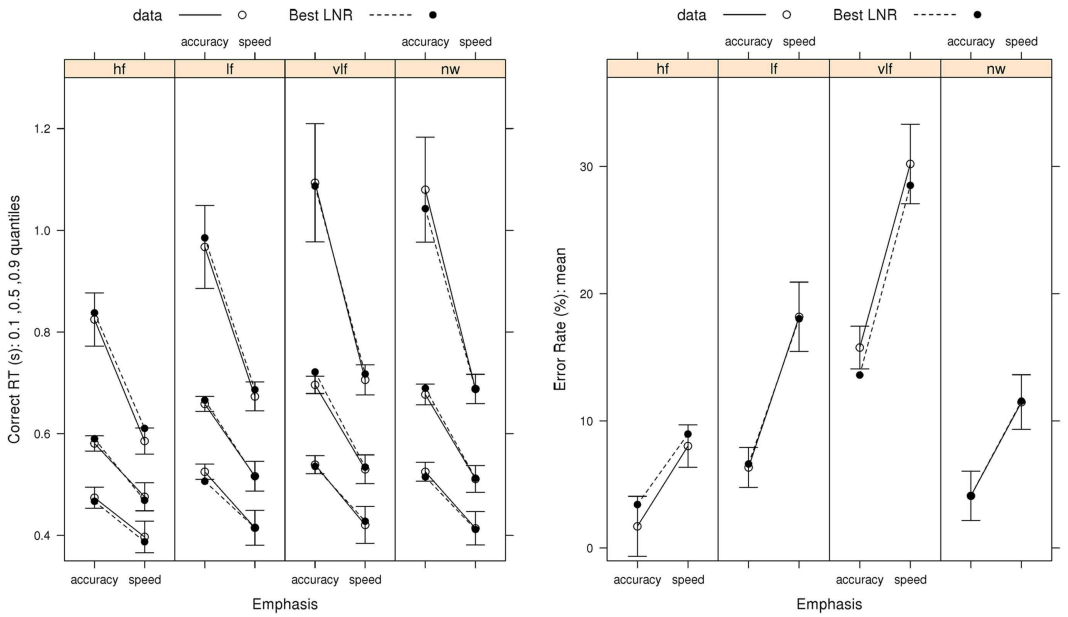


FIGURE 3 | Observed RT distributions (10th, 50th, and 90th percentiles, left column) and error rates (right column) for Wagenmaker et al.'s (2008) Experiment 1 with bias-corrected within-subject 95% confidence intervals (Morey, 2008), and fits averaged over participants for the best (selected by AIC and BIC) LNR ($\mu \sim E*W*C$ & $\sigma^2 \sim E*C$ & $t0 \sim E$) and best (selected by AIC) LBA ($B \sim \mathrm{IR}*E$ & $A \sim E$ & $v \sim W*C$ & $sv \sim E*C$ & $t0 \sim E$) models and the top conventional LBA ($B \sim \mathrm{IR}*E$ & $A \sim \mathrm{IR}*E$ & $v \sim W*C$ & $sv \sim W*C$ & $t0 \sim E$) model: hf, high-frequency words; lf, low-frequency words, vlf, very low-frequency words; nw, non-words.

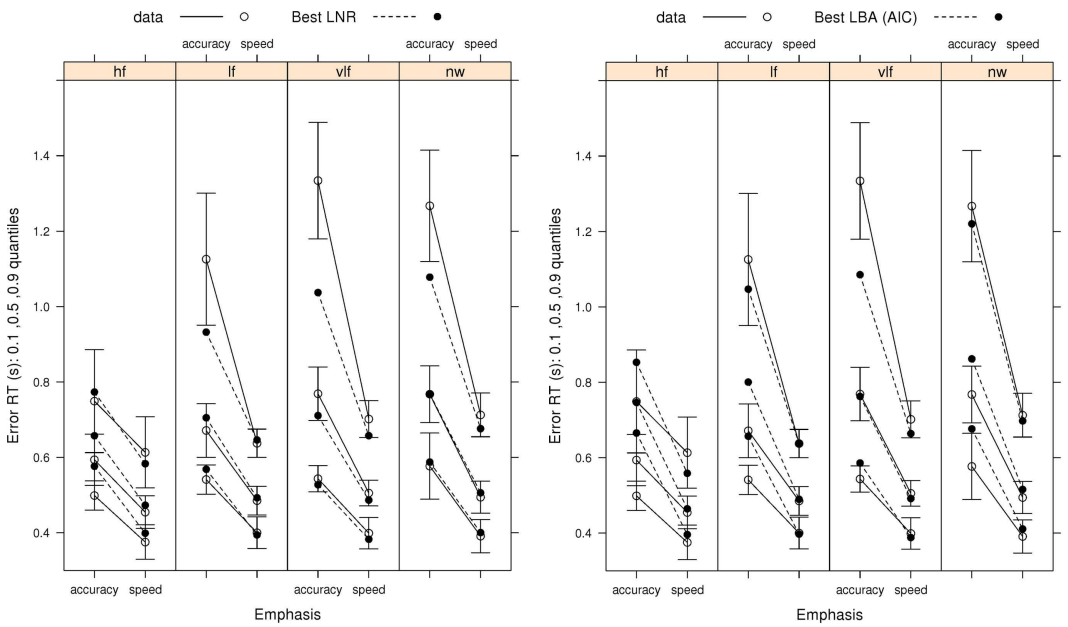


FIGURE 4 | Observed error-response RT distributions (10th, 50th, and 90th percentiles) Wagenmaker et al.'s (2008) Experiment 1 with bias-corrected within-subject 95% confidence intervals (Morey, 2008), and fits averaged over participants for the best (selected by AIC and BIC) LNR ($\mu \sim E*W*C$ & $\sigma^2 \sim E*C$ & $t0 \sim E$) and best (selected by AIC) LBA ($B \sim \mathrm{IR}*E$ & $A \sim E$ & $v \sim W*C$ & $sv \sim E*C$ & $t0 \sim E$) models: hf, high-frequency words; lf, low-frequency words; vlf, very low-frequency words; nw, non-words.

of this possibility, there was a high positive correlation ($r = 0.80$, $p < 0.001$) between the mean and variance estimates for the top LNR model.

For both the LNR mean and variance, **Figure 6** shows that the increase from speed to accuracy conditions is larger for false than true accumulators. This uneven increase cannot be explained by a

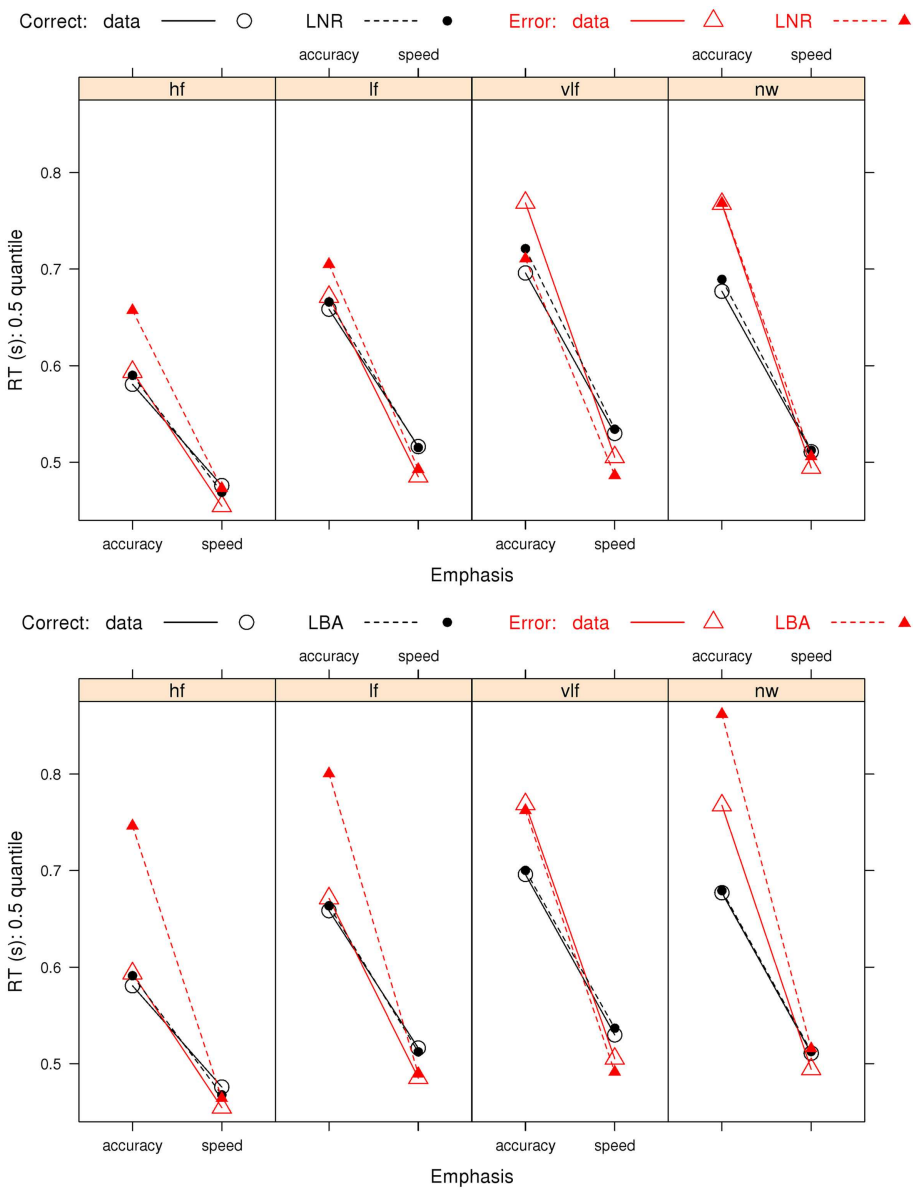


FIGURE 5 | Observed and predicted median RTs for correct and error responses in Wagenmaker et al.'s (2008) Experiment 1 and fits averaged over participants for the best (selected by AIC and BIC) LNR $\mu \sim E*W*C$ & $\sigma^2 \sim E*C$ & $t0 \sim E$) and best (selected by AIC) LBA ($B \sim IR*E$ & $A \sim E$ & $v \sim W*C$ & $sv \sim E*C$ & $t0 \sim E$) models: hf, high-frequency words; lf, low-frequency words, vlf, very low-frequency words; nw, non-words.

selective influence of speed vs. accuracy emphasis on distance (i.e., the LNR analog of a boundary effect), as that must cause an equal effect on false and true accumulators (see Eq. 8). That is, accumulators are only "false" and "true" with respect to the stimulus, so the interactions evident in **Figure 6** indicate that emphasis directly affects the mean and variability of the rate at which information is extracted from a stimulus. Clearly this conclusion is at odds with the way that speed-accuracy trade-off has traditionally been accounted for by evidence accumulation models. Note, however, that these results do not rule out some change in distance under accuracy emphasis, as long as this change is accompanied by appropriate changes in rates to yield the total effect in **Figure 6**.

Perhaps more surprisingly, model selection failed to support the traditional boundary-only account for the LBA; the best model according to both AIC and BIC included emphasis effects on the LBA *sv* parameter. **Figure 7** shows the rate-related parameter estimates for the AIC-best LBA model (BIC-best estimates were almost identical). The *sv* estimates were greater for the true than false accumulator, and vice versa for the *v* estimates, consistent with the findings for the analogous LNR parameters. In contrast to the LNR model, speed emphasis affected false accumulator variability in the opposite direction, increasing it greatly in the speed condition. Regardless, of this difference, clearly neither LBA nor LNR results are in line with conventional assumptions.

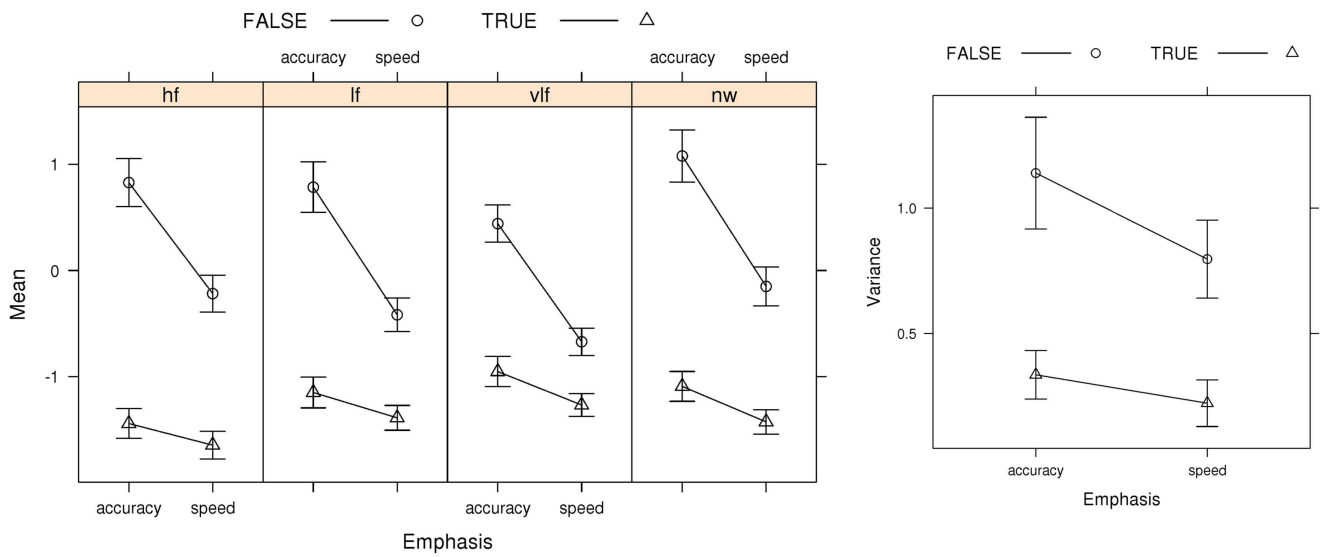


FIGURE 6 | Best LNR ($\mu \sim E*W*C$ & $\sigma^2 \sim E*C$ & $t0 \sim E$) μ (Mean), and σ^2 (Variance) parameter estimates averaged over participants, with bias-corrected within-subject 95% confidence intervals (Morey, 2008). FALSE, error-response accumulator; TRUE, correct-response accumulator; hf, high-frequency words; lf, low-frequency words; vlf, very low-frequency words; nw, non-words.

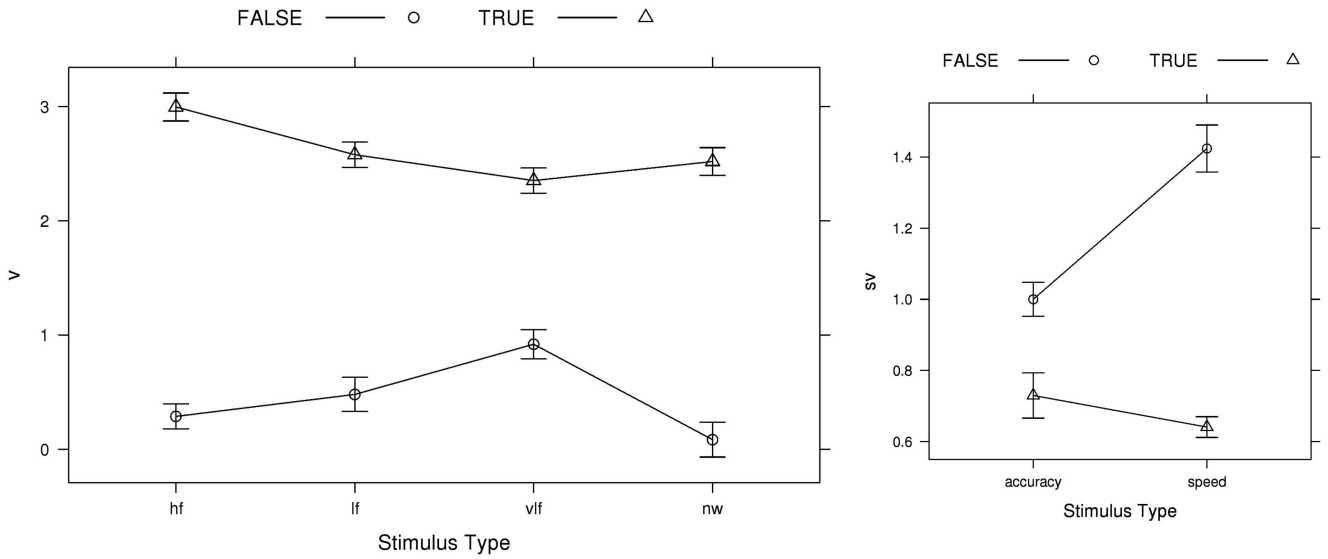


FIGURE 7 | AIC-best LBA ($B \sim IR*E$ & $A \sim E$ & $v \sim W*C$ & $sv \sim E*C$ & $t0 \sim E$) *v* (mean rate) and *sv* (rate standard deviation) parameter estimates averaged over participants, with bias-corrected within-subject 95% confidence intervals (Morey, 2008). FALSE, error-response accumulator; TRUE, correct-response accumulator; hf, high-frequency words; lf, low-frequency words; vlf, very low-frequency words; nw, non-words.

Our analysis of model fit indicated that effects of emphasis on the LBA boundary (B) and start-point noise (A) parameters, as well as on the *sv* parameter, were required to describe the large effects on error rate shown in **Figure 3**. **Figure 8** plots these parameter estimates for the AIC-best LBA model. Speed emphasis caused a decrease in B, and to a lesser degree an increase in A. **Figure 8** also plots results for the top conventional LBA model, which does not allow emphasis to affect any rate parameters. The pattern is similar, except that the effects are larger and the overall level of is B lower.

In summary, when free do to so, the LBA does not attribute the speed-accuracy trade-off induced by instructions in Wagenmaker et al.'s (2008) experiment purely to a change in response caution. Although changes affecting the level of errors caused by start-point noise (i.e., changes in the A and B parameters) could in principle accommodate the large observed differences in error rate, they are not able to do so while also providing an accurate account of RT

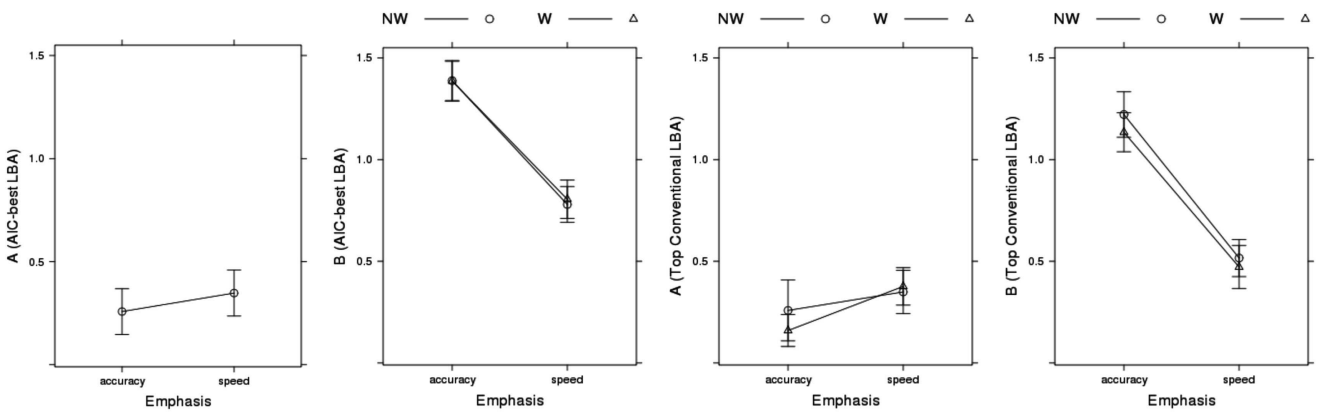


FIGURE 8 | Emphasis effects on AIC-best (*B* ~ IRE* & *A* ~ *E* & *v* ~ *W***C* & *sv* ~ *E***C* & t0 ~ *E*) and top conventional LBA model (*B* ~ IR**E* & *A* ~ IR**E* & *v* ~ *W***C* & *sv* ~ *W***C* & *t*0 ~ *E*) start-point noise (*A*) and boundary (*B*) parameter estimates averaged over participants, with bias-corrected within-subject 95% confidence intervals (Morey, 2008).** NW, non-word accumulator; W, word accumulator.

distribution in this data. In particular, the strength of start-point noise in the LBA is limited because if it is too strong it can produce a more uniform RT distribution than is observed. Instead, the LBA, like the LNR, explains much of the speed-accuracy trade-off effect by an increase in the overlap of false and true distributions under speed emphasis.

GENERAL DISCUSSION

In this paper we have articulated framework for modeling simple-choice behavior using linear deterministic evidence accumulation. Within this framework the evidence is approximated as deterministic during accumulation, and the time for an evidence accumulator to reach its boundary is characterized by a ratio of two variables, one or more of which can vary randomly between-choice trials ("between-choice noise"). The numerator variable is the difference between the level of evidence at the start of accumulation and the boundary ("distance") and the denominator variable is the rate of accumulation. Accumulators participate in a race that is non-interactive in the sense that that the state of one accumulator does not directly affect the state of other accumulators during accumulation. Response selection and RT are determined by the times at which evidence totals first cross one or more accumulator boundaries.

Different types of models within the linear deterministic framework correspond to different assumptions about the distributional forms of the variables in the ratio. Two previously proposed models that fall within the framework, Carpenter's (1981) LATER model and Brown and Heathcote's (2008) LBA model assume normally distributed rates; LATER also assumes distance is a constant whereas the LBA assumes it is uniformly distributed. Model types can also differ in three other ways: in how the random variables are related across accumulators, in the number of racers, and in how boundary crossings determine response selection. In the setup used in most previous applications of the LBA the random variables are uncorrelated, there is a one-to-one mapping of racers to responses, and the winning racer (i.e., the first to cross its boundary) triggers the corresponding response. However, it is also possible to have more racers than responses, to have responding contingent on more than one boundary crossing (see Eidels et al., 2010, for an LBA based example), and to have correlations among the random variables for different accumulators.

In this paper we proposed a new type of linear deterministic model, the LNR, and compared it to a slight variant of previous LBA models through fits of both models to Wagenmaker et al.'s (2008) speed-accuracy trade-off experiment. The version of the LNR that we focused on is mathematically simple because the random-variable ratio has the same distributional form as its constituents, a Lognormal distribution. A Lognormal ratio also results if the numerator (distance) is a constant and only the denominator (rate) is a Lognormal random variable or vice versa. The Lognormal form makes it tractable to allow rates and boundaries to be correlated over accumulators. However, in our initial exploration we assumed no correlation to see if this simple form of the LNR could still provide adequate fits. The LBA variant that we fit has a normal rate distribution truncated at zero, so on every trial all accumulators have a rate greater than zero. The latter property also applies to the LNR, as the Lognormal distribution is positive, so in both cases a response must eventually be selected on every choice trial.

The findings reported in this paper – both theoretical results related to the LNR and empirical results from fitting the LNR and LBA – bear most directly on the deterministic assumption made by our framework. On the theoretical front, the "flow" motivation of the LNR has implications for the division between stimulus encoding and response selection stages, and the effect of assuming a Lognormal distance has implications for the way in which deterministic models explain speed-accuracy trade-off by integrating out start-point noise. On the empirical front, our results point to the utility of the LNR as a tractable descriptive model and also highlight the issue of whether speed vs. accuracy emphasis instructions have effects beyond changes in the evidence boundary. Before addressing these implications, we first discuss the other fundamental assumption made by our framework, that accumulation is linear. As exemplified by the BA model (Brown and Heathcote,

2005a), there is no necessary relationship between the deterministic and linear assumptions, but linearity is closely related to another assumption, that accumulation is non-interactive.

LINEAR ACCUMULATION

The nervous system is pervaded by non-linear dynamics, caused by factors such as imperfect ("leaky") neural integration and recurrent self-excitation and lateral inhibition. Behavioral evidence for non-linear evidence accumulation has been sought using stimuli with non-stationary discriminative information, that is, stimuli that can briefly switch between the choices they support during the time course of accumulation. Usher and McClelland's (2001) seminal experiments investigated the influence of the time within a stream of information favoring one response at which a brief pulse of contradictory information occurred. They found strong differences between participants ranging from leaky integration (i.e., a greater influence for late pulses occurring before response selection) through linear integration (sometimes abbreviated TSI for "time-shift invariance") to results consistent with recurrent interactions (i.e., a greater influence of early pulses).

Huk and Shadlen (2005) performed a similar experiment on temporal integration of motion information in the lateral intraparietal (LIP) area of two rhesus monkeys. They found a greater influence for late pulses and concluded, based on modeling assuming within-choice noise, that: "... the time course of the pulse effects is consistent with the hypothesis that LIP reflects approximately linear integration that stops when the accumulation reaches a bound." (p. 10443). Wong et al. (2007) simulated integration by a recurrent circuit perturbed by within-choice noise and found it displayed a: "violation of ... TSI, similar to the violation observed in the Huk and Shadlen (2005) experiment." (p. 8). Zhou et al. (2009) provide an insightful discussion of difficulties in distinguishing different types of integration using pulse paradigms.

Individual differences, and potential confounding of inferences about the nature of integration, might occur in these experiments if some participants noticed the pulse and employ compensatory strategies such as delaying the onset of sampling (to avoid being mislead by an early pulse) or prematurely terminating sampling (to avoid being mislead by a late pulse). Usher and McClelland (2001) attempted to minimize such problems using fairly brief pulses, as did Huk and Shadlen (2005) by making the pulse unrelated to rewards, but the success of these measures was not directly tested. Brown and Heathcote (2005b) attempted to avoid these problems using a very brief pulse (90 ms) that was meta-contrast masked. Masking was shown to be effective as detection of trials with a pulse was at chance levels for most participants. At the start of the experiment later arriving evidence had greater weight than earlier arriving evidence (i.e., accumulation was leaky), but as subjects practiced at the task integration quickly became linear.

One possible interpretation of why practice might promotes linear integration, is that it makes decision-making more efficient when evidence is stationary. That is, when early arriving evidence is no better guide to the correct choice than late-arriving evidence it is best to weigh each equally. In many simple-choice tasks where stimuli remain available until a choice is made, evidence is likely stationary, particularly as rapid responding means the stimulus is sampled for a relatively brief period of time. Stationarity does, however, require that sampling not begin prematurely (i.e., before stimulus information first becomes available), although Laming (1968) suggested the effects of premature sampling might be mimicked by start-point variability. When the stimulus is only available briefly evidence will be non-stationarity unless it is sampled from a mnemonic representation of the stimulus (e.g., Smith and Ratcliff, 2009). Of course, this type of non-stationary is different to the effects of non-linearity intrinsic to some evidence accumulation models, although they might mimic it.

More recently Tsetsos et al. (2011) introduced an innovative new multiple-pulse paradigm that manipulates temporal correlations among the evidence for three alternatives. In agreement with Usher and McClelland (2001), they found strong individual differences, and their analysis supported the LCA over two other within-choice noise models (a race and diffusion model). Overall, then, it appears that models that can accommodate all types of integration by adjusting the balance of leaky and recurrent dynamics best accommodate the full range of pulse-paradigm data. Such models include not only the LCA but also Wong et al.'s (2007) model, DFT (Busemeyer and Townsend, 1993), and Brown and Heathcote's (2005a) BA. Even so, models assuming linear accumulation, such as the LBA and Ratcliff diffusion, often provide a good fit to data from the simple-choice paradigms used to evaluate evidence accumulation models. A potential reason is that large numbers of responses are collected in these paradigms in order to facilitate model fitting, and so participants are afforded substantial practice. Brown and Heathcote's (2005b) results suggest that practice might cause an adjustment of recurrent interactions so they balance leakage, resulting in efficient integration that is approximately linear.

DETERMINISTIC ACCUMULATION

That the deterministic assumption can support a comprehensive account of simple-choice behavior (i.e., of benchmark findings about the choices made and the times to make them) is perhaps surprising. The nervous system is not deterministic at the level of individual cell activity and most evidence accumulation models assume a dominant role for within-choice noise, although they can only provide a comprehensive account by also assuming a role for between-choice noise. The success of the deterministic assumption maybe less surprising, however, if it is kept in mind that it is an assertion about the importance of different types of noise for explaining choice responding. Importantly, this characterization as an approximation targeting behavioral measurement entails the implication that the deterministic assumption may not be appropriate for explaining other measurements of choice (e.g., single-cell firing rates). Even so, progress in science has often relied on developing approximations appropriate for different levels of description (e.g., the diffusive micro-scale behavior of gas molecules vs. the macro-level interactions between pressure, volume, and temperature captured by the gas laws).

The mechanism by which macro-scale deterministic behavior might emerge from the micro-scale variability intrinsic to the nervous system remains to be determined. Local averaging mechanisms have been argued not to be sufficient because of correlations between the activities of nearby cells that share inputs

(Zohary et al., 1994). One possible explanation is that global rather than local brain dynamics determine behavior, and that it is these global dynamics emerging from the interaction of a variety of widely distributed areas that are deterministic. For example, Ho et al. (2009) suggested that activity consistent with evidence accumulation recorded in a variety of sensorimotor regions reflects a modality independent downstream input from right insula. Similarly, Forstmann et al. (2008) found that activity in the striatum and pre-SMA serves to implement the response boundary envisioned by evidence accumulation models. In both cases LBA model parameters provided a coherent link between observed behavior and neuroimaging measures.

Whatever the mechanism that achieves deterministic accumulation at the scale that determines behavior, it seems clear that deterministic accumulation could potentially give an organism a great adaptive advantage. Optimal methods of integrating out within-choice noise via accumulation have been of great interest on adaptive grounds (e.g., Gold and Shadlen, 2007). On these ground it might be even more desirable to remove noise intrinsic to the nervous system before accumulation (i.e., if accumulation was effectively deterministic), although in doing so there might be some trade-off with the level of between-choice noise. Regardless, given that no serious contender as a model of choice behavior can do without between-choice noise, it would seem to be desirable that optimality analyses take account of the effects of between-choice noise.

Our specific proposal that between-choice noise in the rate of linear accumulation has a Lognormal distribution has a novel implication for the conventional division between stimulus encoding and response selection stages in models of simple choice. This proposal can be derived from Ulrich and Miller's (1993) idea that a cascade or "flow" of linear accumulation that characterizes all stages of processing from sensory processing up to and including response selection. A Lognormal distribution emerges as a Central Limit Theorem approximation when the rates of units in each stage are identically and independently distributed and vary independently between-choices, as assumed by Ulrich and Miller (1993), without strong assumptions on the form of that variation. If stimulus encoding does continuously feed activation into the response selection process, estimates of residual time might not, at least in part, reflect the time for stimulus encoding as is conventionally assumed. Instead, they would largely reflect the response execution stage, with only a small contribution from an initial "dead-time" before sensory neurons begin to respond.

Our mathematical results for an LNR with a Lognormal distance distribution might appear to show that not all linear deterministic models can explain speed-accuracy trade-off by integrating out random biases due to differences in the distance from starting point to boundary between accumulators. In the LBA the tendency for differences in rates to overcome random biases can be increased by an equal increase in the parameter controlling the position of the boundary (b) for all accumulators, resulting in slower but more accurate responses. An analogous effect is not obtained by increasing the LNR parameter determining the mean of the log-distance distribution, μ_D, while holding the variance in log-distance, σ_D^2, constant. This is because and increase in μ_D increases not only the mean distance ($M_D = e^{\mu_D+\sigma_D^2/2}$) but also the variability in distance ($V_D = (e^{\sigma_D^2}-1)e^{2\mu_D+\sigma_D^2}$), and hence the magnitude of random bias. However, as pointed out by a reviewer, a change in M_D while holding V_D constant, although the effect is not exactly analogous to a change the $b$in the LBA (as it also changes the shape of the distance distribution), does result in a speed-accuracy trade-off in the LNR.

Future research could fruitfully explore not only different LNR parameterizations (such as in terms of M_D and V_D) but also a wider variety of distributional assumptions. A range of positive distributions similar to the Lognormal, such as the Gamma distribution and extreme-value distributions such as the Weibull and Gumbel (see Heathcote et al., 2004), are plausible candidates for rate distributions. As they are positive, these distributions ensure that a response will be selected on every trial. Combinations of these rate distributions with different distance distributions (notably the analytically tractable shifted uniform distribution used in the LBA) will then help to provide a better understanding of the general strengths and limitations of the linear deterministic framework. A second area for further exploration concerns correlations in parameters across accumulators. The LNR model is particularly suited to such exploration, both because an LNR with such correlations remains tractable.

Our model fits show that the LNR can provide an accurate and comprehensive account of behavioral data from a simple-choice experiment. If this finding generalizes to other data sets the LNR could provide a useful descriptive model of simple-choice data. It is suited to this role because it sacrifices little of the tractability of distributions commonly applied to simple and correct-choice RTs (e.g., the Wald and ExGaussian, see Ratcliff and Murdock, 1976; Heathcote et al., 1991; Heathcote, 2004; Matzke and Wagenmakers, 2009), and is better able to describe differences in correct and error RT than the tractable EZ and EZ2 simplifications of the RDM (Wagenmakers et al., 2007; Grasman et al., 2009). Although both LBA and LNR likelihoods are easy to compute, the LNR has one distinct advantage: it is also easy to obtain full conditional likelihoods (i.e., likelihoods based on fixing all but one parameter). This is particularly useful for Bayesian approaches as it enables efficient Gibbs sampling steps rather than the less efficient methods required for the LBA (see Donkin et al., 2009a).

Our model fitting results could have the implication that, contrary to Brown and Heathcote's (2005a, 2008) assertions, between-choice noise alone is insufficient to provide a comprehensive account of simple-choice behavior. This implication follows from the assumption that instruction manipulations cannot influence the rate of evidence accumulation. The LBA with this selective influence assumption has previously provided a comprehensive account of emphasis manipulations producing smaller effects on accuracy (e.g., Forstmann et al., 2008, 2010, 2011). In contrast, we found that this type of LBA was clearly unable to fit the changes in accuracy of up to almost 15% in Wagenmaker et al.'s (2008) data.

Although this selective influence assumption is conventional, recently there has been increasing evidence that instructions do influence drift rates, particularly in applications of the RDM. The RDM has always assumed that the relative values of the mean rates for stimuli associated with different responses are determined by a *drift criterion* parameter (Ratcliff and McKoon, 2008). Changes

in the drift criterion and evidence boundaries both affect response bias but in different ways (e.g., Criss, 2010; Starns et al., 2012), and the drift criterion can be influenced by instructions (Leite and Ratcliff, 2011). Factors affecting attentional focus have also been argued to affect drift rates. For example, in White et al.'s (2011) RDM model of the flanker task drift rates change due to changes in "attentional focus" brought about by manipulating the proportion of trials involving response conflict. Arguably speed vs. accuracy emphasis instructions might also affect attentional focus, and hence the rate of information accumulation. Finally, Kleinsorge (2001) demonstrated that, given sufficient warning, participants can, in response to instructions, mobilize an extra effort that genuinely improves performance in a way that cannot be accounted for by a speed-accuracy trade-off. Given Wagenmakers et al. (2008) manipulated speed vs. accuracy instructions between blocks of trials participants had plenty of time to focus attention and/or make an extra effort that could affect performance through changes in drift rates.

On the basis of these recent findings, and the fact that the LBA and LNR did provide an accurate and comprehensive account of Wagenmaker et al.'s (2008) data when emphasis was allowed to affect drift rates, we believe it would be premature to reject the deterministic approximation. However, further research on this point is clearly called for. Neuroscience methods that can provide fairly direct evidence about the effect of instruction manipulations on the statistical characteristics (e.g., mean and variance) of evidence extracted from a stimulus will likely be particularly useful in this regard (see Ho et al., in press). Our results, and the others just reviewed, also suggest that it may also be fruitful to revisit traditional assumptions about the effects of a variety of instruction and expectancy (e.g., for a particular response or type of sensory information) manipulation in order to determine whether they can have multiple effects. That is, can these manipulations cause changes not only evidence accumulation parameters traditionally associated with strategic factors (e.g., boundaries and systematic bias) but also on rate means and on the variability of rates and biases?

ACKNOWLEDGMENTS

This work was supported by Andrew Heathcote's Australian Research Council Professorial Fellowship and an equipment grant from the University of Newcastle's Faculty of Science and IT. We would like to acknowledge the constructive and useful advice provided by all reviewers and in particular Rani Moran for pointing out to us that an increase in mean distance alone is sufficient produce a speed-accuracy trade-off.

REFERENCES

Boucher, L., Palmeri, T. J., Logan, G. D., and Schall, J. D. (2007). Inhibitory control in mind and brain: an interactive race model of countermanding saccades. *Psychol. Rev.* 114, 376–397.

Brown, S. D., and Heathcote, A. (2005a). A ballistic model of choice response time. *Psychol. Rev.* 112, 117–128.

Brown, S. D., and Heathcote, A. (2005b). Practice increases the efficiency of evidence accumulation in perceptual choice. *J. Exp. Psychol. Hum. Percept. Perform.* 31, 289–298.

Brown, S. D., and Heathcote, A. (2008). The simplest complete model of choice reaction time: linear ballistic accumulation. *Cogn. Psychol.* 57, 153–178.

Brown, S. D., Marley, A. A. J., Donkin, C., and Heathcote, A. (2008). An integrated model of choices and response times in absolute identification. *Psychol. Rev.* 115, 396–425.

Burnham, K. P., and Anderson, D. R. (2004). Multimodel inference: Understanding AIC and BIC in model selection. *Sociol. Methods Res.* 33, 261–304.

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.

Carpenter, R. H. S. (1981). "Oculomotor procrastination," in *Eye Movements: Cognition and Visual Perception*, eds D. F. Fisher, R. A. Monty, and J. W. Senders (Hillsdale: Lawrence Erlbaum), 237–246.

Criss, A. H. (2010). Differentiation and response bias in episodic memory: evidence from reaction time distributions. *J. Exp. Psychol. Learn. Mem. Cogn.* 484–499.

Donkin, C., Averell, L., Brown, S. D., and Heathcote, A. (2009a). Getting more from accuracy and response time data: methods for fitting the linear ballistic accumulator model. *Behav. Res. Methods* 41, 1095–1110.

Donkin, C., Brown, S. D., and Heathcote, A. (2009b). The over-constraint of response time models: rethinking the scaling problem. *Psychon. Bull. Rev.* 16, 1129–1135.

Donkin, C., Brown, S. D., and Heathcote, A. (2011a). Drawing conclusions from choice response time models: a tutorial. *J. Math. Psychol.* 55, 140–151.

Donkin, C., Brown, S., Heathcote, A., and Wagenmakers, E. J. (2011b). Diffusion versus linear ballistic accumulation: different models for response time, same conclusions about psychological mechanisms? *Psychon. Bull. Rev.* 55, 140–151.

Eidels, A., Donkin, C., Brown, S. D., and Heathcote, A. (2010). Converging measures of workload capacity. *Psychon. Bull. Rev.* 17, 763–771.

Forstmann, B. U., Dutilh, G., Brown, S. D., 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., Schäfer, A., Anwander, A., Neumann, J., Brown, S. D., 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., Tittgemeyer, M., Wagenmakers, E.-J., Derrfuss, J., Imperati, D., and Brown, S. D. (2011). The speed-accuracy tradeoff in the elderly brain: a structural model-based approach. *J. Neurosci.* 31, 17242–17249.

Gao, J., KongFatt, W.-L., Holmes, P., Simen, P., and Cohen, J. D. (2009). Sequential effects in two-choice reaction time tasks: decomposition and synthesis of mechanisms. *Neural Comput.* 21, 2407–2436.

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

Grasman, R. P. P. P., Wagenmakers, E.-J., and van der Maas, H. L. J. (2009). On the mean and variance of response times under the diffusion model with an application to parameter estimation. *J. Math. Psychol.* 53, 55–68.

Gratton, G., Coles, M. G. H., Sirevaag, E. J., Eriksen, C. W., and Donchin, E. (1988). Pre and post stimulus activation of response channels: a psychophysiological analysis. *J. Exp. Psychol. Hum. Percept. Perform.* 14, 331–344.

Heathcote, A. (2004). Fitting the wald and ex-wald distributions to response time data. *Behav. Res. Methods Instrum. Comput.* 36, 678–694.

Heathcote, A., and Brown, S. D. (2004). Reply to speckman and rouder: a theoretical basis for QML. *Psychon. Bull. Rev.* 11, 577.

Heathcote, A., Brown, S. D., and Cousineau, D. (2004). QMPE: estimating Lognormal, Wald and Weibull RT distributions with a parameter dependent lower bound. *Behav. Res. Methods Instrum. Comput.* 36, 277–290.

Heathcote, A., and Hayes, B. (2012). Diffusion versus linear ballistic accumulation: different models for response time with different conclusions about psychological mechanisms? *Can. J. Exp. Psychol.* 66, 125–136.

Heathcote, A., Popiel, S. J., and Mewhort, D. J. K. (1991). Analysis of response-time distributions: an example using the Stroop task. *Psychol. Bull.* 109, 340–347.

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.

Ho, T. C., Brown, S. D., Abuyo, N. A., Kul, E.-H. J., and Serences, J. T. (in press). Perceptual consequences of feature-based attentional enhancement and suppression. *J. Vis.*

Huettel, S. A., Mack, P. B., and McCarthy, G. (2002). Perceiving patterns in random series: dynamic processing of sequence in prefrontal cortex. *Nat. Neurosci.* 5, 485–490.

Huk, A. C., and Shadlen, M. N. (2005). Neural activity in macaque parietal cortex reflects temporal integration of visual motion signals during perceptual decision making. *J. Neurosci.* 25, 10420–10436.

Karayanidis, F., Mansfield, E. L., Galloway, K. L., Smith, J. L., Provost, A., and Heathcote, A. (2009). Anticipatory reconfiguration elicited by fully and partially informative cues that validly predict a switch in task. *Cogn. Affect. Behav. Neurosci.* 9, 202–215.

Kleinsorge, T. (2001). The time course of effort mobilization and strategic adjustment of response criteria. *Psychol. Res.* 65, 216–223.

Laming, D. R. J. (1968). *Information Theory of Choice-Reaction Times.* London: Academic Press.

Leite, F. P., and Ratcliff, R. (2010). Modeling reaction time and accuracy of multiple-alternative decisions. *Atten. Percept. Psychophys.* 72, 246–273.

Leite, F. P., and Ratcliff, R. (2011). What cognitive processes drive response biases? A diffusion model analysis. *Judgm. Decis. Mak.* 7, 651–687.

Logan, G. D., and Cowan, W. B. (1984). On the ability to inhibit thought and action: a theory of an act of control. *Psychol. Rev.* 91, 295–327.

Luce, R. D. (1986). *Response Times: Their Role in Inferring Elementary Mental Organization.* New York: Oxford University Press.

Marley, A. A. J., and Colonius, H. (1992). The "horse race" random utility model for choice probabilities and reaction times, and its competing risks interpretation. *J. Math. Psychol.* 36, 1–20.

Matzke, D., and Wagenmakers, E.-J. (2009). Psychological interpretation of ex-Gaussian and shifted Wald parameters: a diffusion model analysis. *Psychon. Bull. Rev.* 16, 798–817.

McClelland, J. L. (1979). On the time relations of mental processes: an examination of systems of processes in cascade. *Psychol. Rev.* 86, 287–330.

Morey, R. D. (2008). Confidence intervals from normalized data: a correction to Cousineau (2005). *Tutor. Quant. Methods Psychol.* 4, 61–64.

Myung, I. J., and Pitt, M. A. (1997). Applying Occam's razor in modeling cognition: a Bayesian approach. *Psychon. Bull. Rev.* 4, 79–95.

R Development Core Team. (2012). *R: A Language and Environment for Statistical Computing.* Vienna: R Foundation for Statistical Computing.

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

Ratcliff, R. (2001). Putting noise into neurophysiological models of simple decision making (letter). *Nat. Neurosci.* 4, 336.

Ratcliff, R., Gomez, P., and McKoon, G. (2004). A diffusion model account of the lexical decision task. *Psychol. Rev.* 111, 159–182.

Ratcliff, R., and McKoon, G. (2008). The diffusion decision model: theory and data for two choice decision tasks. *Neural Comput.* 20, 873–922.

Ratcliff, R., and Murdock, B. B. Jr. (1976). Retrieval processes in recognition memory. *Psychol. Rev.* 83, 190–214.

Ratcliff, R., and Rouder, J. N. (1998). Modeling response times for two choice decisions. *Psychol. Sci.* 9, 347–356.

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

Ratcliff, R., and Tuerlinckx, F. (2002). Estimating parameters of the diffusion model: approaches to dealing with contaminant reaction times and parameter variability. *Psychon. Bull. Rev.* 9, 438–481.

Rinkenauer, G., Osman, A., Ulrich, R., Muller-Gethmann, H., and Mattes, S. (2004). On the locus of speed-accuracy trade-off in reaction time: inferences from the lateralized readiness potential. *J. Exp. Psychol. Gen.* 133, 261–282.

Schweikert, R. (1989). Separable effects of factors on activation function in discreet and continuous models. *Psychol. Bull.* 106, 318–328.

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.

Smith, P. L. (1995). Psychophysically principled models of visual simple reaction time. *Psychol. Rev.* 102, 567–591.

Smith, P. L., and Ratcliff, R. (2009). An integrated theory of attention and decision making in visual signal detection. *Psychol. Rev.* 116, 283–317.

Speckman, P. L., and Rouder, J. N. (2004). A comment on Heathcote, Brown, and Mewhort's QMLE method for response time distributions. *Psychon. Bull. Rev.* 11, 574–576.

Starns, J. J., Ratliff, R., and White, C. N. (2012). Diffusion model drift rates can be influenced by decision processes: an analysis of the strength-based mirror effect. *J. Exp. Psychol. Learn. Mem. Cogn.* PMID: 22545609. [Epub ahead of print].

Stone, M. (1960). Models for choice-reaction time. *Psychometrika* 25, 251–260.

Townsend, J. T., and Fikes, T. (1995). *A Beginning Quantitative Taxonomy of Cognitive Activation Systems and Applications to Continuous Flow Processes.* Technical Report, 131. Cognitive Science Program, Bloomington: Indiana University.

Tsetsos, K., Usher, M., and McClelland, J. L. (2011). Testing multi-alternative decision models with non-stationary evidence. *Front. Neurosci.* 5:63. doi:10.3389/fnins.2011.00063

Ulrich, R., and Miller, J. (1993). Information processing models generating lognormally distributed reaction times. *J. Math. Psychol.* 37, 513–525.

Usher, M., and McClelland, J. L. (2001). On 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. C., Serences, J. T., and Forstmann, B. U. (2011). Neural correlates of trial-to-trial fluctuations in response caution. *J. Neurosci.* 31, 17488–17495.

Vickers, D. (1979). *Decision Processes in Visual Perception.* New York, NY: Academic Press.

Vrieze, S. I. (2012). Model selection and psychological theory: a discussion of the differences between the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). *Psychol. Methods* 17, 228–243.

Wagenmakers, E.-J., Ratcliff, R., Gomez, P., and McKoon, G. (2008). A diffusion model account of criterion shifts in the lexical decision task. *J. Mem. Lang.* 58, 140–159.

Wagenmakers, E.-J., van der Maas, H. L. J., and Grasman, R. P. P. P. (2007). An EZ-diffusion model for response time and accuracy. *Psychon. Bull. Rev.* 14, 3–22.

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.

Wong, K. F., Huk, A. C., Shadlen, M. N., and Wang, X. J. (2007). Neural circuit dynamics underlying accumulation of time-varying evidence during decision-making. *Front. Comput. Neurosci.* 1:6. doi:10.3389/neuro.10.006.2007

Woodworth, R. S., and Schlosberg, H. (1954). *Experimental Psychology.* New York: Holt.

Zhou, X., Wong-Lin, K., and Philip, H. (2009). Time-varying perturbations can distinguish among integrate-to-threshold models for perceptual decision making in reaction time tasks. *Neural Comput.* 21, 2336–2362.

Zohary, E., Shadlen, M. N., and Newsome, W. T. (1994). Correlated neuronal discharge rate and its implications for psychophysical performance. *Nature* 370, 140–143.

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: 22 February 2012; accepted: 26 July 2012; published online: 23 August 2012.

*Citation: Heathcote A and Love J (2012) Linear deterministic accumulator models of simple choice. Front. Psychology **3**:292. doi: 10.3389/fpsyg.2012.00292*

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

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 29 May 2012
doi: 10.3389/fpsyg.2012.00132

Do the dynamics of prior information depend on task context? An analysis of optimal performance and an empirical test

Don van Ravenzwaaij[1]*, Martijn J. Mulder[2], Francis Tuerlinckx[3] and Eric-Jan Wagenmakers[1,2]

[1] Psychological Methods, University of Amsterdam, Amsterdam, Netherlands
[2] Cognitive Science Center Amsterdam, Amsterdam, Netherlands
[3] Katholieke Universiteit Leuven, Leuven, Belgium

Edited by:
Marius Usher, Tel-Aviv University, Israel

Reviewed by:
Philip Holmes, Princeton University, USA
Juan Gao, Stanford University, USA

***Correspondence:**
Don van Ravenzwaaij, University of New South Wales, Mathews Building, Room 702, Sydney 2052 Australia.
e-mail: d.vanravenzwaaij@unsw.edu.au

In speeded two-choice tasks, optimal performance is prescribed by the drift diffusion model. In this model, prior information or advance knowledge about the correct response can manifest itself as a shift in starting point or as a shift in drift rate criterion. These two mechanisms lead to qualitatively different choice behavior. Analyses of optimal performance (i.e., Bogacz et al., 2006; Hanks et al., 2011) have suggested that bias should manifest itself in starting point when difficulty is fixed over trials, whereas bias should (additionally) manifest itself in drift rate criterion when difficulty is variable over trials. In this article, we challenge the claim that a shift in drift criterion is necessary to perform optimally in a biased decision environment with variable stimulus difficulty. This paper consists of two parts. Firstly, we demonstrate that optimal behavior for biased decision problems is prescribed by a shift in starting point, irrespective of variability in stimulus difficulty. Secondly, we present empirical data which show that decision makers do not adopt different strategies when dealing with bias in conditions of fixed or variable across-trial stimulus difficulty. We also perform a test of specific influence for drift rate variability.

Keywords: drift diffusion model, decision making, bias

INTRODUCTION

In real-life decision making, people often have *a priori* preferences for and against certain choice alternatives. For instance, some people may prefer Audi to Mercedes, Mac to PC, or Gillette to Wilkinson, even before seeing the product specification. For product preferences, people are influenced by prior experiences and advertising. Here we study the effects of prior information in the context of perceptual decision making, where participants have to decide quickly whether a cloud of dots is moving to the left or to the right. Crucially, participants are given advance information about the likely direction of the dots. In the experiment reported below, for instance, participants were sometimes told that 80% of the stimuli will be moving to the right. How does this advance information influence decision making?

In general, advance information that favors one choice alternative over the other biases the decision process: people will prefer the choice alternative that has a higher prior probability of being correct. This bias usually manifests itself as a shorter response time (RT) and a higher proportion correct when compared to a decision process with equal prior probabilities. Because bias expresses itself in two dependent variables simultaneously (i.e., RT and proportion correct), and because people only have control over a specific subset of the decision environment (e.g., the participant cannot control task difficulty) an analysis of optimal adjustments is traditionally carried out in the context of a *sequential sampling model*. Prototypical sequential sampling models such as the Sequential Probability Ratio Test (SPRT; e.g., Wald and Wolfowitz, 1948; Laming, 1968) or the drift diffusion model (DDM; Ratcliff, 1978) are based on the assumption that the decision maker gradually accumulates noisy information until an evidence threshold is reached.

In such sequential sampling models, the biasing influence of prior knowledge can manifest itself in two ways (e.g., Diederich and Busemeyer, 2006; Ratcliff and McKoon, 2008; Mulder et al., 2012). The first manifestation, which we call *prior bias*, is that a decision maker decides in advance of the information accumulation process to lower the evidence threshold for the biased alternative. The second manifestation, which we call *dynamic bias* (cf. Hanks et al., 2011), is that a decision maker weighs more heavily the evidence accumulated in favor of the biased alternative. There is both theoretical and empirical evidence that shows that both types of bias manifest itself in different situations.

The situation that has been studied most often is one in which task difficulty is fixed across trials. For this case, Edwards (1965) has shown that optimal performance can be achieved by prior bias alone. In empirical support for this theoretical analysis, Bogacz et al. (2006) demonstrated that in three experiments the performance of participants approximated the optimality criterion from Edwards (1965). In a recent paper by Gao et al. (2011), it was demonstrated that for fixed stimulus difficulty, but varied response deadlines over trials, behavioral data was best described by the implementation of prior bias.

Another situation is one in which task difficulty varies across trials. For this case, Hanks et al. (2011) reasoned that people should

tend ever more strongly toward the biased response option as the information accumulation process continues (see also Yang et al., 2005; Bogacz et al., 2006, p. 730). The intuitive argument is that a lengthy decision process indicates that the decision is difficult (i.e., the stimulus does not possess much diagnostic information), and this makes it adaptive to attach more importance to the advance information. In the extreme case, a particular stimulus can be so difficult that it is better to simply go with the advance information and guess that the biased response option is the correct answer.

Thus, in an environment of constant difficulty an optimal decision maker accommodates advance information by prior bias alone, such that a choice for the likely choice alternative requires less evidence than that for the unlikely choice alternative. The conjecture of Hanks et al. (2011) is that, in an environment of variable difficulty, an optimal decision maker accommodates advance information not just by prior bias, but also by dynamic bias.

This paper has two main goals. The first goal is to extend the analytical work of Edwards (1965) and show that prior bias accounts for optimal performance regardless of whether stimulus difficulty is fixed or variable across trials. The second goal is to examine empirically which kind of bias decision makers implement when stimulus difficulty is fixed or variable.

The organization of the paper is as follows. In the first section we briefly introduce the drift diffusion model (DDM), the prototypical sequential sampling model that can be used to model both prior bias and dynamic bias (Bogacz et al., 2006; van Ravenzwaaij et al., 2012). Next we present analytical work that shows how optimal performance in biased decision environments may be achieved with prior bias alone, regardless of whether stimulus difficulty is fixed or variable. We then present an empirical study showing how people accommodate bias for fixed and variable stimulus difficulty in similar fashion.

THE DRIFT DIFFUSION MODEL (DDM)

In the DDM (Ratcliff, 1978; Ratcliff and Rouder, 2000; Wagenmakers, 2009; van Ravenzwaaij et al., 2012), a decision process with two response alternatives is conceptualized as the accumulation of noisy evidence over time. Evidence is represented by a single accumulator, so that evidence in favor of one alternative is evidence against the other alternative. A response is initiated when the accumulated evidence reaches one of two predefined thresholds. For instance, in a *lexical decision task*, participants have to decide whether a letter string is an English word, such as TANGO, or a non-word, such as TANAG (**Figure 1**).

The model assumes that the decision process commences at the starting point z, from which point evidence is accumulated with a signal-to-noise ratio that is governed by mean drift rate v and Wiener noise. Without trial-to-trial variability in drift rate, the change in evidence x is described by the following stochastic differential equation

$$dx(t) = v \cdot dt + s \cdot dW(t), \quad (1)$$

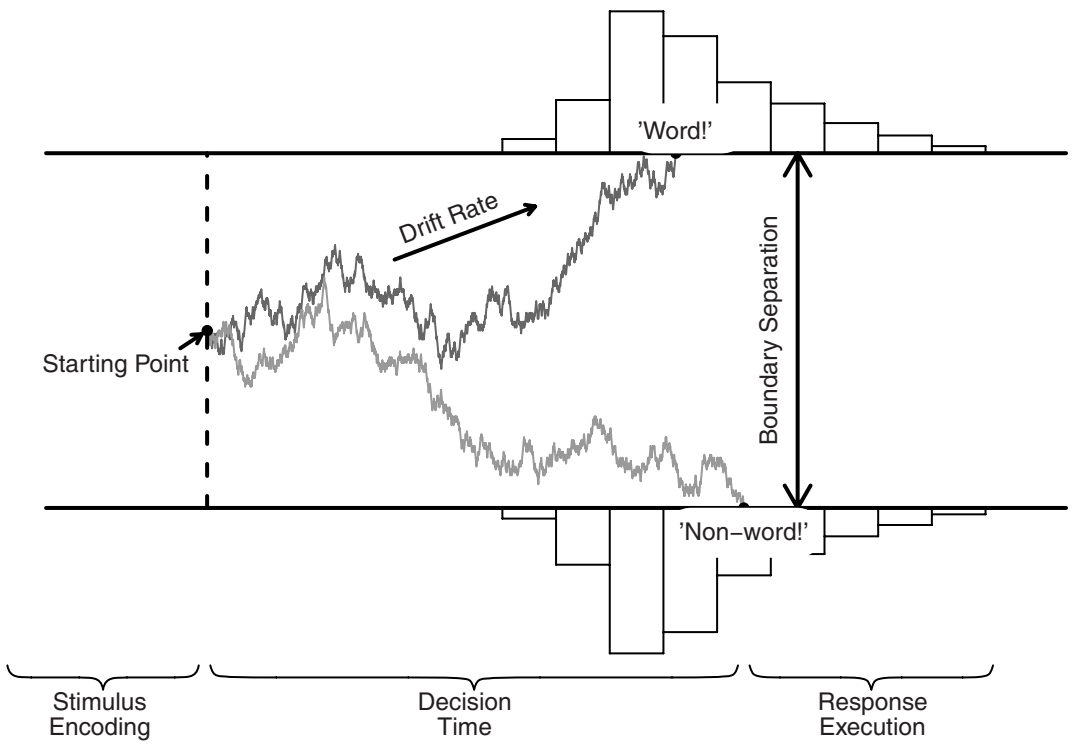


FIGURE 1 | The DDM and its key parameters, illustrated for a lexical decision task. Evidence accumulation begins at starting point z, proceeds over time guided by mean drift rate v, but subject to random noise, and stops when either the upper or the lower boundary is reached. Boundary separation a quantifies response caution. The predicted RT equals the accumulation time plus the time required for non-decision processes T_{er} (i.e., stimulus encoding and response execution).

where W represents the Wiener noise process (i.e., idealized Brownian motion). Parameter s represents the standard deviation of $dW(t)$[1]. Values of v near zero produce long RTs and high error rates. Trial-to-trial variability in drift rate is quantified by η.

Evidence accumulation stops and a decision is initiated once the evidence accumulator hits one of two response boundaries. The difference between these boundaries, boundary separation a, determines the speed–accuracy trade-off; lowering a leads to faster RTs at the cost of a higher error rate. When the starting point, z, is set at $a/2$, bias in the decision process is not manifested in the starting point. Together, these parameters generate a distribution of decision times (DTs). The observed RT, however, also consists of stimulus-non-specific components such as response preparation and motor execution, which together make up non-decision time T_{er}. The model assumes that T_{er} simply shifts the distribution of DT, such that $RT = DT + T_{er}$ (Luce, 1986).

Thus, the five key parameters of the DDM are (1 and 2) speed of information processing, quantified by mean drift rate v and standard deviation of drift rate η; (3) response caution, quantified by boundary separation a; (4) evidence criterion, quantified by starting point z; and (5) non-decision time, quantified by T_{er}. In addition to these five parameters, the full DDM also includes parameters that specify across-trial variability in starting point, and non-decision time (Ratcliff and Tuerlinckx, 2002).

BIAS IN THE DDM

Recall from the introduction that decision makers may implement bias in two ways. A decision maker may decide prior to the start of the decision process that less evidence is required for a response in favor of the biased alternative than for the non-biased alternative. This type of bias, which we call prior bias, is manifested in the DDM as a shift in starting point (see the top panel of **Figure 2**, see also Ratcliff, 1985; Ratcliff and McKoon, 2008; Mulder et al., 2012). Prior bias is most pronounced at the onset of the decision process, but dissipates over time due to the effects of the diffusion noise s. Edwards (1965) showed that when across-trial stimulus difficulty is fixed, it is optimal to shift the starting point an amount proportional to the odds of the prior probabilities of each response alternative.

Alternatively, a decision maker may weigh evidence in favor of the biased response alternative more heavily than evidence in favor of the non-biased response alternative. This type of bias, which we call dynamic bias, is manifested in the DDM as a shift in drift rate criterion (see the bottom panel of **Figure 2**, see also Ratcliff, 1985; Ratcliff and McKoon, 2008; Mulder et al., 2012). With a shift in the drift rate criterion, drift rate for the likely choice alternative is enhanced by a bias component, such that the cumulative effect of dynamic bias grows stronger over time (compare the difference between the biased and neutral lines for shift in z and shift in v_c).

[1]Parameter s is a scaling parameter that is usually fixed. For the remainder of this paper, we set it to 0.1.

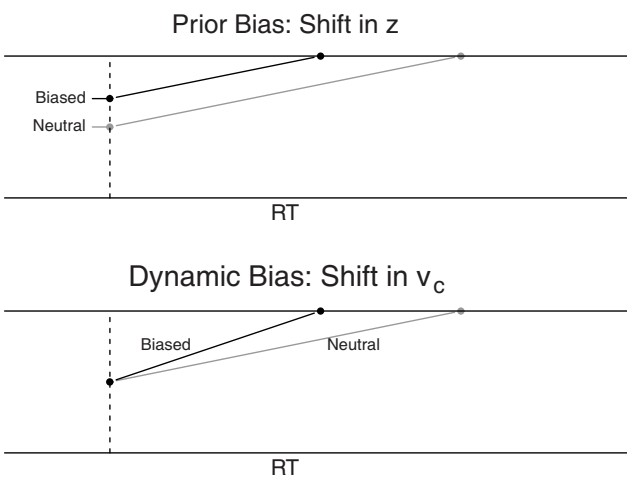


FIGURE 2 | Schematic representation of bias due to a shift in starting point z (top panel) or a shift in drift rate criterion v_c (bottom panel). The gray lines represent neutral stimuli for comparison.

BIAS IN THEORY

Below, we examine analytically which DDM parameter shifts in starting point and drift rate criterion produce optimal performance. We differentiate between fixed and variable difficulty. Two criteria of optimality are discussed: highest mean proportion correct when there is a fixed response deadline (i.e., the interrogation paradigm) and lowest mean RT (MRT) for a fixed mean proportion correct (see also Bogacz et al., 2006). In the next section, we discuss highest mean proportion correct for the interrogation paradigm.

OPTIMALITY ANALYSIS I: THE INTERROGATION PARADIGM

In the interrogation paradigm, participants are presented with a stimulus for a fixed period of time. Once the response deadline T is reached, participants are required to immediately make a response (see **Figure 3**). Thus, for the interrogation paradigm, there are no response boundaries. As such, the unbiased starting point z is 0. In this section, we will look at optimal DDM parameter settings for a biased decision in the interrogation paradigm. The performance criterion is the mean proportion correct. First, we discuss fixed stimulus difficulty across trials, or $\eta = 0$. Second, we discuss variable stimulus difficulty across trials, or $\eta > 0$.

Fixed difficulty

In order to find the maximum mean proportion correct for the interrogation paradigm, we assume that participants base their response depending on whether the evidence accumulator is above or below zero when the accumulation process is interrupted (see, e.g., van Ravenzwaaij et al., 2011; **Figure 9**). We choose parameter settings such that if the accumulator is above zero at time T, the biased response is given, whereas if the accumulator is below zero at time T, the non-biased response is given.

Across trials, the final point of the evidence accumulator will be normally distributed with mean $vT + z$ and standard deviation $s\sqrt{T}$ (e.g., Bogacz et al., 2006; van Ravenzwaaij et al., 2011). In what follows, we assume that stimuli corresponding to either

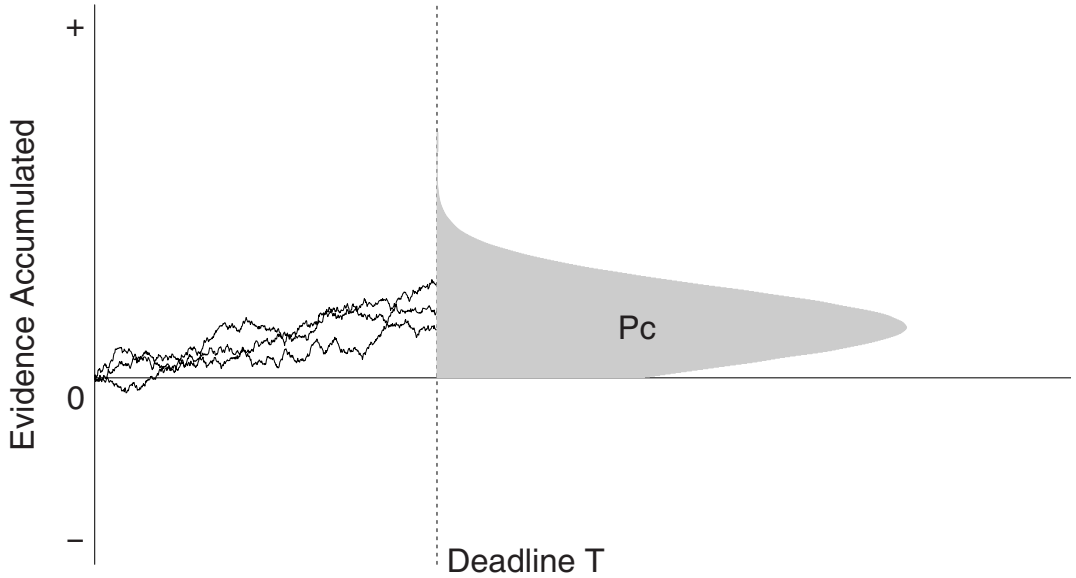


FIGURE 3 | The interrogation paradigm. At deadline *T*, decision makers choose a response alternative depending on the sign of the evidence accumulator. The shaded area under the distribution represents the proportion of correct answers.

of the two response alternatives are equally difficult[2]. For fixed difficulty, the analytical expression for mean proportion correct for an unbiased decision in the interrogation paradigm, denoted by $PcI_{F,U}$, is given by

$$PcI_{F,U} = \Phi\left[\frac{vT + z}{s\sqrt{T}}\right], \tag{2}$$

where Φ denotes the standard normal cumulative distribution[3]. In $PcI_{F,U}$, the I denotes "Interrogation," the F denotes "Fixed Difficulty," and the U denotes "Unbiased." For a biased decision, this expression becomes:

$$PcI_{F,B} = \beta\Phi\left[\frac{(v + v_c)\,T + z}{s\sqrt{T}}\right] + (1 - \beta)\,\Phi\left[\frac{(v - v_c)\,T - z}{s\sqrt{T}}\right]. \tag{3}$$

where β denotes the proportion of stimuli that are consistent with the prior information (i.e., in the experiment below, $\beta = 0.80$) and v_c denotes the shift in drift rate criterion. In $PcI_{F,B}$, the I denotes "Interrogation," the F denotes "Fixed Difficulty," and the B denotes "Biased." Equation (3) is derived from equation (2) as follows. To incorporate a shift in starting point z toward the biased alternative, the left-hand side of equation (3) contains z (cf. equation (2)). In the right-hand side, z is replaced by $-z$ to account for the fact that the correct answer lies at the non-biased threshold. To incorporate a shift in drift rate criterion, the left-hand side of equation (3) adds a shift v_c to mean drift rate v. In the right-hand side, the shift v_c is subtracted from mean drift rate v.

Maxima for $PcI_{F,B}$ occur for:

$$z_{max} + v_c \cdot T = \frac{s^2 \log\left(\frac{\beta}{1-\beta}\right)}{2v}, \tag{4}$$

where z_{max} denotes the value of the starting point that leads to the highest $PcI_{F,B}$ (e.g., Edwards, 1965; Bogacz et al., 2006, for expressions without $v_c \cdot T$).

The addition of $v_c \cdot T$ to the left-hand side of equation (4) follows from inspection of both numerators in equation (3): it shows a trade-off between the starting point z and the shift in drift rate criterion v_c, such that $\Delta z = \Delta v_c \cdot T$, where Δ denotes a parameter shift. Therefore, the maximum value of $PcI_{F,B}$ does not belong to a unique set of parameters, but exists along an infinite combination of values for z and v_c. This trade-off is graphically displayed for different sets of parameter values in **Figure 4**.

Figure 4 shows that for a drift rate v of 0.2 or 0.3, combined with a deadline T of 300 or 500 ms, every maximum that occurs for a particular value of z with $v_c = 0$ may also be reached for different combinations of z and v_c (bias β is set to 0.8). Results for different sets of parameters look qualitatively similar.

Taken at face value, this result challenges the claim that optimal behavior in biased decision problems is *exclusively* accomplished by a shift in starting point z: the same level of accuracy may be accomplished by, for instance, setting $z = 0$ and $v_c = z_{max}/T$ or any other combination of values that is consistent with the parameter trade-off $\Delta z = \Delta v_c \cdot T$. Importantly, however, participants have to be aware of the exact moment of the deadline T to be able to utilize the trade-off between starting point and

[2]This is commonly the case in perceptual decision tasks, such as the random dot motion task (Newsome et al., 1989), which we will be using in the experiment reported below.

[3]Note that for an unbiased decision, $z = 0$.

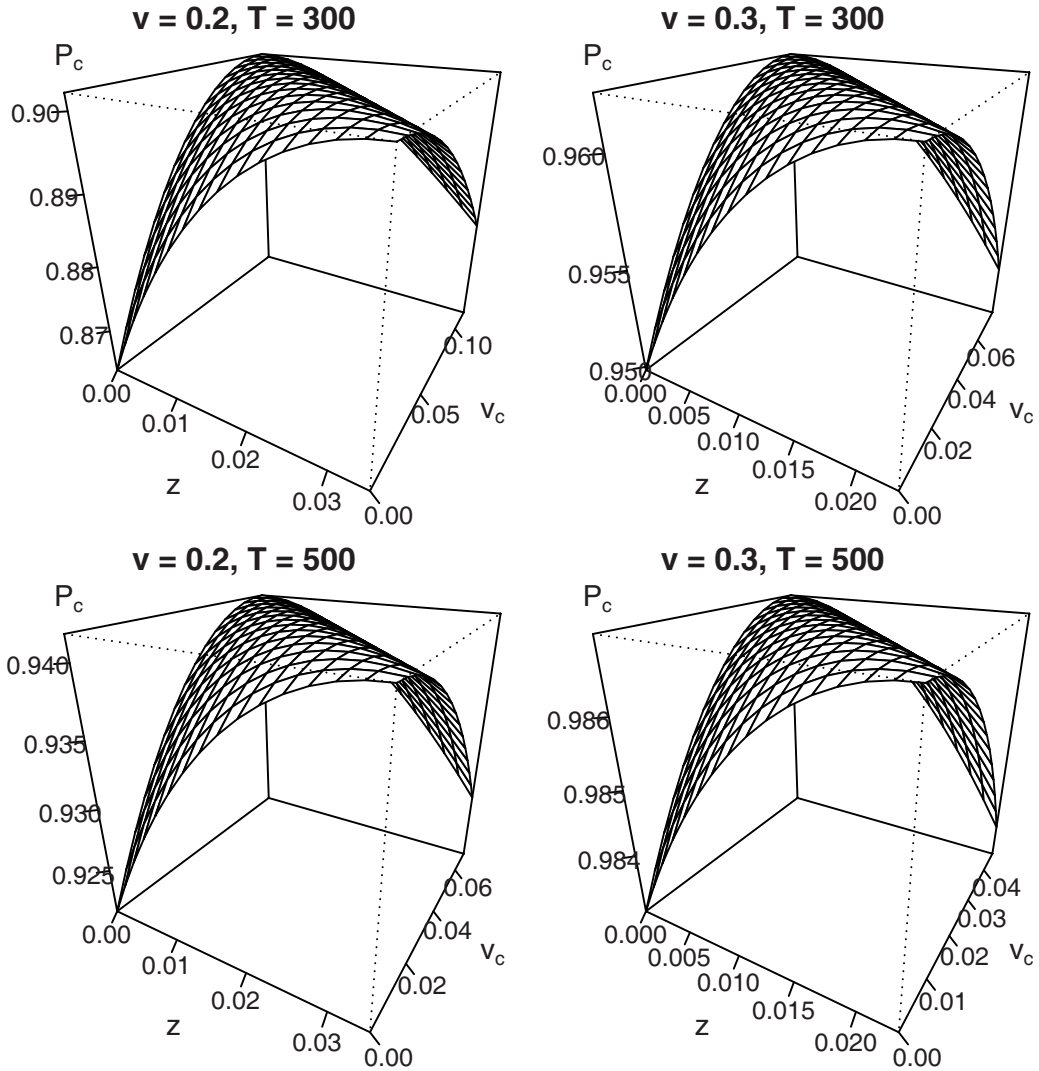


FIGURE 4 | Mean percentage correct for fixed difficulty ($\eta = 0$) in the interrogation paradigm for different combinations of starting point z and shift in mean drift rate v_c. Due to the trade-off $\Delta z = \Delta v_c \cdot T$, no unique maximum exists.

drift rate criterion. Recent work by Gao et al. (2011) examined the type of bias people implement when the deadline T is varied across blocks of trials, so that exact knowledge of the deadline is absent. The authors used the leaky competing accumulator model (LCA), a model akin to the DDM with separate accumulators for each response alternative. Using the LCA, the authors were able to differentiate between dynamic bias (shift in the input of the biased accumulator), and two accounts of prior bias (shift in starting point of the biased accumulator and shift in response threshold of the biased accumulator). The results showed that for varying deadline T, people implement a shift in starting point for the biased accumulator. This is indicative of prior bias.

Next, we examine optimal decision making in the interrogation paradigm when participants have no *a priori* knowledge of the difficulty of the decision problem. In other words, stimulus difficulty varies across trials.

Variable difficulty

To obtain an expression for the mean proportion correct in the interrogation paradigm when there is across-trial variability in stimulus difficulty, it is necessary to include across-trial variability in drift rate, or η. The mean proportion correct for this situation, or $PcI_{V,B}$, is derived by multiplying equation (3) by a Gaussian distribution of drift rates with mean v and standard deviation η. The resulting expression needs to be integrated over the interval $(-\infty, \infty)$ with respect to drift rate ξ:

$$PcI_{V,B} = \int_{-\infty}^{\infty} \left(\beta \Phi \left[\frac{(\xi + v_c)\, T + z}{s\sqrt{T}} \right] + (1 - \beta)\, \Phi \right. \\ \left. \times \left[\frac{(\xi - v_c)\, T - z}{s\sqrt{T}} \right] \right) \frac{1}{\sqrt{2\pi\eta^2}} \exp\left(-\frac{(\xi - v)^2}{2\eta^2} \right) d\xi, \quad (5)$$

which simplifies to

$$PcI_{V,B} = \beta\Phi\left[\frac{(v+v_c)\,T+z}{s\sqrt{T+T^2\eta^2/s^2}}\right] + (1-\beta)\,\Phi\left[\frac{(v-v_c)\,T-z}{s\sqrt{T+T^2\eta^2/s^2}}\right]. \quad (6)$$

In $PcI_{V,B}$, the I denotes "Interrogation," the V denotes "Variable Difficulty," and the B denotes "Biased." The derivation can be found in the appendix.

By differentiating with respect to z in a way analogous to the derivation of equation (4) from equation (3), maxima for $PcI_{V,B}$ occur for:

$$z_{max} + v_c \cdot T = \frac{\left(s^2+\eta^2 T\right)\log\left(\frac{\beta}{1-\beta}\right)}{2v}, \quad (7)$$

where z_{max} denotes the value of the starting point that leads to the highest $PcI_{V,B}$.

Once again, the addition of $v_c{\cdot}T$ to the left-hand side of equation (7) follows from inspection of both numerators in equation (6): it shows a trade-off between the starting point z and the shift in drift rate criterion v_c, such that $\Delta z = \Delta v_c{\cdot}T$, where Δ denotes a parameter shift. **Figure 5** graphically displays the parameter trade-off for the same sets of parameter values that were used in **Figure 4** (η is set to 0.1). Results for different sets of parameters look qualitatively similar.

In sum, our derivations and figures show that in the interrogation paradigm with variable across-trial stimulus difficulty, optimal decisions may again be reached by shifting either starting point z, drift rate criterion v_c, or a combination of the two. For an Ornstein–Uhlenbeck process, the fact that both types of bias cannot be distinguished in the interrogation paradigm for both fixed and variable stimulus difficulty had already been demonstrated (Feng et al., 2009, see also Rorie et al., 2010). Contrary to the fixed difficulty situation, however, participants need to know the value

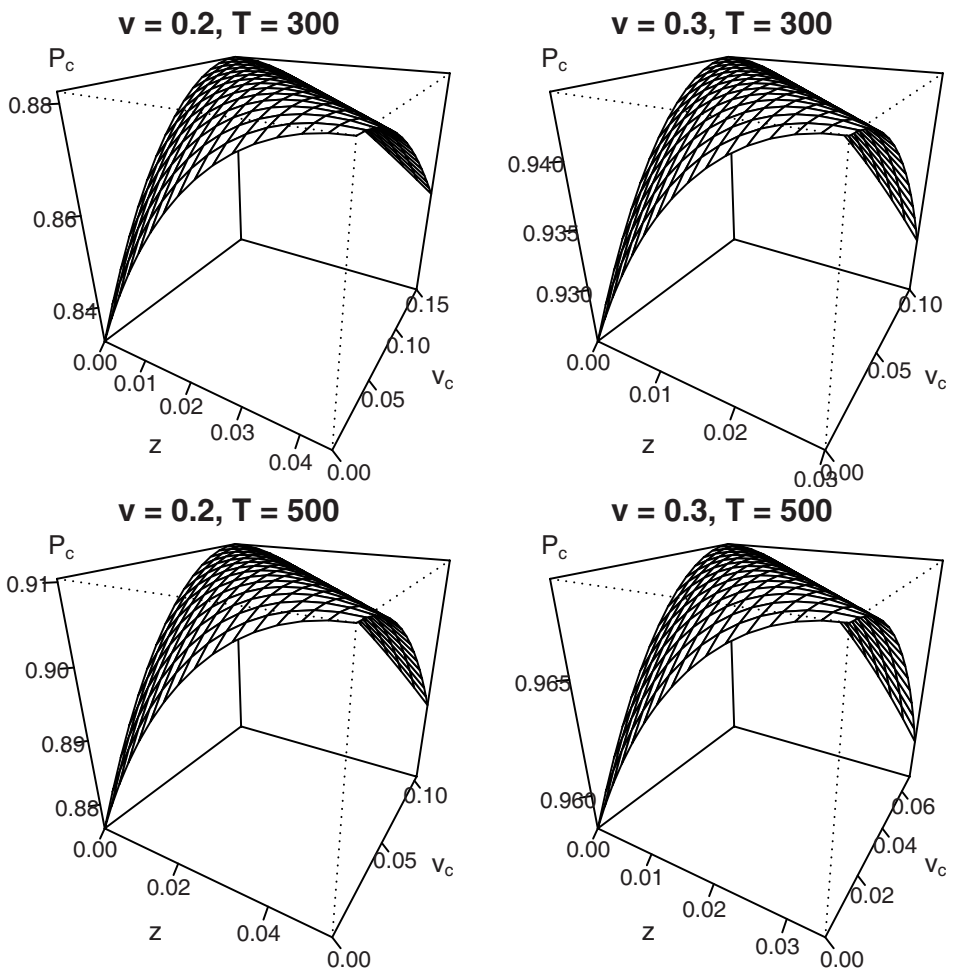


FIGURE 5 | Mean percentage correct for variable difficulty ($\eta = 0.1$) in the interrogation paradigm for different combinations of starting point z and shift in mean drift rate v_c. Due to the trade-off $\Delta z = \Delta v_c{\cdot}T$, no unique maximum exists.

of deadline T to optimize performance even when just shifting starting point.

As noted in the introduction, Hanks et al. (2011) suggested that bias should manifest itself as shifts in both starting point and the drift rate criterion when stimulus difficulty varies over trials. The authors reasoned that decision makers should decide in favor of the biased alternative when the decision process is lengthy, because slow decisions are likely to be difficult decisions. However, in the interrogation paradigm, all decisions take equally long. In order to more thoroughly investigate the claim of Hanks et al. (2011), it is necessary to eliminate the deadline T and examine a different criterion of optimality: minimum MRT for a fixed percentage correct.

OPTIMALITY ANALYSIS II: MINIMUM MRT FOR FIXED ACCURACY

In this section we consider the minimum mean RT for fixed accuracy in a free response paradigm. In the free response paradigm, a response is made once the evidence accumulator hits an upper boundary a or a lower boundary 0; the unbiased starting point z is $a/2$. First, we discuss the situation in which across-trial stimulus difficulty is fixed.

Fixed difficulty

In order to find the minimum MRT for a given level of accuracy, we need expressions for both MRT and accuracy in the DDM for a biased decision. We can then calculate combinations of starting point z and drift criterion shift v_c that yield a given level of accuracy for a given set of parameter values for drift rate v, boundary separation a, and diffusion noise s. The last step is to calculate the MRT for each combination of starting point z and drift criterion shift v_c.

For fixed difficulty, the analytical expression for mean proportion correct for an unbiased decision without a deadline, denoted by $Pc_{F,U}$, is given by

$$Pc_{F,U} = \frac{\exp\left(\frac{2av}{s^2}\right) - \exp\left(\frac{2(a-z)v}{s^2}\right)}{\exp\left(\frac{2av}{s^2}\right) - 1}, \tag{8}$$

(see, e.g., Wagenmakers et al., 2007, equation (2))[4]. In $Pc_{F,U}$, the F denotes "Fixed Difficulty," and the U denotes "Unbiased." Transforming equation (8) to an expression for a biased decision occurs in a similar fashion as the derivation of equation (3) from equation (2):

$$Pc_{F,B} = \beta Pc+_{F,B} + (1-\beta)\, Pc-_{F,B}, \tag{9}$$

where

$$\begin{aligned} Pc+_{F,B} &= \left[\frac{\exp\left(\frac{2a(v+v_c)}{s^2}\right) - \exp\left(\frac{2(a-z)(v+v_c)}{s^2}\right)}{\exp\left(\frac{2a(v+v_c)}{s^2}\right) - 1}\right] \\ Pc-_{F,B} &= \left[\frac{\exp\left(\frac{2a(v-v_c)}{s^2}\right) - \exp\left(\frac{2z(v-v_c)}{s^2}\right)}{\exp\left(\frac{2a(v-v_c)}{s^2}\right) - 1}\right]. \end{aligned} \tag{10}$$

In $Pc_{F,B}$, the B denotes "Biased."

[4]Note that for an unbiased decision, $z = a/2$.

To incorporate a shift in starting point z toward the biased alternative, $Pc+_{F,B}$ contains z as in equation (8). For $Pc-_{F,B}$, z is replaced by $a - z$ to account for the fact that the correct answer lies at the non-biased threshold. To incorporate a shift in drift rate criterion, $Pc+_{F,B}$ adds a shift v_c to mean drift rate v. For $Pc-_{F,B}$, the shift v_c is subtracted from mean drift rate v.

Now that we have an expression for mean proportion correct for a biased decision without a response deadline, we need an expression for MRT. For an unbiased decision without deadline, MRT is given by

$$MRT_U = -\frac{z}{v} + \frac{a\left[\exp\left(-2vz/s^2\right) - 1\right]}{v\left[\exp\left(-2va/s^2\right) - 1\right]}, \tag{11}$$

(e.g., Grasman et al., 2009, equation (5)). For a biased decision, this expression becomes

$$\begin{aligned} MRT_B = {} & \beta\left(-\frac{z}{v+vc} + \frac{a\left[\exp\left(-2(v+vc)z/s^2\right) - 1\right]}{(v+vc)\left[\exp\left(-2(v+vc)a/s^2\right) - 1\right]}\right) \\ & + (1-\beta)\left(-\frac{a-z}{v-vc} + \frac{a\left[\exp\left(-2(v-vc)(a-z)/s^2\right) - 1\right]}{(v-vc)\left[\exp\left(-2(v-vc)a/s^2\right) - 1\right]}\right). \end{aligned} \tag{12}$$

In what follows, we will fix $Pc_{F,B}$ to two percentages (i.e., 90 and 95%). Then, we calculate MRT_B for each combination of starting point z and drift rate criterion v_c that yields the predetermined value of $Pc_{F,B}$ and determine which combination is optimal in the sense that it results in the lowest MRT_B.

Firstly we considered the following set of parameters: mean drift rate $v = 0.2$, boundary separation $a = 0.12$, and bias $\beta = 0.8$[5]. The top-left panel of **Figure 6** shows that an accuracy level of 90% may be achieved by a combination of shifts for starting point z and shifts of drift rate criterion v_c. Again, both parameters exist in a trade-off relationship, such that higher values of z combined with lower values of v_c produce the same mean proportion correct.

The bottom-left panel of **Figure 6** shows MRT for each calculated combination of starting point z and drift rate criterion v_c. The x-axis shows the value for starting point z (as a proportion of boundary separation a), the associated value for drift rate criterion v_c can be found in the top-left panel. The MRT results show that the lowest value for MRT is reached when all bias is accounted for by a shift in starting point z.

Secondly we considered a mean drift rate $v = 0.3$ and an accuracy level of 95%. The results, shown in the right two panels of **Figure 6**, are qualitatively similar[6].

In sum, when there is no response deadline and across-trial difficulty is fixed, the optimal way to deal with bias is by shifting the starting point toward the biased response alternative, without shifting the drift rate criterion. In an empirical paper by Simen et al. (2009), it was demonstrated that human participants do indeed shift starting point toward the biased response alternative.

[5]For the remainder of this article we set non-decision time $T_{er} = 0$.

[6]We explored a range of values for Bias β, accuracy, and drift rate. The results were always qualitatively similar: the lowest MRT was achieved when $v_c = 0$.

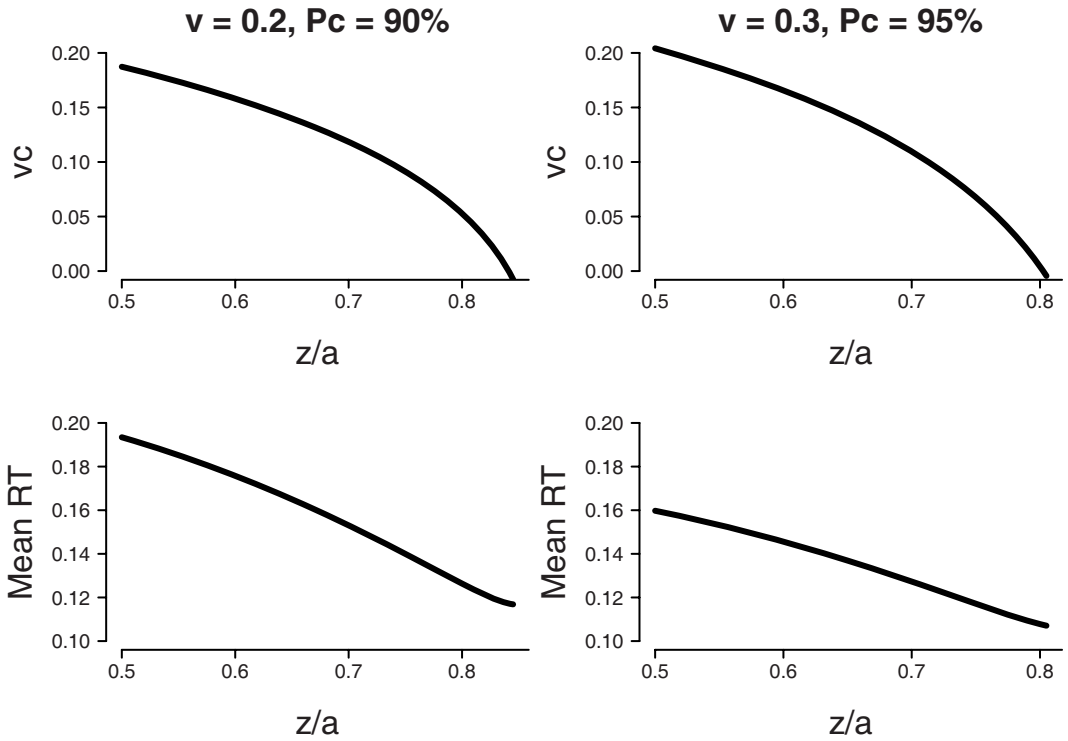


FIGURE 6 | Fixed difficulty ($\eta = 0$) in the free response paradigm. Minimum MRT is achieved when all bias is accounted for by a shift in starting point z. Bias $\beta = 0.8$, boundary separation $a = 0.12$. Top panel: Parameter combinations of starting point as a ratio of boundary separation z/a and shift in drift rate criterion v_c that lead to the same fixed level of accuracy (90% for $v = 0.2$, 95% for $v = 0.3$). Bottom panel: Mean RT that corresponds to the parameter combinations of the top panel.

No explicit mention of shifts in drift rate criterion are made. In the next subsection, we examine optimal performance for variable stimulus difficulty across trials.

Variable difficulty

Unfortunately, for variable difficulty there are no expressions for mean percentage correct and MRT. As such, we approximate the results for fixed difficulty by numerically obtaining combinations of starting point z and drift rate criterion v_c that yield two percentages correct (i.e., 90 and 95%). Then, we determine which combination is optimal in the sense that it results in the lowest MRT.

The top-left panel of **Figure 7** shows the set of parameter values of starting point z and drift rate shift v_c that lead to a percentage correct of 90% with a mean drift rate $v = 0.2$. The top-right panel shows the set of parameter values of starting point z and drift rate shift v_c that lead to a percentage correct of 95% with a mean drift rate $v = 0.3$. For both panels, boundary separation $a = 0.12$, bias $\beta = 0.8$, and standard deviation of drift rate $\eta = 0.1$. The bottom panels of **Figure 7** show corresponding values of MRT. The MRT results show that as for fixed difficulty, the lowest value for MRT is reached when all bias is accounted for by a shift in starting point z.

In sum, in paradigms with variable across-trial difficulty, but no response deadline, the optimal way to deal with bias is by shifting the starting point toward the biased response alternative, without shifting the drift rate criterion. This result mirrors the result for fixed difficulty.

Interim conclusion

Hanks et al. (2011) claimed that for biased decisions with variable across-trial stimulus difficulty, optimal performance requires not just a shift in starting point but also a shift in drift rate criterion (i.e., dynamic bias). Our results challenge this claim: regardless of whether stimulus difficulty is fixed or variable, optimal performance can be obtained by having bias only shift the starting point, and not the drift rate criterion. In the next section, we will investigate how people perform in practice.

BIAS IN PRACTICE

We have demonstrated that optimal performance in decision conditions with a biased response alternative can be achieved by shifting only the starting point criterion. However, people may not accommodate bias in an optimal manner. For instance, Hanks et al. (2011) demonstrated that in a decision environment with variable across-trial stimulus difficulty, participants accommodate advance information by dynamic bias. The authors did not, however, directly compare the performance of participants in conditions with fixed and variable across-trial stimulus difficulty. In this section, we perform such a comparison and address the question whether the inclusion of variability in stimulus difficulty alters the way in which people accommodate advance information. We

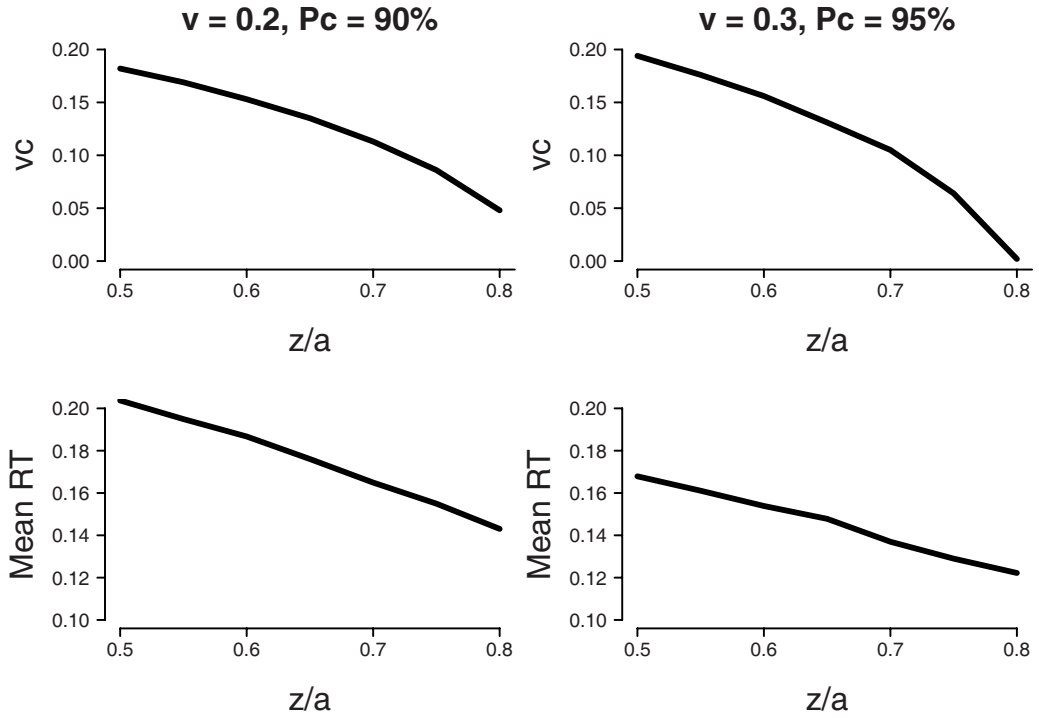


FIGURE 7 | Variable difficulty ($\eta = 0.1$) in the free response paradigm. Minimum MRT is achieved when all bias is accounted for by a shift in starting point z. Bias $\beta = 0.8$, boundary separation $a = 0.12$. Top panel: Parameter combinations of starting point as a ratio of boundary separation z/a and shift in drift rate criterion v_c that lead to the same fixed level of accuracy (90% for $v = 0.2$, 95% for $v = 0.3$). Bottom panel: Mean RT that corresponds to the parameter combinations of the top panel.

also examine if the performance of people in practice corresponds to the theoretical optimality indicated in the previous sections.

Our experiment also allows us to test a prediction from the DDM, namely that increasing the variability in across-trial stimulus difficulty results in a higher estimate of across-trial drift rate variability η. The experiment used a random dot motion task (Newsome et al., 1989) with advance information about the upcoming direction of movement. In a within-subjects design, each participant was administered a condition with fixed stimulus difficulty (i.e., identical coherence of movement from trial-to-trial) and a condition with variable stimulus difficulty (i.e., variable coherence).

PARTICIPANTS

Eleven healthy participants (8 female), aged 19–40 years (mean 24.6) performed a random-dots motion (RDM) paradigm in exchange for course credit or a monetary reward of 28 euros. Participants were recruited through the University of Amsterdam and had normal or corrected-to-normal vision. The procedure was approved by the ethical review board at the University of Amsterdam and informed consent was obtained from each participant. According to self-report, no participant had a history of neurological, major medical, or psychiatric disorder.

Materials

Participants performed an RT version of an RDM task. Participants were instructed to maintain fixation on a cross at the middle of the screen and decide the direction of motion of a cloud of partially randomly moving white dots on a black background. The decision was made at any time during motion viewing with a left or right button press. The stimulus remained on screen until a choice was made. The motion stimuli were similar to those used elsewhere (e.g., Newsome and Paré, 1988; Britten et al., 1992; Gold and Shadlen, 2003; Palmer et al., 2005; Ratcliff and McKoon, 2008; Mulder et al., 2010): white dots, with a size of 3×3 pixels, moved within a circle with diameter of 5° with a speed of 5°/s and a density of 16.7 dots/degree^2/s on a black background. On the first three frames of the motion stimulus, the dots were located in random positions. For each of these frames the dots were repositioned after two subsequent frames (the dots in frame one were repositioned in frame four, the dots in frame two were repositioned in frame five, etc.). For each dot, the new location was either random or in line with the motion direction. The probability that a dot moved coherent with the motion direction is defined as coherence. For example, at a coherence of 50%, each dot had a probability of 50% to participate in the motion–stimulus, every third frame (see also Britten et al., 1992; Gold and Shadlen, 2003; Palmer et al., 2005).

Visual stimuli were generated on a personal computer (Intel Core2 Quad 2.66 GHz processor, 3 GB RAM, two graphical cards: nvidia GeForce 8400 GS and a nvidia GeForce 9500 GT, running MS Windows XP SP3) using custom software and the Psychophysics Toolbox Version 3.0.8 (Brainard, 1997) for Matlab (version 2007b, Mathworks, 1984).

Difficulty

To get acquainted with the task, each participant performed a practice block of 40 easy trials (60% coherence). To match the difficulty level of the motion stimuli across participants, each participant performed an additional block of 400 trials of randomly interleaved stimuli with different motion strengths (resp. 0, 10, 20, 40, and 80% coherence, 80 trials each). We fitted the DDM to the mean response times and accuracy data of this block using a maximum likelihood procedure, constraining the drift rates to be proportional to the coherence settings (Palmer et al., 2005). For each participant, the motion strength at 75% accuracy was then interpolated from the psychometric curve (predicted by the proportional-rate diffusion model) and used in the experimental blocks with fixed coherence across trials. For blocks with mixed coherences, we randomly sampled performance levels from a uniform distribution (range: 51–99), and interpolated for each randomly chosen performance level the associated coherence from the psychometric curve.

Design

Each participant performed four sessions of the RDM task. In each session, participants performed six blocks of 100 trials: three blocks with the coherence fixed across trials, and three blocks with coherence varied across trials. For each condition (fixed and variable coherence) there were two biased and one neutral block. Prior information was given at the start of each experimental block. In the first experimental block, participants were told that there was a larger probability that the dots will move to the left (left-bias). In the second block, participants were told that there was an equal probability that the dots will move to the left or to the right (neutral). In a third block, the instructions indicated that there was a larger probability that the dots will move to the right (right-bias). For the biased blocks, prior information was consistent with the stimulus direction in 80% of the trials. The sequence of conditions was counterbalanced across participants.

Analyses

In order to quantify bias in the starting point and the drift rate criterion, we fit the diffusion model to the data with the Diffusion Model Analysis Toolbox (DMAT, Vandekerckhove and Tuerlinckx, 2007). We estimated the following parameters: mean drift rate v, boundary separation a, non-decision time T_{er}, starting point z, standard deviation of drift rate η, range of starting point s_z, and range of non-decision time s_t.

For both the fixed difficulty condition and the variable difficulty condition we estimated a mean drift rate v for consistent, neutral, and inconsistent stimuli. This resulted in six different estimates for mean drift rate v. In addition, for both the fixed difficulty condition and the variable difficulty condition we estimated a starting point z for left-biased, neutral, and right-biased stimuli. This resulted in six different estimates for starting point z. Furthermore, for the fixed difficulty condition and the variable difficulty condition we estimated a boundary separation a and a standard deviation of drift rate η, resulting in two different estimates for these parameter. Finally, we constrained non-decision time T_{er}, range of starting point s_z, and range of non-decision time s_t to be equal over stimulus type, conditions, and sessions, resulting in a single estimate for each of those parameters.

Starting point bias was calculated as half of the difference between the starting point z for left-biased stimuli and the starting point z for right-biased stimuli, scaled by boundary separation a. The maximum bias in starting point was therefore 50%. Drift rate criterion bias was calculated as half of the difference between mean drift rate v for consistent stimuli and mean drift rate v for inconsistent stimuli, scaled by the sum of mean drift rate v for consistent and inconsistent stimuli. The maximum bias in drift rate criterion was therefore 50% as well[7].

RESULTS

For the presented analyses, we report Bayesian posterior probabilities in addition to conventional p-values. When we assume, for fairness, that the null-hypothesis and the alternative hypothesis are equally plausible *a priori*, a default Bayesian t-test (Rouder et al., 2009) allows one to determine the posterior plausibility of the null-hypothesis and the alternative hypothesis. We denote the posterior probability for the null-hypothesis as p_{H0}^{Bayes}. When, for example, $p_{H0}^{Bayes} = 0.9$, this means that the plausibility for the null-hypothesis has increased from 0.5 to 0.9. Posterior probabilities avoid the problems that plague p-values, allow one to directly quantify evidence in favor of the null-hypothesis, and arguably relate more closely to what researchers want to know (e.g., Wagenmakers, 2007).

In order to assess whether our manipulation of across-trial parameter difficulty was successful, we compared the estimate of across-trial drift rate variability η between the fixed and variable difficulty conditions. The results show that across-trial drift rate variability η was larger in the variable difficulty condition (mean: 0.23) then in the fixed difficulty condition (mean: 0.14; $t(20) = 2.25$, $p < 0.05$, $p_{H0}^{Bayes} = 0.34$), suggesting that the across-trial difficulty manipulation was successful[8].

The left panel of **Figure 8** shows both starting point bias and drift rate criterion bias for the fixed difficulty condition and the variable difficulty condition. In the fixed difficulty condition, both the bias in starting point and the bias in drift rate are larger than zero ($t(10) = 5.85$, $p < 0.01$, $p_{H0}^{Bayes} < 0.01$, and $t(10) = 3.43$, $p < 0.01$, $p_{H0}^{Bayes} = 0.10$, respectively). In the variable difficulty condition, both the bias in starting point and the bias in drift rate also are larger than zero ($t(10) = 6.89$, $p < 0.01$, $p_{H0}^{Bayes} < 0.01$, and $t(10) = 4.14$, $p < 0.01$, $p_{H0}^{Bayes} = 0.04$, respectively). The right panel of **Figure 8** shows MRT and proportion correct for consistent, neutral, and inconsistent stimuli for the fixed and variable difficulty conditions.

Since both types of bias are measured on different scales, they cannot be compared directly. The important claim by Hanks et al. (2011) is that bias in the drift rate criterion is larger in the variable difficulty condition than in the fixed difficulty condition. In order to assess this claim we compared both types of bias directly

[7] In theory, negative drift rates could lead to a larger bias. However, none of our participants had negative drift rates for any of the stimulus types.

[8] Note that the alternative hypothesis is roughly twice as likely as the null-hypothesis according to p_{H0}^{Bayes}.

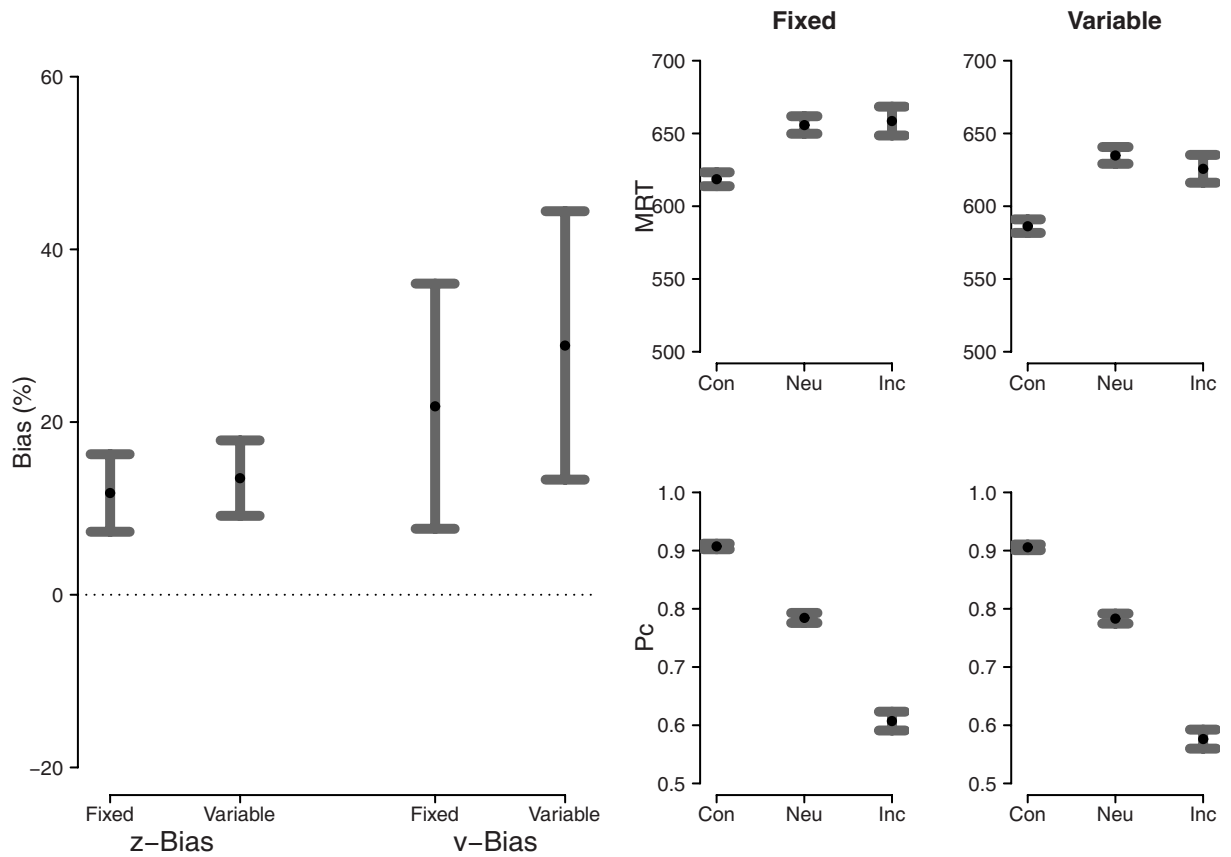


FIGURE 8 | Left panel: starting point bias and drift rate criterion bias. Right four panels: MRT (top) and proportion correct (bottom) for fixed (left) and variable (right) difficulty. Dots represent the mean, error bars represent 95% confidence intervals. Con, consistent; Neu, neutral; Inc, inconsistent.

between the two conditions. Consistent with the visual impression from the left panel of **Figure 8**, the statistical analysis reveal no differences for starting point bias and drift rate criterion bias ($t(20) = 0.61$, $p > 0.05$, $p_{H0}^{Bayes} = 0.74$, and $t(20) = 0.75$, $p > 0.05$, $p_{H0}^{Bayes} = 0.73$, respectively).

In sum, both in the fixed difficulty condition and in the variable difficulty condition, participants exhibit bias in starting point and drift rate criterion. The results suggest there is no difference between the two conditions in the amount of each type of bias, challenging the claim by Hanks et al. (2011) that bias in the drift rate criterion should be more pronounced when stimulus difficulty is variable than when it is fixed.

MODEL PREDICTIVES

In cognitive modeling, model fit can be assessed by means of model predictives. Model predictives are simulated data generated from the cognitive model, based on the parameter estimates for the real data. If the generated data closely resemble the real empirical data, then the model fit is deemed adequate (e.g., Gelman and Hill, 2007).

For this experiment, we drew 100 samples from the real data set and generated diffusion model parameter estimates for each participant and each condition separately using a bootstrap procedure. For each of these samples, we generated synthetic data, for which we calculated the 0.1, 0.3, 0.5, 0.7, and 0.9 RT quantiles for both the real data set and the synthetic data set. The real and synthetic RT quantiles are shown in **Figure 9**.

Figure 9 shows a quantile probability plot (e.g., Ratcliff, 2002), where the left-hand side represents error RTs for the five quantiles, and the right-hand side represents correct RTs for those same quantiles. The circles represent stimuli that were consistent with the biased response direction, the squares represent stimuli that were neutral, and the triangles represent stimuli that were inconsistent with the biased response direction. The filled symbols in the figure show the empirical data, the open symbols show the simulated data that were generated using the DMAT parameter estimates.

For response accuracy, the correspondence between the empirical data and the synthetic data can be judged by the horizontal disparity between the data points and the model points. **Figure 9** shows that the diffusion model captures the error rate reasonably well for most of the stimulus types, as indicated by the horizontal disparity between the filled and open symbols. The diffusion model does capture the RTs well, as can be judged from the vertical disparity between the filled and open symbols. The quantile probability plot nicely shows how correct responses are fastest for consistent stimuli and slowest for inconsistent stimuli, whereas error responses are fastest for inconsistent stimuli and slowest for

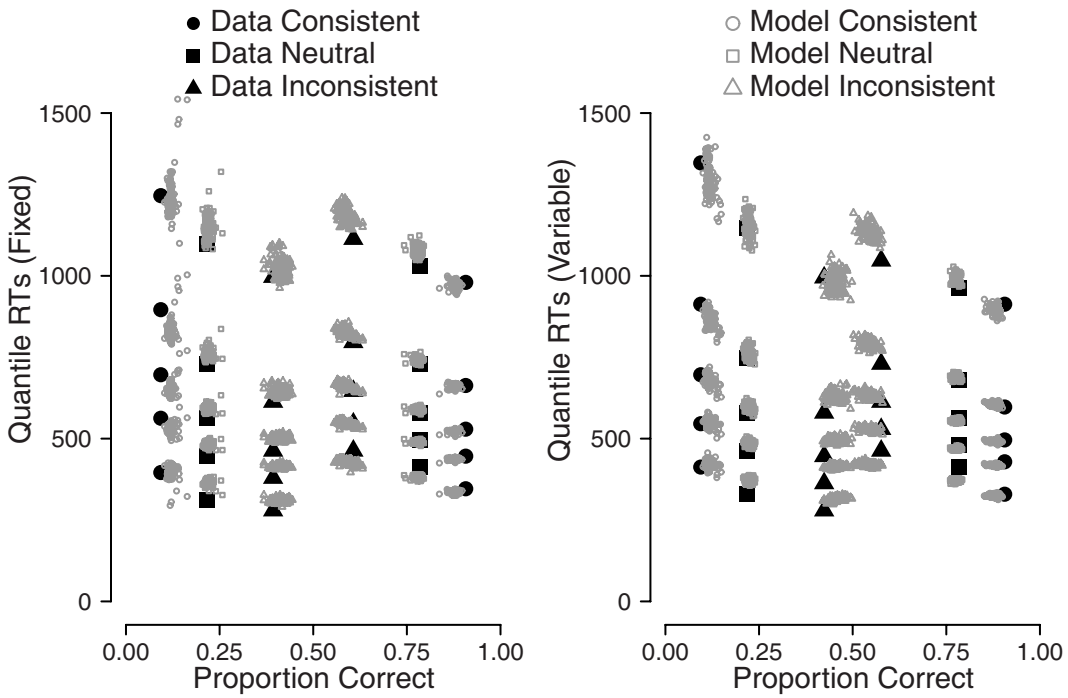


FIGURE 9 | Model predictives indicate that the parameter estimates of DMAT describe the data well. Filled symbols: empirical data. Open symbols: bootstrapped synthetic data, based on the model parameter estimates.

consistent stimuli, which is consistent with prior bias (see, e.g., Mulder et al., 2012).

CONCLUSION

It is not straightforward to perform optimally in response time tasks in which one response option is more likely than the other. When stimulus difficulty is fixed across trials, Edwards (1965) has demonstrated that optimal performance requires advance information to be accommodated solely by a shift in starting point. Recently, Hanks et al. (2011) claimed that when stimulus difficulty varies from trial-to-trial, optimal performance also requires a shift in drift rate criterion (i.e., dynamic bias).

The contribution of this paper is twofold. Firstly, we demonstrated that in theory, optimal performance can be achieved by a shift in starting point only. This result holds regardless of whether stimulus difficulty is fixed or variable from trial-to-trial. Secondly, we presented empirical data showing that people accommodate bias similarly for conditions of fixed and variable across-trial difficulty. Specifically, decision makers incorporate both prior and dynamic bias, and no evidence suggested that the presence of variability in stimulus difficulty made participants rely more on shifts in the drift rate criterion.

In the theoretical part, we first considered the interrogation paradigm. We demonstrated that optimal performance can be achieved by entertaining a bias in starting point, by entertaining a bias in the drift rate criterion, or by a combination of the two. There was no qualitative difference between conditions of fixed and variable across-trial stimulus difficulty.

It could be argued that the theoretical results of the interrogation paradigm are unlikely to apply in practice. Specifically, participants need to know the exact moment of the deadline T to be able to utilize the trade-off between starting point and drift rate criterion. Also, the argument of Hanks et al. (2011) depends on the time course of the decision process: long decisions are most likely difficult decisions, so the longer a decision takes, the more adaptive it becomes to select the biased response alternative. In the interrogation paradigm, however, each decision process takes an identical amount of time.

Another complication when interpreting results for the interrogation paradigm is that for variable difficulty, the optimal setting of the starting point depends on the time deadline (see equation (7)). Because the decision maker does not know the exact value of the time deadline, a more realistic expression should integrate over some unknown distribution of time deadlines that describes the uncertainty in T on the part of the decision maker.

In a task setting without response deadline, our results show that optimal performance is achieved by shifting only the starting point; additionally shifting the drift rate criterion only serves to deteriorate performance. Crucially, we found that this result is true for both fixed and variable across-trial stimulus difficulty. As such, our results conflict with the claim by Hanks et al. (2011) that optimal decision makers should entertain a shift in drift rate criterion to accommodate bias under conditions of variable stimulus difficulty.

In the empirical part, we conducted an experiment in which we manipulated across-trial stimulus difficulty in order to investigate

if performance of decision makers is optimal. We also wanted to test whether decision makers accommodate bias differently in decision environments with fixed and variable across-trial stimulus difficulty. We successfully manipulated across-trial variability in difficulty, as evidenced by a higher value of across-trial drift rate variability η for variable across-trial difficulty than for fixed-trial difficulty. Our results showed that performance of participants was not optimal: decision makers implemented both a shift in starting point and a shift in drift rate criterion in order to deal with bias in prior information about the decision alternatives. Contrary to the theory of Hanks et al. (2011), there was no difference in the implementation of dynamic bias between conditions of fixed and variable across-trial stimulus difficulty.

In sum, we conclude that dynamic bias is not needed for optimal performance, not when stimulus difficulty is fixed and not when it is variable. From the perspective of optimality, advance information should affect only prior bias (i.e., starting point), such that the evidence threshold is lowered for the choice alternative that is most likely to be correct *a priori*. In practice it appears that people use both prior bias and dynamic bias; our experiment suggests that increasing the variability of stimulus difficulty does not cause participants to accommodate advance information preferentially by shifting the drift rate criterion.

ACKNOWLEDGMENTS

This research was supported by a Vidi grant from the Netherlands Organisation for Scientific Research (NWO). This paper is part of the Research Priority Program Brain and Cognition at the University of Amsterdam. We thank Anja Sommavilla for testing the participants. This publication was supported by an Open Access grant from the Netherlands Organisation for Scientific Research (NWO).

REFERENCES

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., and Cohen, J. D. (2006). The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced choice tasks. *Psychol. Rev.* 113, 700–765.

Brainard, D. H. (1997). The psychophysics toolbox. *Spat. Vis.* 10, 433–436.

Britten, K. H., Shadlen, M. N., Newsome, W. T., and Movshon, J. A. (1992). The analysis of visual motion: a comparison of neuronal and psychophysical performance. *J. Neurosci.* 12, 4745–4765.

Diederich, A., and Busemeyer, J. R. (2006). Modeling the effects of payoff on response bias in a perceptual discrimination task: bound-change, drift-rate-change, or two-stage-processing hypothesis. *Percept. Psychophys.* 68, 194–207.

Edwards, W. (1965). Optimal strategies for seeking information: models for statistics, choice reaction times, and human information processing. *J. Math. Psychol.* 2, 312–329.

Feng, S., Holmes, P., Rorie, A., and Newsome, W. T. (2009). Can monkeys choose optimally when faced with noisy stimuli and unequal rewards? *PLoS Comput. Biol.* 5, e1000284. doi:10.1371/journal.pcbi.1000284

Gao, J., Tortell, R., and McClelland, J. L. (2011). Dynamic integration of reward and stimulus information in perceptual decision-making. *PLoS ONE* 6, e16749. doi:10.1371/journal.pone.0016749

Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (2004). *Bayesian Data Analysis*, 2nd Edn. Boca Raton, FL: Chapman and Hall/CRC.

Gelman, A., and Hill, J. (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge: Cambridge University Press.

Gold, J. I., and Shadlen, M. N. (2003). The influence of behavioral context on the representation of a perceptual decision in developing oculomotor commands. *J. Neurosci.* 23, 632–651.

Grasman, R. P. P. P., Wagenmakers, E.-J., and van der Maas, H. L. J. (2009). On the mean and variance of response times under the diffusion model with an application to parameter estimation. *J. Math. Psychol.* 53, 55–68.

Hanks, T. D., Mazurek, M. E., Kiani, R., Hopp, E., and Shadlen, M. N. (2011). Elapsed decision time affects the weighting of prior probability in a perceptual decision task. *J. Neurosci.* 31, 6339–6352.

Laming, D. R. J. (1968). *Information Theory of Choice-Reaction Times*. London: Academic Press.

Luce, R. D. (1986). *Response Times*. New York: Oxford University Press.

Mathworks. (1984). *Matlab*. Available at: http://www.mathworks.com/products/matlab

Mulder, M. J., Bos, D., Weusten, J. M. H., van Belle, J., van Dijk, S. C., Simen, P., van Engeland, H., and Durston, S. (2010). Basic impairments in regulating the speed-accuracy tradeoff predict symptoms of attention-deficit/hyperactivity disorder. *Biol. Psychiatry* 68, 1114–1119.

Mulder, M. J., Wagenmakers, E. -J., Ratcliff, R., Boekel, W., and Forstmann, B. U. (2012). Bias in the brain: a diffusion model analysis of prior probability and potential payoff. *J. Neurosci.* 32, 2335–2343.

Newsome, W. T., Britten, K. H., and Movshon, J. A. (1989). Neuronal correlates of a perceptual decision. *Nature* 341, 52–54.

Newsome, W. T., and Paré, E. B. (1988). A selective impairment of motion perception following lesions of the middle temporal visual area (MT). *J. Neurosci.* 8, 2201–2211.

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.

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

Ratcliff, R. (1985). Theoretical interpretations of the speed and accuracy of positive and negative responses. *Psychol. Rev.* 92, 212–225.

Ratcliff, R. (2002). A diffusion model account of response time and accuracy in a brightness discrimination task: fitting real data and failing to fit fake but plausible data. *Psychon. Bull. Rev.* 9, 278–291.

Ratcliff, R., and McKoon, G. (2008). The diffusion decision model: theory and data for two-choice decision tasks. *Neural Comput.* 20, 873–922.

Ratcliff, R., and Rouder, J. N. (2000). A diffusion model account of masking in two-choice letter identification. *J. Exp. Psychol. Hum. Percept. Perform.* 26, 127–140.

Ratcliff, R., and Tuerlinckx, F. (2002). Estimating parameters of the diffusion model: approaches to dealing with contaminant reaction times and parameter variability. *Psychon. Bull. Rev.* 9, 438–481.

Rorie, A. E., Gao, J., McClelland, J. L., and Newsome, W. T. (2010). Integration of sensory and reward information during perceptual decision-making in lateral intraparietal cortex (LIP) of the macaque monkey. *PLoS ONE* 5, e9308. doi:10.1371/journal.pone.0009308

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.

Simen, P., Contreras, D., Buck, C., Hu, P., Holmes, P., and Cohen, J. D. (2009). Reward rate optimization in two-alternative decision making: empirical tests of theoretical predictions. *J. Exp. Psychol. Hum. Percept. Perform.* 35, 1865–1897.

van Ravenzwaaij, D., Brown, S., and Wagenmakers, E.-J. (2011). An integrated perspective on the relation between response speed and intelligence. *Cognition* 119, 381–393.

van Ravenzwaaij, D., van der Maas, H. L. J., and Wagenmakers, E. –J. (2012). Optimal decision making in neural inhibition models. *Psychol. Rev.* 119, 201–215.

Vandekerckhove, J., and Tuerlinckx, F. (2007). Fitting the Ratcliff diffusion model to experimental data. *Psychon. Bull. Rev.* 14, 1011–1026.

Wagenmakers, E.-J. (2007). A practical solution to the pervasive problems of p-values. *Psychon. Bull. Rev.* 14, 779–804.

Wagenmakers, E.-J. (2009). Methodological and empirical developments for the Ratcliff diffusion model of response times and accuracy. *Eur. J. Cogn. Psychol.* 21, 641–671.

Wagenmakers, E.-J., van der Maas, H. J. L., and Grasman, R. P. P. P. (2007). An EZ-diffusion model for response time and accuracy. *Psychon. Bull. Rev.* 14, 3–22.

Wald, A., and Wolfowitz, J. (1948). Optimal character of the sequential probability ratio test. *Ann. Math. Stat.* 19, 326–339.

Yang, T., Hanks, T. D., Mazurek, M., McKinley, M., Palmer, J., and Shadlen, M. N. (2005). "Incorporating prior probability into decision-making in the face of uncertain reliability of evidence," in

Poster Presented at the Meeting of the Society for Neuroscience, Washington, DC.

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 January 2012; paper pending published: 16 February 2012; accepted: 13 April 2012; published online: 29 May 2012.
Citation: van Ravenzwaaij D, Mulder MJ, Tuerlinckx F and Wagenmakers E-J (2012) Do the dynamics of prior information depend on task context? An analysis of optimal performance and an empirical test. Front. Psychology ***3****:132. doi: 10.3389/fpsyg.2012.00132*
This article was submitted to Frontiers in Cognitive Science, a specialty of Frontiers in Psychology.

APPENDIX

DERIVATION OF EQUATION (6)

To understand the derivation of equation (6) from equation (5), let us focus on the first part $\Phi\left[\frac{(\xi+v_c)T+z}{s\sqrt{T}}\right]$ (the treatment of the second part is analogous) and write it as $\Phi[A\xi+B]$, with $A=\frac{\sqrt{T}}{s}$ and $B=\frac{v_cT+z}{s\sqrt{T}}$. We then have the following integral I:

$$I=\int_{-\infty}^{+\infty}\Phi\left[A\xi+B\right]\phi\left(\xi;v,\eta^2\right)d\xi,$$

where $\phi(\xi;v,\eta^2)$ stands for the normal density function for ξ with mean v and variance η^2. We can continue as follows:

$$\begin{aligned}I&=\int_{-\infty}^{+\infty}\Phi\left[A\xi+B\right]\phi\left(\xi;v,\eta^2\right)d\xi\\&=\int_{-\infty}^{+\infty}\int_{-\infty}^{A\xi+B}\phi\left(x;0,1\right)dx\phi\left(\xi;v,\eta^2\right)d\xi\\&=\int_{-\infty}^{+\infty}\int_{-\infty}^{B}\phi\left(x;-A\xi,1\right)\phi\left(\xi;v,\eta^2\right)dxd\xi\\&=\int_{-\infty}^{+\infty}\int_{-\infty}^{B/A}\phi\left(x;-\xi,1/A^2\right)\phi\left(\xi;v,\eta^2\right)dxd\xi\\&=\int_{-\infty}^{B/A}\int_{-\infty}^{+\infty}\phi\left(x;-\xi,1/A^2\right)\phi\left(\xi;v,\eta^2\right)d\xi\,dx.\end{aligned}$$

The inner integral is a known integral in Bayesian statistics (see, e.g., Gelman et al., 2004). The kernel of the product of the two normal distributions contains an exponent with quadratic terms in ξ and x (Gelman et al., 2004). Thus, ξ and x have a bivariate normal distribution and thus marginalizing over ξ results in a normal distribution for x.

Using the double expectation theorem (see Gelman et al., 2004), we can find the marginal mean of x:

$$\begin{aligned}E(x)&=E\left[E(x|\xi)\right]\\&=E[-\xi]\\&=-v.\end{aligned}$$

Applying a similar theorem for the marginal variance of x gives (Gelman et al., 2004):

$$\begin{aligned}\mathrm{Var}(x)&=E\left[\mathrm{Var}(x|\xi)\right]+\mathrm{Var}\left[E(x|\xi)\right]\\&=E\left[1/A^2\right]+\mathrm{Var}(-\xi)\\&=1/A^2+\eta^2.\end{aligned}$$

Therefore, we can simplify the inner integral to $\phi(x;-v,1/A^2+\eta^2)$:

$$\begin{aligned}I&=\int_{-\infty}^{B/A}\phi\left(x;-v,1/A^2+\eta^2\right)dx\\&=\int_{-\infty}^{\frac{B/A+v}{\sqrt{1/A^2+\eta^2}}}\phi\left(x;0,1\right)dx\\&=\Phi\left[\frac{B/A+v}{\sqrt{1/A^2+\eta^2}}\right]\\&=\Phi\left[\frac{\frac{v_cT+z}{T}+v}{\sqrt{s^2/T+\eta^2}}\right]\\&=\Phi\left[\frac{(v_c+v)\,T+z}{s\sqrt{T+T^2\eta^2/s^2}}\right],\end{aligned}$$

which is, after multiplication with β, equal to the first term of equation (6). The second part can be found in a similar way.

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 16 July 2012
doi: 10.3389/fpsyg.2012.00213

Can post-error dynamics explain sequential reaction time patterns?

***Stephanie Goldfarb*[1,2]*, *KongFatt Wong-Lin*[3], *Michael Schwemmer*[1,4], *Naomi Ehrich Leonard*[1,2,4] and *Philip Holmes*[1,2,4]**

[1] *Princeton Neuroscience Institute, Princeton University, NJ, USA*
[2] *Department of Mechanical and Aerospace Engineering, Princeton University, NJ, USA*
[3] *Intelligent Systems Research Centre, University of Ulster, Magee Campus, Northern Ireland, UK*
[4] *Program in Applied and Computational Mathematics, Princeton University, NJ, USA*

Edited by:
Marius Usher, Tel-Aviv University, Israel

Reviewed by:
Eddy J. Davelaar, Birkbeck College, UK
Don Van Ravenzwaaij, University of Amsterdam, Netherlands

***Correspondence:**
Stephanie Goldfarb, Department of Mechanical and Aerospace Engineering, Princeton University, NJ 08544, USA.
e-mail: sgoldfar@princeton.edu

We investigate human error dynamics in sequential two-alternative choice tasks. When subjects repeatedly discriminate between two stimuli, their error rates and reaction times (RTs) systematically depend on prior sequences of stimuli. We analyze these sequential effects on RTs, separating error and correct responses, and identify a sequential RT trade-off: a sequence of stimuli which yields a relatively fast RT on error trials will produce a relatively slow RT on correct trials and vice versa. We reanalyze previous data and acquire and analyze new data in a choice task with stimulus sequences generated by a first-order Markov process having unequal probabilities of repetitions and alternations. We then show that relationships among these stimulus sequences and the corresponding RTs for correct trials, error trials, and averaged over all trials are significantly influenced by the probability of alternations; these relationships have not been captured by previous models. Finally, we show that simple, sequential updates to the initial condition and thresholds of a pure drift diffusion model can account for the trends in RT for correct and error trials. Our results suggest that error-based parameter adjustments are critical to modeling sequential effects.

Keywords: drift diffusion model, error rate, perceptual decision making, post-error slowing, reaction time, sequential effects

1. INTRODUCTION

Efforts to model and predict human behavior are informed by an understanding of the dynamics of error rates (ERs) and reaction times (RTs) in simple tasks. In particular, in two-alternative forced-choice (TAFC) tasks (e.g., Laming, 1968; Link, 1975; Link and Heath, 1975; Ratcliff and Rouder, 1998) human participants are known to slow down after committing an error, and generally to exhibit RTs and ERs that systematically depend on prior stimulus sequences (Bertelson, 1961; Capaldi, 1966; Laming, 1968; Remington, 1969; Kirby, 1976; Vervaeck and Boer, 1980; Soetens et al., 1984, 1985). However, while much previous work has considered post-error slowing and sequential effects separately, we are not aware of studies that explicitly account for interactions among these effects.

Patterns in RTs for individual trials are well documented in the literature. In particular, relative to their mean RTs on correct trials, subjects are known to respond faster on error trials and more slowly immediately following errors (Rabbitt, 1966; Laming, 1979a,b). On average it has been shown that participants return to their mean RT values within two trials after an error (Rabbitt, 1968b). Various models of TAFC tasks have accounted for this post-error slowing (Ratcliff and Rouder, 1998; Dudschig and Jentzsch, 2009). However, to our knowledge the mean RTs on trials following specific sequences of stimuli have not been studied independently for trials ending in an error, and deliberate post-error adjustments have not been incorporated into models of sequential effects.

Moreover, the characteristic patterns in speed and accuracy following sequences of repetitions and alternations are well documented only for tasks in which the stimuli are equally likely. While overall trends in speed and accuracy have received significant attention (Carpenter and Williams, 1995; Ratcliff and Smith, 2004; Bogacz et al., 2006; Simen et al., 2009), for tasks in which the stimuli are not equally likely, such sequential patterns in mean RT have not been considered.

In a majority of TAFC studies, participants are either rewarded equally for overall participation or they are rewarded for each correct response. However, several studies (e.g., Corrado et al., 2005; Feng et al., 2009) have investigated tasks which reward correct responses to one stimulus more highly than another and have shown that reward contingencies influence choice behavior. When reward probabilities or reward values are unequal, participants are known to select the stimulus corresponding to the most probable or most valuable reward more frequently (Corrado et al., 2005; Feng et al., 2009), and they may do so almost optimally (Gao et al., 2011).

When stimuli are equally probable and correct responses are equally rewarded, several effects are known. For small (<500 ms) response to stimulus intervals (RSIs), the behavior typically illustrates automatic facilitation (AF): mean RTs on the current trial are

faster if the previous trial was a repetition, regardless of whether the current trial is a repetition or an alternation. For slow RSIs (≈1000 ms), mean RTs on the current trial after a series of alternations are faster if the current trial is a repetition and slower if the current trial is an alternation (Bertelson, 1961; Laming, 1968; Kirby, 1976). This effect is called strategic expectancy (SE), suggesting a relationship between a participant's expectations and his or her reaction time. Moreover, a transition occurs from AF to SE as RSI increases (Soetens et al., 1985; Jentzsch and Sommer, 2002) and can be illustrated graphically (Audley, 1973). Prior to the present paper, it was unknown whether such a transition from AF to SE could also occur for a constant RSI with increasing probability of alternations, or, more generally, how sequential effects carry over from equally probable to biased stimuli.

In this paper, we study sequential patterns in ERs as well as in RTs for error and correct responses independently in TAFC tasks in which stimuli are equally probable or strongly biased toward repetitions or alternations, focusing on sequences of three trials. Stimulus sequences can be biased by specifying stimulus probabilities (state orientation) or by specifying transition probabilities between states (transition orientation), and it is known that these processes produce distinct response patterns in RT (Brodersen et al., 2008). Since we are interested in sequential effects and expectancy, we generate stimuli by a first-order Markov process with unequal (as well as equal) transition probabilities (see **Figure 1**): The unequal case sets the probability of an alternation (P_A) to be unequal to the probability of a repetition ($1 - P_A$). Transition probabilities P_A and $1 - P_A$ are held fixed over blocks of trials, and we use relatively long RSIs (800 and 1,000 ms mean), for which SE is most apparent. We reanalyze behavioral data from an equal-probability experiment (Cho et al., 2002), and we collect and analyze a new data set with P_A set to 10, 50, and 90%. We find significant transition probability effects on RTs for error and correct responses and on ERs.

To further study patterns in RT and ER we extend the pure drift diffusion model (DDM) to account for sequential patterns. The pure DDM describes choice between two alternatives by representing the noisy accumulation of the difference in evidence (logarithmic likelihood) from a given initial condition to one of two decision thresholds. This process is known to mimic aspects of neural integration (Carpenter and Williams, 1995; Gold and Shadlen, 2000; Bogacz et al., 2006; Gold et al., 2008). Adapting the DDM, we propose two simple update mechanisms to vary the initial condition and thresholds from trial to trial, depending on previous stimuli and response correctness. We show how our adapted DDM can account for the observed trends in RT for correct and error trials.

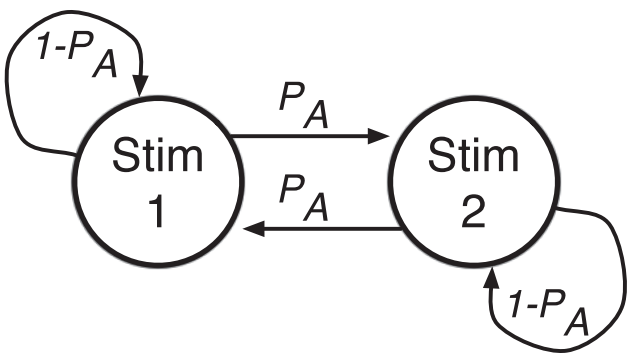


FIGURE 1 | Stimulus order is generated by a transition-oriented Markov process. Given current stimulus 1, the next stimulus will be stimulus 2 (an alternation) with probability P_A and stimulus 1 (a repetition) with probability $1 - P_A$.

Related TAFC models frequently involve a variant of the leaky competing accumulator (LCA; Usher and McClelland, 2001), featuring two coupled stochastic differential equations which contain multiple parameters to account for leakage (decay of previous evidence) and for the interaction between neural populations. LCA models have been shown to capture sequential effects for equally probable stimuli (Cho et al., 2002; Gao et al., 2009). For certain parameter ranges, it can be shown that the LCA, along with race, inhibition, and other models, reduces to a DDM, and the DDM itself may be extended to account for variability in the model parameters (Ratcliff and Smith, 2004; Bogacz et al., 2006). However, we are aware only of modeling studies that predict both ERs and RTs for sequential effects (Cho et al., 2002; Gao et al., 2009), and these studies did not analyze patterns in error RTs, nor did they incorporate post-error parameter adjustments into the analysis. Bayesian models of TAFC, which can also be represented by DDMs for certain parameter ranges (Liu et al., 2009), have also been used to model sequential effects (Yu and Cohen, 2009; Wilder et al., 2009), but none of these models yet accounts for patterns in errors.

Physiological evidence suggests sources of systematic changes in behavior from trial to trial, providing some neurobiological basis for our proposed update mechanisms. An electroencephalogram (EEG) study has identified a SE pattern in the P300 response (Sommer et al., 1999), an event related potential signal which follows 300–600 ms after unexpected, alternating, stimuli. The prefrontal cortex is also activated following an alternation after frequent repetitions, with greater activation following a longer run of repetitions prior to the alternation (Huettel et al., 2005). In addition, the anterior cingulate cortex (ACC) is known to show increased activity with increased conflict in representation, or alternation of stimuli, and ACC activity has been linked to cognitive control and post-error corrections and corresponding increase in RT (Botvinick et al., 2001). Prior work has incorporated ACC conflict signals into models of sequential and error effects (Jones et al., 2002).

An understanding of the relationship between error correction and sequential biasing mechanisms may allow us to further differentiate between corresponding physiological processes. Such an understanding could have broad implications. Indeed, recent work suggests that the same mechanisms that account for sequential effects also account for sequence learning (Soetens et al., 2004): a general mechanism may therefore lend insight into sequence learning (Soetens et al., 1985; Pashler and Baylis, 1991a,b; Frensch and Miner, 1994), including linguistic processes. Further, better understanding of the mechanisms behind simple discrimination tasks may also allow for improved prediction and prevention of errors.

This paper is organized as follows. In Section 2, we describe two experiments: the first originally reported in Cho et al. (2002) and the second conducted specifically for the present study. We then describe a diffusion model account of participant behavior in the tasks. In Section 3, we describe the experimental results and discuss diffusion model fits to participant behavior. Finally, Section 4 contains further discussion and our conclusions, and it identifies directions for future experimental and modeling work.

2. MATERIALS AND METHODS

In this section, we describe the protocol followed for the two experiments presented in this paper. We then describe a general model of decision making, which accounts for choice behavior with two simple mechanistic adaptations to the pure drift diffusion model (DDM). Finally, we describe a procedure for fitting the model to match participant data in our adapted DDM.

2.1. EXPERIMENT 1: ERROR DYNAMICS IN UNBIASED TASKS

The first experiment (reanalyzed from Cho et al., 2002) served as a control in which stimulus probabilities were equal and transition probabilities were held constant at 50% throughout the experiment. As the details of the experiment have been described in the literature previously, we discuss them only briefly here. In Experiment 1, six Princeton University undergraduates participated in a task over a single session by identifying the upper or lowercase "o" character on the screen with the appropriate keypress. The index finger was used to identify the uppercase letter, and the middle finger to identify the lowercase letter. Each session consisted of 13 blocks of 120 trials each, and a response to stimulus interval (RSI) of 800 ms was used. Participants received course credit in exchange for their participation in the study. For additional details see Cho et al. (2002). No trials were omitted from our reanalysis.

2.2. EXPERIMENT 2: ERROR DYNAMICS IN TRANSITION-BIASED TASKS

In the second experiment stimulus transition probabilities were varied from block to block, so that in a given block a participant would have a constant high, medium, or low probability of alternations. That is, given the current stimulus 1, a participant would next see the other stimulus 2 with probability P_A and the same stimulus 1 again with probability $1 - P_A$, and the sequence of stimuli would be drawn from a transition-oriented Markov process, as shown in **Figure 1**.

2.2.1. Participants

Sixteen adults (6 males) participated in exchange for a standard payment of $12 per session of 9 blocks. Participants were recruited from the Princeton University community via announcements posted online and on campus. The experiment was approved by the Institutional Review Panel for Human Subjects of Princeton University, and all participants provided their informed consent prior to participation.

2.2.2. Stimuli

Participants performed an RT version of a motion discrimination task. The visual stimulus consisted of a black screen showing a cloud of white moving dots with a red, stationary fixation dot at its center. The red dot had size 0.30° visual angle, and the white dots had size 0.15° each and moved within a circle of diameter 10° at a speed of 7°/s and a density of 20 dots/degree^2. On each trial 90% of the white dots would move coherently in a given, "correct" direction, and the remaining white dots would move randomly. The high coherence of motion was selected to ensure that some processing was necessary but that the difficulty of the task would remain low, consistent with other studies of sequential effects. A decision could be indicated with a left or right keypress at any point after dots appeared on the screen. Responses were collected via the standard Macintosh computer keyboard, with the "Z" key used to indicate leftward motion and the "M" key used to indicate rightward motion. The experiment was performed on a Macintosh computer using the Psychophysics Toolbox extension (Brainard, 1997).

2.2.3. Procedure

The participants were instructed to fixate upon the red dot and then determine the direction of the moving dots. They were also instructed to complete the session as quickly and as accurately as possible. Each participant completed 1 session of approximately 60 min duration.

Each session consisted of 9 blocks of 200 trials each in which the P_A remained fixed at 10, 50, or 90% (3 blocks for each condition). The order of the blocks was constrained to follow a Latin Square design. Participants were allowed a short break between blocks. To minimize anticipatory responding, response to stimulus intervals were drawn from a gamma distribution with a mean of 1 s for each trial, following the convention set in previous sequential RT tasks (Rabbitt, 1966; Simen et al., 2006; Brodersen et al., 2008; Balci et al., 2011). Outlier RTs (less than 100 ms or greater than 900 ms, comprising less than 1.5% of the data) were not included in the analysis. We note that only the outlier was removed from the RT and error analysis; it was included in sorting RR, AR, RA and AA sequences, since it precedes a trial that is included in the analysis. In addition, one participant failed to follow instructions and the corresponding data were omitted from the analysis.

During each block in the session, the subjects received the following feedback. Correct responses were denoted with a short beep sound, and error and premature, anticipatory responses were denoted with a buzz sound. In addition, on every fifth trial, the number of correct responses provided in that block so far was briefly flashed across the screen. This was the only feedback that was provided. Participants were seated at a viewing distance of approximately 60 cm from the screen. Our protocol in Experiment 2 is similar to others in the literature (e.g., Newsome and Pare, 1988; Britten et al., 1992; Ratcliff and McKoon, 2008).

2.3. AN ADAPTED DRIFT DIFFUSION MODEL

To account for sequential effects and error effects, we consider a simple adaptation of the pure drift diffusion model (DDM) in which the initial condition and thresholds are updated sequentially following each trial (Ratcliff and Rouder, 1998; Ratcliff and Smith, 2004; Bogacz et al., 2006). In the pure DDM, information is accumulated stochastically according to the following equation:

$$dx = \mu\, dt + \sigma\, dW, x(0) = x_0. \quad (1)$$

Here $x(t)$ represents the difference in logarithmic likelihood ratio for the two choices, the drift rate μ (conventionally taken to be positive) represents the difference in incoming evidence for the correct alternative relative to the incorrect alternative, and $\sigma\,dW$ is a Wiener (white noise) process with mean 0 and variance σ^2. The evidence thresholds are set at $\pm z$, and noisy accumulation continues until $x(t)$ first crosses either $+z$ (a correct decision) or $-z$ (an error). If the non-decision time is then given by T_{nd} such that $\mathrm{RT} = \mathrm{DT} + T_{\mathrm{nd}}$ where DT is the decision time, it can be shown that the mean DT and ER are (Gardiner, 1985; Busemeyer and Townsend, 1992):

$$\langle \mathrm{DT} \rangle = \tilde{z}\tanh(\tilde{z}\tilde{\mu}) + \left\{ \frac{2\tilde{z}\left(1 - \exp\left(-2\,\tilde{x}_0\,\tilde{\mu}\right)\right)}{\exp\left(2\tilde{z}\tilde{\mu}\right) - \exp\left(-2\tilde{z}\tilde{\mu}\right)} - \tilde{x}_0 \right\}, \tag{2}$$

and

$$\langle \mathrm{ER} \rangle = \frac{1}{1 + \exp(2\tilde{z}\tilde{\mu})} - \left\{ \frac{1 - \exp(-2\,\tilde{x}_0\,\tilde{\mu})}{\exp(2\tilde{z}\tilde{\mu}) - \exp(-2\tilde{z}\tilde{\mu})} \right\}, \tag{3}$$

in which the parameters have been scaled so that

$$\tilde{z} = \frac{z}{\mu}, \quad \tilde{x}_0 = \frac{x_0}{\mu}, \quad \text{and} \quad \tilde{\mu} = \left(\frac{\mu}{\sigma}\right)^2. \tag{4}$$

Given a non-zero initial condition $\tilde{x}_0$, mean DTs are different for correct and error trials:

$$\langle \mathrm{DT}_{\mathrm{correct}} \rangle = \frac{\exp((\tilde{z} - \tilde{x}_0)\tilde{\mu})}{1 - \mathrm{ER}} \times \frac{[(\tilde{z} - \tilde{x}_0)\cosh((\tilde{z} + \tilde{x}_0)\tilde{\mu})\,\sinh(2\tilde{z}\tilde{\mu}) - 2\tilde{z}\sinh((\tilde{z} - \tilde{x}_0)\tilde{\mu})]}{\sinh^2(2\tilde{z}\tilde{\mu})}, \tag{5}$$

$$\langle \mathrm{DT}_{\mathrm{error}} \rangle = \frac{\exp\left(-\left(\tilde{z} + \tilde{x}_0\right)\tilde{\mu}\right)}{\mathrm{ER}} \times \frac{[(\tilde{z} + \tilde{x}_0)\cosh\left((\tilde{z} - \tilde{x}_0)\tilde{\mu}\right)\sinh(2\tilde{z}\tilde{\mu}) - 2\tilde{z}\sinh((\tilde{z} + \tilde{x})\,\tilde{\mu})]}{\sinh^2(2\tilde{z}\tilde{\mu})}. \tag{6}$$

See the Appendix for derivations of equations (5 and 6).

The simplicity and analytical tractability of the DDM is a motivating factor in our decision to use it as a basis for our study. We note that the DDM is much simpler than the leaky competing accumulator (LCA) Model (Usher and McClelland, 2001), which has been used in prior models of sequential effects (Cho et al., 2002; Jones et al., 2002; Gao et al., 2009). LCA processes involve two or more coupled non-linear and stochastic differential equations. We compare the adapted DDM with the LCA-based Cho et al. (2002), Jones et al. (2002), and Gao et al. (2009) models in Section 3.1, using the data of Experiment 1.

2.3.1. Priming mechanism

As with other sequential effects models (e.g., Cho et al., 2002; Jones et al., 2002; Gao et al., 2009), parameters are updated by a priming mechanism to reflect the stimulus history of repetitions and alternations. In the Cho, Jones, and Gao Models, priming is implemented by small history-based changes to the drift parameter, μ. In contrast, in our adapted DDM we update the initial conditions at trial $n+1$ by setting

$$\tilde{x}_0\,(n+1) = \pm k\left(M(n) - \frac{1}{2}\right) \pm \tilde{x}_{\mathrm{offset}}, \tag{7}$$

in which n is the previous trial number, $k > 0$ is a scaling constant, and $M(n)$ serves as a dynamic memory of repetitions and updates at the start of each new trial. $M(n)$ is confined to the interval [0,1], so that $M(n) - 1/2$ ranges from $-1/2$ to 1/2. A symmetry between R and A biases is then enforced: a positive value of $M(n) - 1/2$ corresponds to bias toward R trials and a negative $M(n) - 1/2$ corresponds to bias toward A trials. Moreover, updates to $M(n)$ are defined such that an increase in bias toward R trials will correspond to a decrease in bias toward A trials, and vice versa. Without loss of generality, we define our model terms such that the positive direction for $\tilde{x}_{\mathrm{offset}}$ always corresponds to the correct response. The normalized drift parameter $\tilde{\mu}$ must then always take a positive value, and the sign of the offset bias $\tilde{x}_{\mathrm{offset}}$ and the scaling constant k will vary from trial to trial, with positive coefficients selected if the current trial is a repetition of the previous stimulus and negative coefficients if it is an alternation.

The memory function is updated as follows:

$$M\,(n) = \Delta M\,(n-1) + \begin{cases} 1 - \Delta, & \text{if repetition from } n-1 \text{ to } n, \\ 0, & \text{if alternation from } n-1 \text{ to } n, \end{cases} \tag{8}$$

where $0 < \Delta < 1$. The Δ parameter determines the dependence of behavior on previous trials, with higher values corresponding to the level of influence of trials further back in the sequence and lower values corresponding to dependence on only recent trials. A Δ value of 0.5 corresponds to a memory length of approximately four trials ($\Delta^4 = 0.0625$), after which history dependence goes below 5%. A single update parameter Δ can then account for responses to both R and A trials. In contrast, the Cho, Jones, and Gao models used a memory function $M(n)$ but separately tracked R and A trials. Our model is always initialized with no bias, so that $M(1) = M(2) = 1/2$, after which $M(n)$ updates according to the above expression. This mechanism allows for large adjustments to initial conditions to follow the termination of strings of repetitions or alternations. The updating mechanism is similar to updates to biasing terms proposed in previous work (Cho et al., 2002; Gao et al., 2009), where initial conditions and drift rates are updated.

2.3.2. Error-correcting mechanism

We also employ error-correction threshold modulation. Threshold modulation has been studied in the context of several sequential choice tasks (Bogacz et al., 2006; Simen et al., 2006). In particular, models have used variable thresholds in describing optimal behavior, as well as to account for variability in reaction time. Increased caution is attributed to a higher threshold, which is understood to follow error commission. However, prior models of sequential effects have not included threshold modulation.

In the adapted DDM, the thresholds are adjusted after every trial and constrained to remain symmetric at $\pm\tilde{z}$. After a correct

trial, $\tilde{z}$ is reduced by $\tilde{z}_{\text{down}} > 0$, and after an error trial, increased by $\tilde{z}_{\text{up}} > 0$:

$$\tilde{z}(n) = \tilde{z}(n-1) + \begin{cases} -\tilde{z}_{\text{down}}, & \text{if correct at } n-1, \\ \tilde{z}_{\text{up}} & \text{if error at } n-1. \end{cases} \quad (9)$$

The range of $\tilde{z}(n)$ is constrained so that the thresholds always have a magnitude greater than or equal to the magnitude of the initial conditions, i.e., such that $\tilde{z}(n) \geq \frac{k}{2} + \tilde{x}_{\text{offset}}$; $\tilde{z}(n)$ is also constrained so that $\tilde{z}(n) \leq \tilde{z}_{max}$. The thresholds are initialized conservatively such that $\tilde{z}(1) = \tilde{z}_{max}$. If an update causes $\tilde{z}(n)$ to fall outside its bounds, $\tilde{z}(n)$ is then set to the value of the nearest bound until the next trial.

Sequential, error-correcting variations in the evidence thresholds $\tilde{z}(n)$ can produce significant differences between reaction times for correct and error trials. Trials with lower thresholds have higher ERs and faster RTs; thus, on average, error trials are faster and correct trials slower. This effect is modulated by adjustments to the initial condition $\tilde{x}_0$, which result in faster correct or incorrect responses by biasing the system asymmetrically toward one of the choices. The memory function and initial condition and threshold updates add six parameters to the model: k, $\tilde{x}_{\text{offset}}$, Δ, $\tilde{z}_{\text{down}}$, $\tilde{z}_{\text{up}}$, and $\tilde{z}_{\text{max}}$, in addition to $\tilde{\mu}$ and T_{nd}, for a total of eight parameters.

2.4. MODEL SIMULATION AND DATA FITTING PROCEDURE

Fitted model parameters were used to validate the adapted DDM against data from the two experiments. Separate analysis and fitting was conducted for Experiments 1 and 2. In each case, the data were sorted by sequence, RT, and ER. Model behavior was computed for each parameter set and then sorted similarly. The model was run using the same stimulus sequences that each participant had encountered. Parameters were selected by attempting to minimize the sum of squares error between model prediction and participant data,

$$\text{Err} = \sum_{i}^{N} \left(r_{i,\text{model}} - r_{i,\text{data}}\right)^2, \quad (10)$$

in which the elements r_i include unweighted overall mean RTs for each of the four possible second-order sequences for repeating (R) and alternating (A) stimuli. We considered RR, AR, RA, and AA sequences for correct trials, for error trials, and for trials overall, mean ERs for these sequences, as well as mean RTs before error trials, on error trials, and after error trials. For Experiment 1, r had $N = 19$ elements. For Experiment 2, r had $N = 3 \times 19 = 57$ elements, because data was included for each of the 3 values of P_{A}. Time was considered in units of seconds and ERs in decimal fractions of trials, so that range of parameters for elements of r were comparable.

The search for parameters was conducted using a Trust-Region-Reflective Optimization (TRRO) algorithm (Coleman and Li, 1994, 1996). The function lsqnonlin in Matlab was used with default options to search and select parameters that minimize equation (10). For each parameter set and experimental condition, the model ran at least 5 times through the stimulus sequence that each participant had encountered in a given block of trials. (Thus, if a participant were to see left, then left, then right stimuli, the model was presented with those same stimuli in sequence left-left-right, along with the stimuli preceding and following them, and these entire sequences would be repeated for the model subject at least 5 times.) For each trial the probability of error was computed from equation (3) and from this number the correctness or error of that trial was decided by biased coin flip. The expected correct or error RT for the trial was then obtained from equation (5) or equation (6), and parameter updates were implemented according to equations (7–9). The individual trial results were then sorted and averaged in the same manner as the experimental data, model predictions were inserted into equation (10), and model parameters were updated by the TRRO algorithm. This was repeated until the lsqnonlin convergence criterion was met. To produce the model data plots in Section 3, each model with its best fit parameters was rerun 10 times and the resulting RTs and ERs computed by averaging over these runs.

Use of the analytical expressions of equations (3–6) for expected ERs and RTs substantially speeds up the fitting process, since direct numerical simulations of equation (1) are avoided. The final parameter selections are listed in **Table 1**, and the results and implications of the fitting process are considered in the results and discussion sections of this paper.

A study of the differences between the two tasks can lend some insight into the different fit parameterizations for each of the experiments. We note that the choice tasks presented in Experiment 2 are more challenging than those of Experiment 1, in which stimuli were highly discernable. The difference in signal to noise ratios ($\tilde{\mu}$) in the fits to the two experiments is therefore to be expected. In addition, more difficult tasks generally incur more conservative or cautious behavior in subjects, even when it is not in the subjects' best interests (Bogacz et al., 2006). Increased caution (and consequently higher thresholds in DDM fits) have been shown to correspond to more difficult tasks (e.g., Ratcliff et al., 2000, 2001, 2004). Thus, after correct responses in Experiment 2, model threshold adjustments ($\tilde{z}_{\text{down}}$) are small, whereas in Experiment 1 they are larger, and corrections after errors ($\tilde{z}_{\text{up}}$) are smaller in Experiment 1 than in Experiment 2. Our Δ values are consistent with studies showing stimulus history dependence of up to 4 trials back (e.g., Soetens et al., 1984). The remaining parameters are relatively closer in magnitude for both experiments.

3. RESULTS

In order to better understand the relationship between sequential effects and error effects, data from the two experiments were sorted by stimulus sequence and response correctness and compared with model predictions. We first note several trends from this analysis in the Experiment 1 data. We then analyze data from Experiment 2, and we consider how error and sequential effects are influenced by the relative frequencies of repetition (R) and alternation (A) trials. At the same time, we validate our model fits by comparing them with the data from the two experiments.

In our analysis, we refer to RA and AR sequences as *unexpected sequences*, and RR and AA sequences as *expected sequences*. The RT for an RA sequence is the RT corresponding to the A trial, and for an AR sequence, the RT corresponding to the R trial. We call an *R line* one which connects plotted data for RR and AR, and an *A line* one which connects plotted data for RA and AA. We consider only

the two most recent trials in each sequence in our calculations, as the effects of errors are known to have a limited duration (Rabbitt, 1966); moreover, for the strongly biased stimuli ($P_A = 10, 90\%$) of Experiment 2, longer sequences of A's, respectively, R's, occur too rarely to yield sufficient data. The degrees of freedom for the F-tests were Greenhouse-Geisser adjusted for all reported main effects and interactions in which there were significant violations of sphericity.

3.1. EXPERIMENT 1: ERROR DYNAMICS IN UNBIASED TASKS

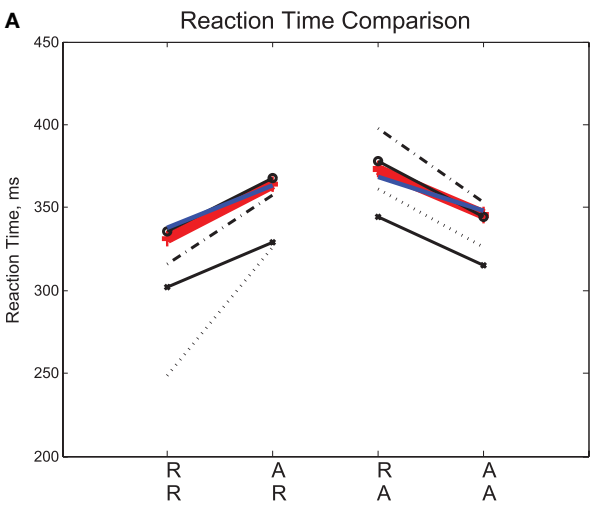
We consider sequential effects and error effects in Experiment 1 data (referred to as Cho Data), in which R and A trials were equally likely, and as has been customary, we initially average over all responses, correct and incorrect. We first discuss overall sequential effects in RT and ER, as shown in **Figure 2**. As expected, we find the smallest mean RT and ER for expected trials (RR, AA), and the largest mean RT and ER for unexpected trials (AR, RA). The effects of sequence on RT [$F(3,15) = 14.81, p < 0.001, \eta^2 = 0.13$] and ER [$F(3,15) = 8.80, p < 0.01, \eta^2 = 0.25$] were significant in two one-way, within-groups ANOVAs. We consider also three published, generative models of the data in Experiment 1, which we refer to as the Cho et al. (2002), Jones et al. (2002), and Gao et al. (2009) Models, respectively. These models were designed to account for these basic sequential effects, and we note that they, as well as the adapted drift diffusion model (DDM) described in Section 2.3, account for trends in mean RTs and ERs.

We next consider the data separated into correct and error trials, shown in **Figure 3A** as solid and dotted lines, respectively. Splitting the data in this way reveals a separation in mean RT for correct and error responses that is greatest for unexpected trials (AR, RA) and least for expected ones (RR, AA). For unexpected trials, error responses are fast, and correct responses are slow. A two-way within-groups ANOVA shows the effect of response correctness is significant [$F(1,5) = 113.93, p < 0.001, \eta^2 = 0.35$)], along with the interaction of response correctness and expectedness of a stimulus [$F(1,5) = 16.51, p < 0.01, \eta^2 = 0.19$]. We note a slight asymmetry in the responses such that RTs for error and correct trials are closer for the R lines than for the A lines. **Figures 3B,C** illustrate the results of the Cho and Jones Models, respectively, and **Figures 3D,E** those of the two versions of the Gao Model. While all four of these models capture the trends in RT for correct trials, none of them predicts the qualitative patterns or quantitative results for error trials. Since the ERs are generally low, RTs averaged over both correct and error trials are close to the mean RTs for the correct trials alone, and this failure of the models becomes apparent only when error trials are considered separately (cf. **Figure 2**, which displays fairly good fits, and see Cho et al., 2002; Jones et al., 2002; Gao et al., 2009). This analysis also reveals that the errors, while fast on average, are not uniformly so, being significantly faster for unexpected sequences (AR, RA). Moreover, as shown in **Figure 3F**, the adapted DDM accounts for all the RT data.

Strikingly, we note that when plotted against each other as in **Figure 4**, RTs for correct and error trials for the sequences RR, AR, RA, and AA display a monotonic and nearly linear relationship, which we call the *sequential RT tradeoff*. As we shall see, such a tradeoff also holds for Experiment 2. In **Figure 4** we show the

Table 1 | DDM parameterization.

	$\tilde{\mu}$	T_{nd}	K	$\tilde{x}_{0,offset}$	Δ	$\tilde{z}_{down}$	$\tilde{z}_{up}$	$\tilde{z}_{max}$
Experiment 1	38.1747	0.2626	0.0943	0.0051	0.6860	0.0058	0.0348	0.2857
Experiment 2	19.3312	0.3359	0.1181	0.0034	0.6882	0.0034	0.1635	0.2062



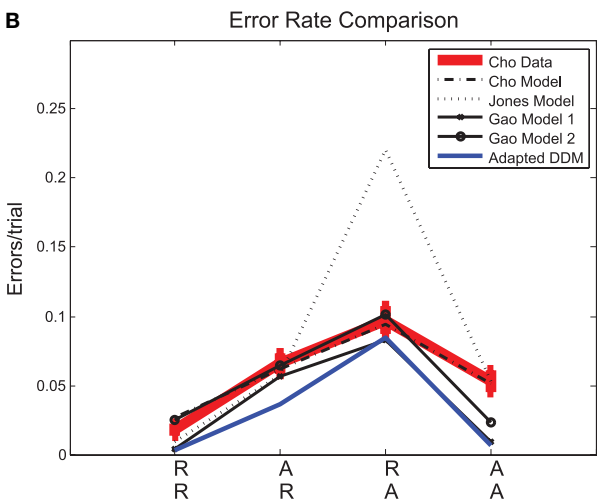


FIGURE 2 | Mean (A) RTs and (B) ERs for the data in Experiment 1 (Cho Data), the three published fits to the data (Cho, Jones, Gao), and the fit presented for this data in the adapted DDM of the present study. In the diagram, an RR sequence refers to the RT on the second repetition of a stimulus (e.g., left, left, *left*) and an AR sequence refers to the RT on the first repetition of a stimulus (e.g., left, right, *right*), etc. The adapted DDM provides the best fit to RTs, but underestimates errors, especially for AA sequences.

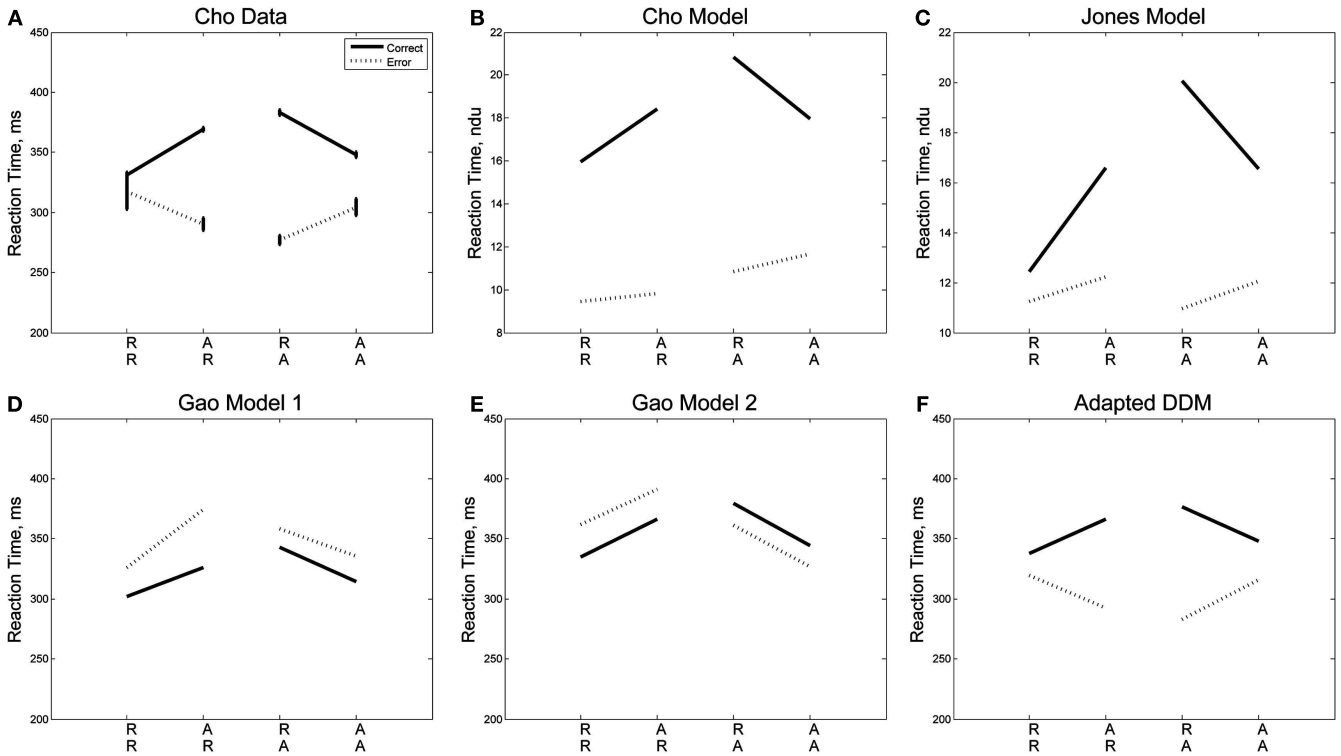


FIGURE 3 | Sequential effects in Experiment 1: (A) data, (B) results of the Cho Model (Cho et al., 2002), (C) results of the Jones Model (Jones et al., 2002), (D,E) results of the Gao Models (Gao et al., 2009), and (F) results of the adapted DDM. The Cho and Jones models predict a dimensionless reaction time, which we give here in non-dimensional units **(B,C)**. The adapted DDM captures the slopes of the R and A curves for error and correct trials. In these figures, the correctness or lack thereof of a given trial corresponds only to that trial itself, so a left-right-left sequence is tabulated as correct for the final left stimulus if and only if the final trial were identified correctly as a left stimulus.

data from Experiment 1 ($R^2 = 0.995$, $p < 0.01$) and the adapted DDM, and from a separate study by Jentzsch and Sommer, 2002; $R^2 = 0.96, p < 0.05$). The area of the circles are proportional to the ERs for the given sequences. We note that the smallest ERs correspond to sequences with relatively fast correct responses and slow errors, while the high ERs occur with relatively fast errors and slow correct responses. While the overall ordering of the sequences (RR, AR, RA, AA) in the tradeoff differs between the two experimental studies, in both cases the points corresponding to unexpected trials (AR, RA) lie at the upper left, and those corresponding to expected trials (RR, AA) lie at the lower right.

The ordering of the tradeoffs is influenced by the nature of the task. However, in each task we see that an increase in time to respond correctly (or a bias toward the correct response) is correlated with a decrease in time to respond in error, and vice versa. Our proposed biasing mechanism achieves a similar effect.

Finally, we consider the RTs before, during, and after an error in Experiment 1, as shown in **Figure 5**. Mean RTs for trials immediately following an error are longer than both those for the error trial itself and for the trial immediately before the error. A one-way within-groups ANOVA confirms that this effect on RT is significant [$F(2,10) = 16.37$, $p < 0.001$, $\eta^2 = 0.48$]. We again compare the behavior with the adapted DDM and the three previous models. In the Cho Model, the RT after an error is slower than the RT on the error trial but faster than the trial immediately prior to the error. The Jones Model maintains the trends in the data but parameter values are skewed so that the range of RTs is larger. In the two Gao Models, mean RTs for trials immediately preceding

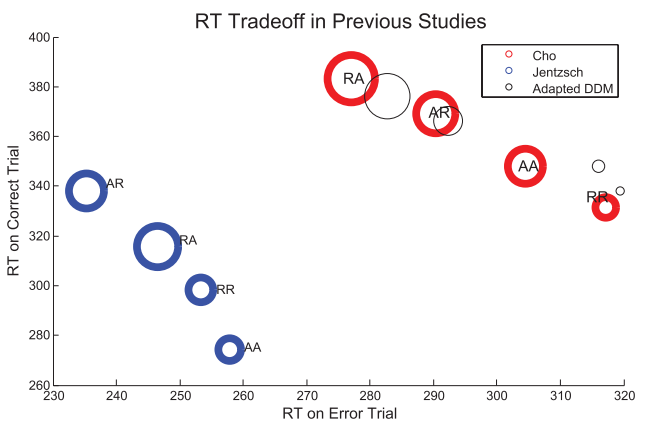


FIGURE 4 | Sequential RT tradeoff for unbiased tasks: a slower RT for correct trials corresponds to a faster RT for error trials for the sequences RR, AR, RA, and AA. The RT tradeoff for Experiment 1 is shown in red. Also shown, in blue: the RT tradeoff from a prior study by Jentzsch and Sommer (2002). Adapted DDM fits to Experiment 1 data are shown in black. The areas of the circles are proportional to the ERs.

and following an error are faster than those on the error trial itself: opposite to the data. The adapted DDM provides the best fit, with the RTs for error trials and post-error trials closely matching the data, although it underestimates RTs on the pre-error trial.

We compare the adapted DDM with the other models using the Akaike Information Criterion (AIC; Akaike, 1974; Stone, 1979), corrected AIC (AIC_c; Hurvich and Tsai, 1989; Burnham and Anderson, 2002), and Bayesian Information Criterion (BIC; Akaike, 1980; Smith and Spiegelhalter, 1980), which provide model fit comparisons that account for the number of parameters included in each model. Scores for the different model fits are shown in **Table 2**. The adapted DDM receives the best overall and relative scores on all three metrics, confirming the fit qualities shown in **Figures 2–5**. AIC_c values cannot be computed for the Gao model, because the number of means being compared is too close to the number of parameters used in the model itself.

3.2. EXPERIMENT 2: INFLUENCE OF STIMULUS ALTERNATION FREQUENCY

We now consider the role that alternation frequency plays in sequential and error effects. We first address overall trends in RT and ER, as shown in **Figures 6A,B**, respectively, following the convention in the sequential effects literature (e.g., Soetens et al., 1985; Jentzsch and Sommer, 2001; Cho et al., 2002). Trends for the $P_A = 50\%$ blocks match trends from Experiment 1 with longer RTs and higher ERs for unexpected trials, and shorter RTs and lower ERs for expected trials. Trends for the $P_A = 10\%$ blocks and $P_A = 90\%$ blocks are clearly distinguishable from the trends for $P_A = 50\%$ blocks, notably in the magnitudes of the slopes of R and A lines. Further, there is an approximate symmetry between the $P_A = 10\%$ case and the $P_A = 90\%$ case.

Sequential effects in mean RTs are clearly influenced by the probability of alternations, with respect to both overall mean RTs and ERs (**Figures 6A,B**). Mean RTs for unexpected sequences (AR, RA) remain similar for all P_A conditions but there are significant differences in mean RTs for expected sequences (RR, AA). For the highest P_A, RT is faster on AA trials than the corresponding sequence RTs for lower P_As, and for the lowest P_A, the RT is faster on RR trials than the corresponding sequence RTs for higher P_As. As expected, we find that the effects of sequence [$F(3,42) = 50.62$, $p < 0.001$, $\eta^2 = 0.26$] and its interaction with P_A [$F(3.36, 47.04) = 43.09$, $p < 0.001$, $\eta^2 = 0.26$] on RT are both significant. Error rates are greatest for AR trials at the highest P_A and RA trials at the lowest P_A. The effects of sequence [$F(1.95,$

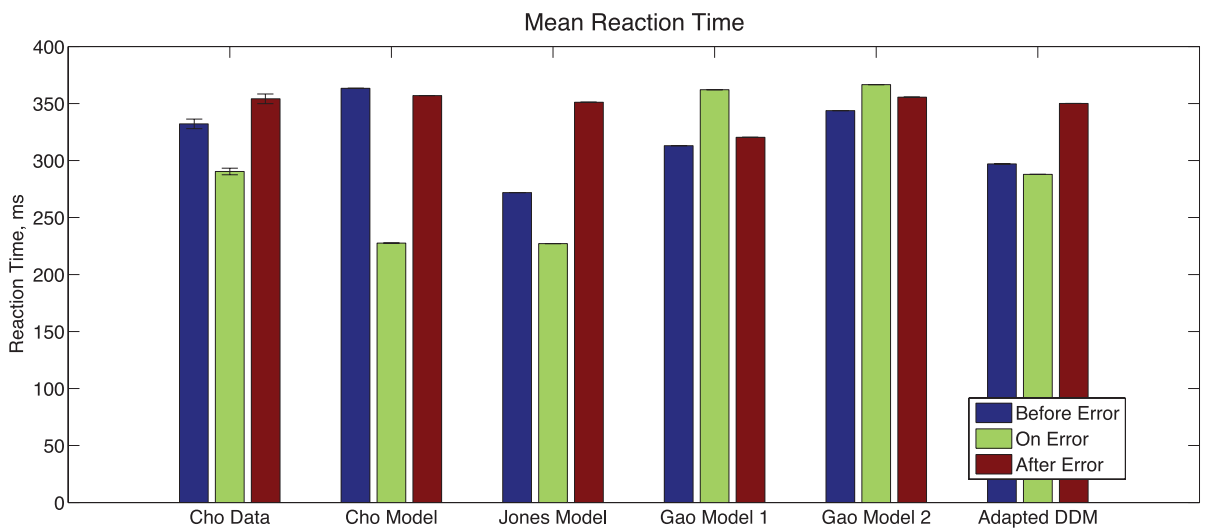


FIGURE 5 | Post-error slowing in Experiment 1 data and in the models of Cho et al. (2002), Jones et al. (2002), Gao (first model; Gao et al., 2009), and the adapted DDM of the present study. In the Experiment 1 data, the mean RT for a trial immediately following the error trial is slower than that for the trial before the error, and the mean RT for the error trial itself is fast. The Cho and Gao models fail to account for both trends. The Jones Model accounts for the proper trends but overestimates the magnitude of the post-error slowing. The adapted DDM accounts for both trends but underestimates post-error RTs.

Table 2 | Model performance comparison.

Model	Total parameters	R^2	AIC	AIC_c	BIC
Adapted DDM	8	0.996	84.7	115.1	108.2
Cho	13	0.936	138.2	237.0	176.5
Gao 1	18	0.915	140.4	–	193.4
Gao 2	18	0.951	137.0	–	190.0
Jones	16	0.943	146.9	450.9	194.1

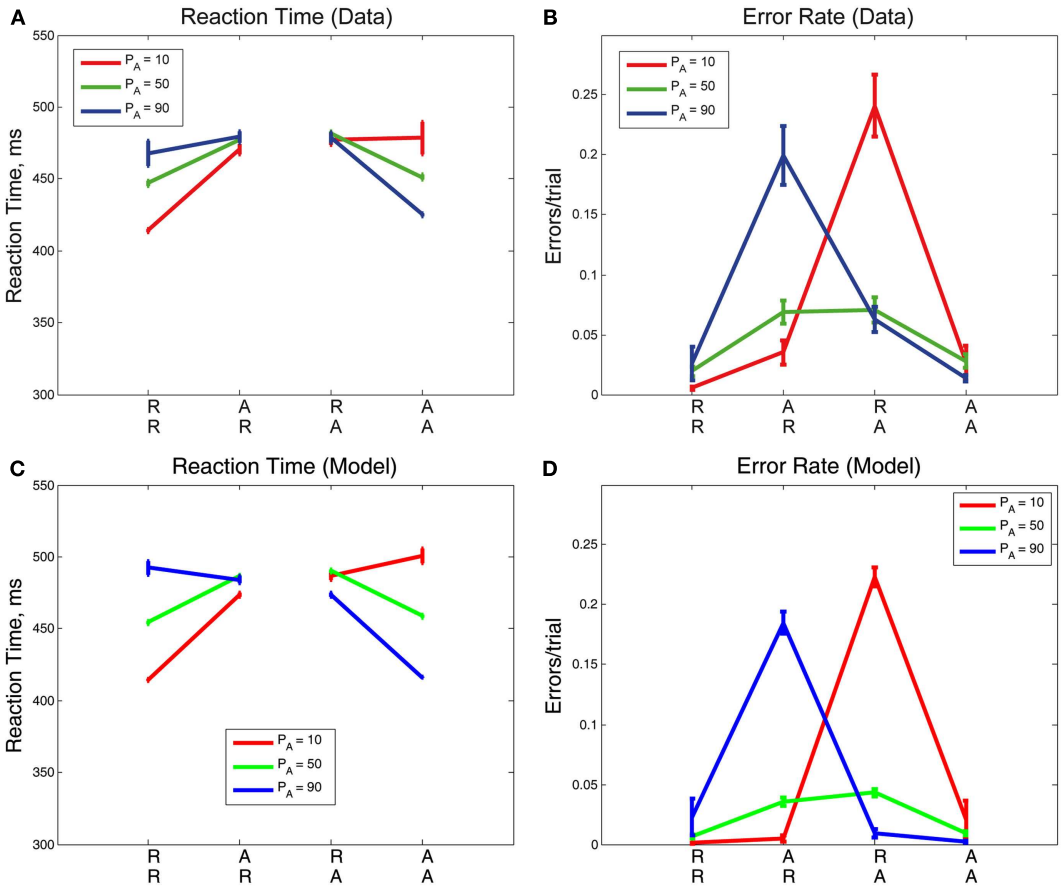


FIGURE 6 | Mean (A) RTs and (B) ERs for the three values of P_A in Experiment 2, averaged over correct and error trials. The influence of P_A is most apparent in the mean RT plot on expected trials (RR, AA) and in the mean ER on unexpected trials (AR, RA). Model fits for **(C)** RTs and **(D)** ERs re-create behavioral trends in RTs and ERs but overestimate RTs for expected trials (RR, AA). The error bars in plots **(A,B)** represent the standard error of the mean, and in **(C,D)** the average value of standard error of the mean over 10 simulation runs (see Section 2.4 for details).

$27.30) = 20.86$, $p < 0.001$, $\eta^2 = 0.31$] and its interaction with P_A [$F(2.46, 34.44) = 20.19$, $p < 0.001$, $\eta^2 = 0.32$] on ER are also both significant. The adapted DDM reproduces the qualitative patterns in the data, but overestimates RTs for expected sequences when their probabilities are low (RR, with $P_A = 90\%$; AA, with $P_A = 10\%$), and underestimates ERs for unexpected sequences (AR, RA): **Figures 6C,D**.

We also found that the overall sequential effects are influenced by the probability of alternations. The relationship between the time to respond to sequences ending in R versus A on the final sequence is known to indicate relative preference for R or A trials (Audley, 1973). Prior work has shown that preference for A trials varies with RSI, but the role of the likelihood of A trials in determining the relative preference for A has not been studied.

The green lines corresponding to $P_A = 50\%$ in **Figures 7A,C** show no preference for R or A: expected sequences (RR, AA) yield faster RTs symmetrically in R and A than unexpected sequences (AR, RA). The red $P_A = 10\%$ blocks show a strong preference for R: the mean RT after an R is faster in the case of RR than it is for AR, whereas the RT after A is similar for both RA and AA. The blue lines corresponding to $P_A = 90\%$ show a strong preference for A: the RT after an A is faster in the case of AA than it is for RA, whereas the RT after an R is similar for both RR and AR. For $P_A = 10\%$, the model predicts, as in the data, that the repetition RT is faster for RR than it is for AR, but the model predicts a slower alternation RT for AA than for RA, and it shows a symmetric trend for $P_A = 90\%$. In summary, both data and model exhibit increases in preference for A with increased probability of alternations, showing that relative preferences for R or A trials can be influenced by transition probabilities in addition to task properties such as RSI.

In **Figures 7A–C** we replot the mean RT data, separated into correct and incorrect responses, thus revealing differing sequential effects for each P_A. A two-way within-groups ANOVA shows that the effects of correctness [$F(1,14) = 249.64$, $p < 0.001$, $\eta^2 = 0.80$], whether or not the trial was expected [$F(1,14) = 54.70$, $p < 0.001$, $\eta^2 = 0.44$], and the interaction of these two factors [$F(1,14) = 88.38$, $p < 0.001$, $\eta^2 = 0.62$] are all significant. For unbiased sequences ($P_A = 50\%$), sequential effects are again similar to those for correct and error trials in Experiment 1 (**Figure 7B**, cf. **Figure 3A**). For both low and high P_A blocks, the orientations of the R and A lines are maintained, with correct R lines sloping upwards from RR to AR and correct A lines sloping down from RA

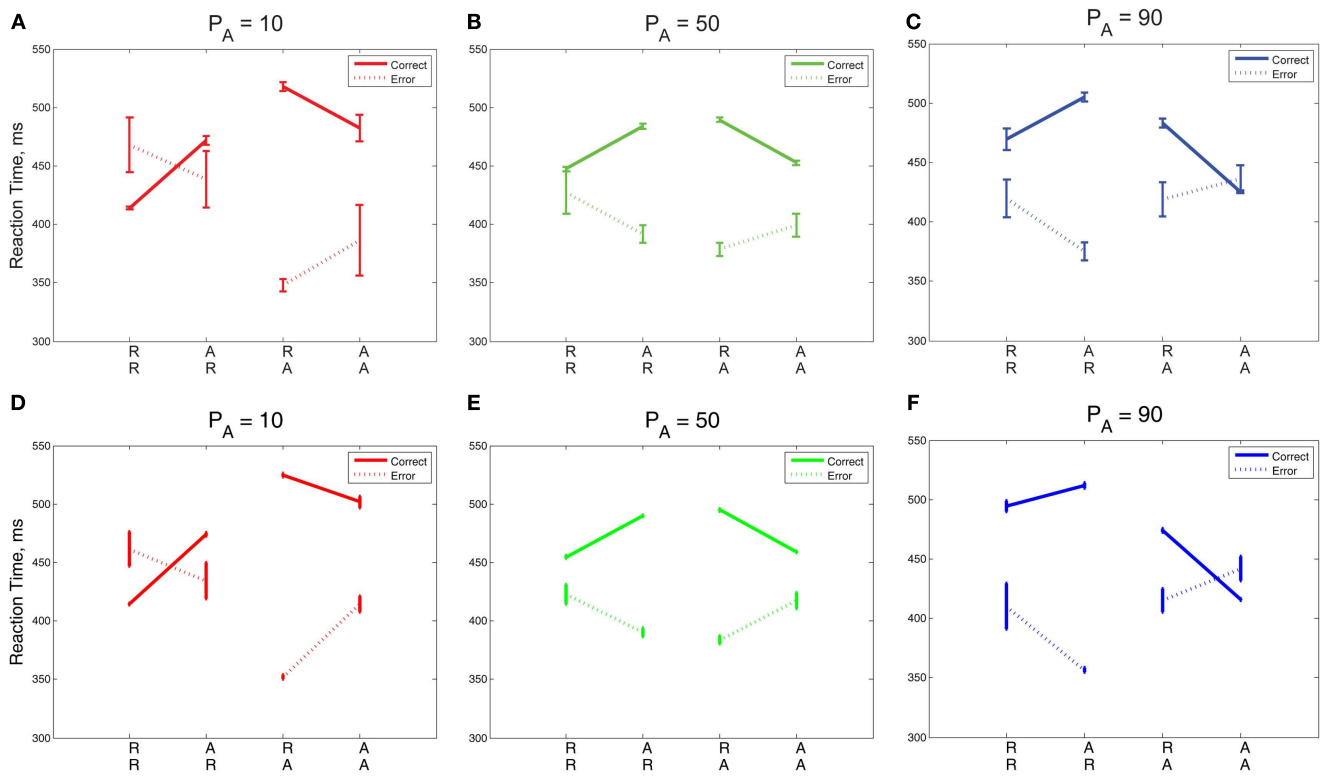


FIGURE 7 | (A–C) RT data for error and correct trials in Experiment 2 compared with the adapted DDM **(D–F)**. The slopes of the R and A lines are reversed for correct and incorrect responses. Error trials incur faster responses on unexpected trials (RA, AR) than on expected trials (RR, AA); this trend is reversed for correct responses. For low P_A, the R lines overlap, and for high P_A, the A lines overlap, resulting in an approximate reflectional symmetry between data for the high and low P_A blocks, so that sometimes the mean time for an error trial is slower than for a correct trial. The error bars in plots **(A–C)** represent the standard error of the mean, and in **(D–F)** the average value of standard error of the mean over 10 simulation runs (see Section 2.4 for details).

to AA. For $P_A = 50\%$, the slopes of the R and A lines for incorrect responses are nearly opposite the slopes of the R and A lines for correct responses. For $P_A = 10\%$ blocks, the R lines cross and the A lines are further apart than in the $P_A = 50\%$ blocks. For $P_A = 90\%$ blocks, we see a mirrored trend, in which the A lines cross and the R lines are further apart than in the $P_A = 50\%$ block. We also note a striking asymmetry for the biased stimuli: for $P_A = 10\%$ R lines are, on average, closer together than A lines, and for $P_A = 90\%$ this relationship is mirrored, so that A lines are closer than R lines. However, the mirroring is not perfect: the degree of overlap in R lines is greater for $P_A = 10\%$ than the corresponding overlap in A lines for $P_A = 90\%$.

The trends in correct and error trial RTs, including the crossover of the R and A lines, are generally captured by the adapted DDM, as shown in **Figures 7D–F**. However, the steepness of slope of the R (respectively, A) lines are underestimated for correct trials for $P_A = 90\%$ (10%), due to overestimation of the RR (AA) RTs.

Next, we note that the sequential RT tradeoff between correct and error responses is also observed in Experiment 2, as shown in **Figure 8A**. As in **Figure 4**, the areas of the circles are proportional to the corresponding ERs. The relationship between RTs for correct and error trials for each of the sequences RR, AR, RA, and AA is monotone (and nearly linear) for all points shown in the figure ($R^2 = 0.75$, $p < 0.001$), and this correlation is also captured by our model ($R^2 = 0.74$, $p < 0.001$). The sequences with the largest ER have relatively fast RTs for errors and relatively slow RTs for correct trials. Note, however, that data for individual P_A's of 10, 50, and 90% is not quite as strongly correlated. Differences in order can be expected because the sequential effects for each probability of alternation are influenced by the probability of alternation.

Finally, we note that post-error slowing occurs for all P_A blocks with the same trend: the error trial itself incurs a slightly faster RT than the trial which precedes it, and the post-error trial incurs an RT significantly slower than RTs for the preceding two trials, as shown in **Figure 9**. A two-way within-groups ANOVA indicates that the effects of time before, upon, and after an error commission [$F(1.36, 19.04) = 68.25$, $p < 0.001$, $\eta^2 = 0.57$] and on P_A [$F(2,28) = 5.83$, $p < 0.01$, $\eta^2 = 0.07$] are both significant, but the effect of their interaction is not significant. Thus, in Experiment 2, pre- and post-error RTs share the pattern of RTs in Experiment 1, and this pattern is preserved over all three values of P_A. The bottom panel shows that our model both qualitatively and quantitatively captures the post-error slowing in Experiment 2. However, as in Experiment 1 (**Figure 5**), the model fails to produce the observed speed-up on the error trial itself.

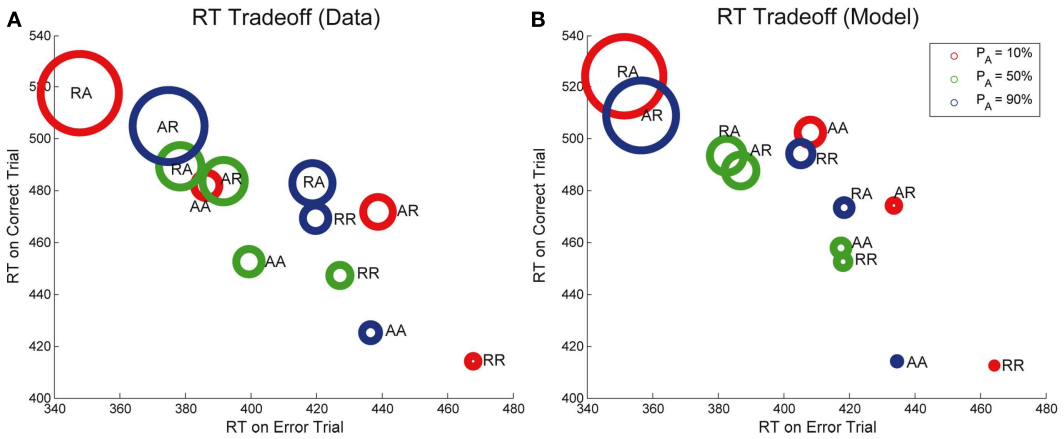


FIGURE 8 | Sequential RT tradeoff for Experiment 2. Mean RTs for correct trials are strongly correlated with mean RTs for error trials for each of the sequences RR, AR, RA, AA, for each value of P_A for both **(A)** data and **(B)** adapted DDM.

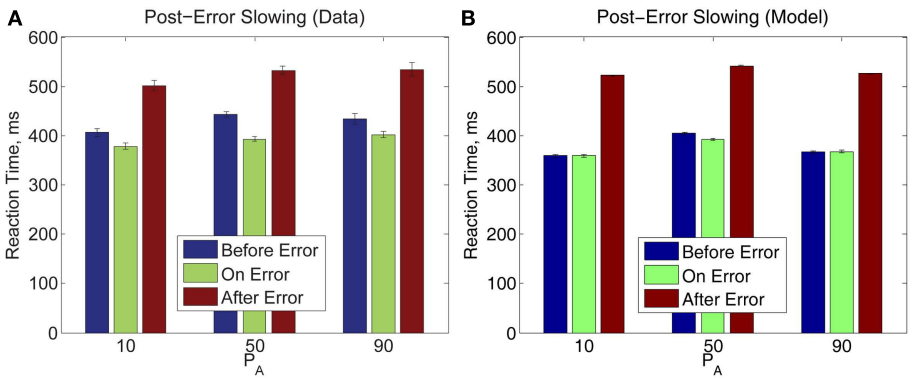


FIGURE 9 | (A) Post-error slowing in Experiment 2 data is independent of P_A. **(B)** The model fit also predicts post-error slowing but does not fully account for pre-error speeding. The error bars in plot **(A)** represent the standard error of the mean, and in **(B)** the average value of standard error of the mean over 10 simulation runs (see Section 2.4 for details).

4. DISCUSSION

In this paper, we propose priming and error-correcting mechanisms to account for sequential effects and post-error slowing, respectively. Each mechanism, on its own, is commonplace in models of decision making. Indeed, various priming mechanisms have been previously proposed to account for sequential effects (Cho et al., 2002; Jones et al., 2002; Gao et al., 2009). Post-error slowing is also known to occur and exert a significant influence on RT patterns (Rabbitt, 1966, 1968a; Laming, 1979b). The implementation of post-error slowing is understood to be a simple one: in an accumulator model, the response thresholds can be raised following an error to increase the necessary processing time before a decision is reached (Rabbitt and Rodgers, 1977; Rabbitt and Vyas, 1981; Brewer and Smith, 1984; Jentzsch and Dudschig, 2009). To the best of our knowledge, no prior model of sequential effects has explicitly incorporated such an error-correcting mechanism to also account for post-error slowing. We consider sequential effects for both high and low probabilities of alternations, a consideration unique to this paper: previously, sequential effects for sequences of alternating and repeating stimuli had been studied only for stimuli in which the probabilities of alternations and repetitions were equal.

Our model is informed by previous work: the initial conditions are varied according to a priming function similar to those in other models (Cho et al., 2002; Jones et al., 2002; Gao et al., 2009), and the thresholds are raised after incorrect responses and lowered after correct ones (Simen et al., 2006). Variability in thresholds of drift diffusion processes during a trial can result in fast errors (Ratcliff and Rouder, 1998). Our implementation, however, is unique: we use both priming and error-correcting mechanisms in the same model. In doing so, we can account for many of the observed trends in behavior.

Our adaptation of the pure drift diffusion model has multiple advantages. The pure DDM is analytically simple, and explicit expressions exist for both RT distributions and accuracy, and separate and closed-form expressions for mean RTs can be derived for correct and error responses, as shown in the Appendix. With non-zero initial conditions, the pure DDM can also account for

RT distributions for correct and error trials. Moreover, the priming and error-correction mechanisms that we have proposed are conceptually straightforward. With the error-correction mechanism, our model accounts for post-error slowing: the RT for the trial which immediately follows an error trial is not only significantly slower than the error trial but also slower than the RT for the trial immediately preceding the error. We show that when thresholds are systematically adjusted to account for error and correct responses and priming is implemented, sequential patterns in error and correct response trial RTs emerge and are consistent with participant behavior, as shown in **Figures 5** and **9**.

Indeed, our adapted DDM predicts the characteristic trends in mean RTs for sequences ending in correct or incorrect responses whereas several other models do not. We show experimentally and for the first time that unexpected trials (AR or RA) result in relatively slow correct responses and fast errors, whereas expected trials (RR or AA) result in relatively fast correct responses and slow errors as shown again in **Figures 3** and **7**. Our model captures aspects of this behavior with the incorporation of post-error adjustments to the model thresholds: priming accounts for the sequential patterns in RT for correct trials, and error-correction accounts for the patterns in RT for the error trials.

The relationship between RTs for correct and error trials is central to our model: biasing the initial conditions toward expected sequences automatically biases them *against* unexpected sequences. Subjects biased against an unexpected stimulus will then respond to it slowly if they are to respond correctly, and rapidly if they are to respond in error. In contrast, in previous work (Cho et al., 2002; Jones et al., 2002; Gao et al., 2009), the biasing was instead applied to sensitivity to stimulus, so that the relationship between RT for error and correct trials was less direct. Moreover, when biasing is coupled with explicit post-error adjustments, further nuances in the relationship between mean time to respond correctly versus in error may be realized.

Significantly, we also identify a sequential RT tradeoff, in which the correlation between the mean RTs for error and correct trials for each of the sequences (RR, AR, RA, AA) is quite strong: a faster RT on an error response corresponds to a slower RT on a correct response. The correlation between mean RTs for correct and error trials is captured by our model, as shown in **Figures 4** and **8**.

We then show that sequential effects in mean RTs overall, as well as in mean RTs for correct and error trials, are significantly influenced by the probability of alternations. Our data reveals remarkable near-mirror-symmetry between RT patterns for alternations when the probability of alternations is low and repetitions when the probability of alternations is high: incorrect responses are fast and correct responses are significantly slower. Sequential effects in ER also vary with the probability of alternations. Our model captures this near-symmetry in **Figures 6** and **7**.

Moreover, we have shown, both in our data and in our model, that an increase in the likelihood of alternations corresponds to an increase in relative preference for alternations. This can be inferred from the RT versus sequence plots in **Figure 6A**. The change in alternation preference with changing likelihood of alternations suggests that choice behavior can be informed and even manipulated by the probabilistic structure of the environment.

The sequential effects in RT and ER for various probabilities of alternation are of particular interest due to their relevance to prior physiological studies. In particular, previous work has shown that the anterior cingulate cortex (ACC) is sensitive to alternations in a sequence of stimuli and identified corresponding neural signals (e.g., Botvinick et al., 2001). Prior models of sequential effects, such as those of Jones et al. (2002) and Gao et al. (2009) have included a "conflict" signal informed by activity in the ACC, and the signal increases in strength with frequent alternations. However, the near-symmetry of behavior at high and low probabilities of alternations in our data suggests a comparable sensitivity to repetitions and alternations, rather than to alternations alone. Indeed, prior work has suggested that the role of the conflict signal in trials with long RSI, such as those considered in this paper, is a minor one (Jones et al., 2002; Jentzsch and Leuthold, 2005) and secondary to that of explicit error correction. Jones et al. (2002) found that the incorporation of a conflict signal in their model resulted in a small but significant improvement in model fit. For short RSI, however, the role of response conflict is more significant (Jentzsch et al., 2007; Jentzsch and Dudschig, 2009). Future work could further clarify the respective roles of response caution (thresholds) and response conflict (ACC) co-varying RSIs and probabilities of alternation.

Additional directions for future work include a consideration of alternative error-correction and priming mechanisms. For example, the magnitude of adjustments made due to our priming mechanism varies from trial to trial, while adjustments from the error-correction mechanism are consistent. Alternate models in which different update schemes are employed are worthy of consideration. Such a study could allow for further model simplification and provide a stronger account of behavior in choice tasks. Moreover, sufficient data should be gathered so that sequential and error effects can be studied and described for individual participants, by fitting RT distributions for different stimulus sequences and individual participants. Finally, a consideration of human behavior in more difficult tasks, such as those with low or variable stimulus discriminability, or tasks in which the probability of alternations varies during blocks of trials, can build upon our work.

In this paper, we have presented a neurally plausible and conceptually straightforward account of sequential effects and post-error slowing by developing a simple repetition-based priming mechanism, coupled with an error-correction mechanism. We implemented these mechanisms within the context of a pure DDM, so the behavior can be described analytically and in closed form. Despite its simplicity, our implementation of the DDM accounts for nuances in behavior which are not found in previous models. In particular, we identified in our data, and our model accounted for, sequential effects for correct and error trials, as well as for trials during blocks with high and low probabilities of alternations. This suggests that an error-correction process, such as a simple adjustment of response thresholds after each trial, plays an instrumental role in sequential patterns in RT. Future work may further clarify the implementation of the error-correction process and its implications for perceptual decision making tasks.

ACKNOWLEDGMENTS

This research was partially supported by the Air Force Office of Scientific Research under grants FA 9550-07-1-0528 and FA 9550-07-1-0537, Multi-disciplinary University Research Initiatives. The authors would like to thank Fuat Balci, Jonathan D. Cohen, Juan Gao, Patrick Simen, Marieke van Vugt, and Robert C. Wilson for helpful comments, as well as the reviewers of this paper and Raymond Cho, Leigh Nystrom, and Ines Jentzsch for sharing their data. Stephanie Goldfarb is supported by a National Science Foundation Graduate Research Fellowship and was previously supported by a National Defense Science and Engineering Graduate Fellowship. Michael Schwemmer is currently supported by NIH grant T32-MH065214-1 through the Princeton Neuroscience Institute.

REFERENCES

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Trans. Automat. Contr.* 19, 716–723.

Akaike, H. (1980). Likelihood and the Bayes procedure. *Trab. Estad. Investig. Oper.* 31, 143–166.

Audley, R. (1973). "Some observations on theories of choice reaction time: tutorial review," in *Attention and Performance IV*, ed. S. Kornblum (New York: Academic Press), 509–545.

Balci, F., Simen, P., Niyogi, R., Saxe, A., Hughes, J., Holmes, P., and Cohen, J. (2011). Acquisition of decision making criteria: reward rate ultimately beats accuracy. *Atten. Percept. Psychophys.* 73, 640–657.

Bertelson, P. (1961). Sequential redundancy and speed in a serial two-choice responding task. *Q. J. Exp. Psychol.* 13, 90–102.

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., and Cohen, J. (2006). The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. *Psychol. Rev.* 113, 700–765.

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

Brainard, D. (1997). The psychophysics toolbox. *Spat. Vis.* 10, 433–436.

Brewer, N., and Smith, G. (1984). How normal and retarded individuals monitor and regulate speed and accuracy of responding in serial choice tasks. *J. Exp. Psychol. Gen.* 113, 71.

Britten, K., Shadlen, M., Newsome, W., and Movshon, J. (1992). The analysis of visual motion: a comparison of neuronal and psychophysical performance. *J. Neurosci.* 12, 4745–4765.

Brodersen, K., Penny, W., Harrison, L., Daunizeau, J., Ruff, C., Duzel, E., Friston, K., and Stephan, K. (2008). Integrated bayesian models of learning and decision making for saccadic eye movements. *Neural Netw.* 21, 1247–1260.

Burnham, K., and Anderson, D. (2002). *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach.* New York: Springer Verlag.

Busemeyer, J., and Townsend, J. (1992). Fundamental derivations from decision field theory. *Math. Soc. Sci.* 23, 255–282.

Capaldi, E. (1966). Partial reinforcement: a hypothesis of sequential effects. *Psychol. Rev.* 73, 459.

Carpenter, R., and Williams, M. (1995). Neural computation of log likelihood in control of saccadic eye movements. *Nature* 377, 59–62.

Cho, R., Nystrom, L., Brown, E., Jones, A., Braver, T., Holmes, P., and Cohen, J. (2002). Mechanisms underlying dependencies of performance on stimulus history in a two-alternative forced-choice task. *Cogn. Affect. Behav. Neurosci.* 2, 283–299.

Coleman, T., and Li, Y. (1994). On the convergence of interior-reflective Newton methods for nonlinear minimization subject to bounds. *Math. Program.* 67, 189–224.

Coleman, T., and Li, Y. (1996). An interior trust region approach for minimization subject to bounds. *SIAM J. Optim.* 6, 418–445.

Corrado, G., Sugrue, L., Seung, H., and Newsome, W. (2005). Linear-nonlinear-poisson models of primate choice dynamics. *J. Exp. Anal. Behav.* 84, 581.

Dudschig, C., and Jentzsch, I. (2009). Speeding before and slowing after errors: is it all just strategy? *Brain Res.* 1296, 56–62.

Feller, W. (1968). *An Introduction to Probability Theory and its Applications.* New York: Wiley.

Feng, S., Holmes, P., Rorie, A., and Newsome, W. (2009). Can monkeys choose optimally when faced with noisy stimuli and unequal rewards? *PLoS Comput. Biol.* 5, e1000284. doi:10.1371/journal.pcbi.1000284

Frensch, P., and Miner, C. (1994). Effects of presentation rate and individual differences in short-term memory capacity on an indirect measure of serial learning. *Mem. Cognit.* 22, 95–110.

Gao, J., Tortell, R., and McClelland, J. (2011). Dynamic integration of reward and stimulus information in perceptual decision-making. *PLoS ONE* 6, e16749. doi:10.1371/journal.pone.0016749

Gao, J., Wong-Lin, K., Holmes, P., Simen, P., and Cohen, J. (2009). Sequential effects in two-choice reaction time tasks: decomposition and synthesis of mechanisms. *Neural. Comput.* 21, 2407–2436.

Gardiner, C. (1985). *Handbook of Stochastic Methods for Physics, Chemistry, and the Natural Sciences.* Berlin: Springer Verlag.

Gold, J., Law, C., Connolly, P., and Bennur, S. (2008). The relative influences of priors and sensory evidence on an oculomotor decision variable during perceptual learning. *J. Neurophysiol.* 100, 2653.

Gold, J., and Shadlen, M. (2000). Representation of a perceptual decision in developing oculomotor commands. *Nature* 404, 390–394.

Huettel, S., Song, A., and McCarthy, G. (2005). Decisions under uncertainty: probabilistic context influences activation of prefrontal and parietal cortices. *J. Neurosci.* 25, 3304–3311.

Hurvich, C., and Tsai, C. (1989). Regression and time series model selection in small samples. *Biometrika* 76, 297–307.

Jentzsch, I., and Dudschig, C. (2009). Why do we slow down after an error? mechanisms underlying the effects of posterror slowing. *Q. J. Exp. Psychol.* 62, 209–218.

Jentzsch, I., and Leuthold, H. (2005). Response conflict determines sequential effects in serial response time tasks with short response-stimulus intervals. *J. Exp. Psychol. Hum. Percept. Perform.* 31, 731–748.

Jentzsch, I., Leuthold, H., and Ulrich, R. (2007). Decomposing sources of response slowing in the prp paradigm. *J. Exp. Psychol. Hum. Percept. Perform.* 33, 610.

Jentzsch, I., and Sommer, W. (2001). Sequence-sensitive subcomponents of P300: topographical analyses and dipole source localization. *Psychophysiology* 38, 607–621.

Jentzsch, I., and Sommer, W. (2002). Functional localization and mechanisms of sequential effects in serial reaction time tasks. *Percept. Psychophys.* 64, 1169.

Jones, A., Cho, R., Nystrom, L., Cohen, J., and Braver, T. (2002). A computational model of anterior cingulate function in speeded response tasks: effects of frequency, sequence, and conflict. *Cogn. Affect. Behav. Neurosci.* 2, 300–317.

Kirby, N. (1976). Sequential effects in two-choice reaction time: automatic facilitation or subjective expectancy. *J. Exp. Psychol.* 2, 567–577.

Laming, D. (1968). *Information Theory of Choice-Reaction Times.* New York: Academic Press.

Laming, D. (1979a). Autocorrelation of choice-reaction times. *Acta Psychol. (Amst.)* 43, 381–412.

Laming, D. (1979b). Choice reaction performance following an error. *Acta Psychol. (Amst.)* 43, 199–224.

Link, S. (1975). The relative judgment theory of two choice response time. *J. Math. Psychol.* 12, 114–135.

Link, S., and Heath, R. (1975). A sequential theory of psychological discrimination. *Psychometrika* 40, 77–105.

Liu, Y., Yu, A., and Holmes, P. (2009). Dynamical analysis of Bayesian inference models for the Eriksen task. *Neural. Comput.* 21, 1520–1553.

Newsome, W., and Pare, E. (1988). A selective impairment of motion perception following lesions of the middle temporal visual area. *J. Neurosci.* 8, 2201–2211.

Pashler, H., and Baylis, G. (1991a). Procedural learning: 1. Locus of practice effects in speeded choice tasks. *Contract* 14, 0281.

Pashler, H., and Baylis, G. (1991b). Procedural learning: 2. Intertrial repetition effects in speeded-choice tasks. *Learn. Mem.* 17, 33–48.

Prudnikov, A., Brychkov, Y., and Marichev, O. (1986). *Integrals and Series. Vol. 1: Elementary Functions.* New York: Gordon and Breach.

Rabbitt, P. (1966). Errors and error correction in choice-response tasks. *J. Exp. Psychol.* 71, 264–272.

Rabbitt, P. (1968a). Repetition effects and signal classification strategies in serial choice-response tasks. *Q. J. Exp. Psychol.* 20, 232.

Rabbitt, P. (1968b). Three kinds of error-signalling responses in a serial choice task. *Q. J. Exp. Psychol.* 20, 179.

Rabbitt, P., and Rodgers, B. (1977). What does a man do after he makes an error? An analysis of response programming. *Q. J. Exp. Psychol.* 29, 727–743.

Rabbitt, P., and Vyas, S. (1981). Processing a display even after you make a response to it. how perceptual errors can be corrected. *Q. J. Exp. Psychol.* 33, 223–239.

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

Ratcliff, R., and McKoon, G. (2008). The diffusion decision model: theory and data for two-choice decision tasks. *Neural. Comput.* 20, 873–922.

Ratcliff, R., and Rouder, J. (1998). Modeling response times for two-choice decisions. *Psychol. Sci.* 9, 347.

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

Ratcliff, R., Spieler, D., and McKoon, G. (2000). Explicitly modeling the effects of aging on response time. *Psychon. Bull. Rev.* 7, 1–25.

Ratcliff, R., Thapar, A., Gomez, P., and McKoon, G. (2004). A diffusion model analysis of the effects of aging in the lexical-decision task. *Psychol. Aging* 19, 278.

Ratcliff, R., Thapar, A., and McKoon, G. (2001). The effects of aging on reaction time in a signal detection task. *Psychol. Aging* 16, 323.

Remington, R. (1969). Analysis of sequential effects on choice reaction times. *J. Exp. Psychol.* 82, 250.

Simen, P., Cohen, J., and Holmes, P. (2006). Rapid decision threshold modulation by reward rate in a neural network. *Neural Netw.* 19, 1013–1026.

Simen, P., Contreras, D., Buck, C., Hu, P., Holmes, P., and Cohen, J. (2009). Reward rate optimization in two-alternative decision making: empirical tests of theoretical predictions. *J. Exp. Psychol. Hum. Percept. Perform.* 35, 1865.

Smith, A., and Spiegelhalter, D. (1980). Bayes factors and choice criteria for linear models. *J. R. Stat. Soc. Series B Stat. Methodol.* 42, 213–220.

Soetens, E., Boer, L., and Hueting, J. (1985). Expectancy or automatic facilitation? Separating sequential effects in two-choice reaction time. *J. Exp. Psychol. Hum. Percept. Perform.* 11, 598–616.

Soetens, E., Deboeck, M., and Hueting, J. (1984). Automatic aftereffects in two-choice reaction time: a mathematical representation of some concepts. *J. Exp. Psychol.* 10, 581–598.

Soetens, E., Melis, A., and Notebaert, W. (2004). Sequence learning and sequential effects. *Psychol. Res.* 69, 124–137.

Sommer, W., Leuthold, H., and Soetens, E. (1999). Covert signs of expectancy in serial reaction time tasks revealed by event-related potentials. *Percept. Psychophys.* 61, 342–353.

Stone, M. (1979). Comments on model selection criteria of Akaike and Schwarz. *J. R. Stat. Soc. Series B Stat. Methodol.* 40, 276–278.

Tuerlinckx, F. (2004). The efficient computation of the cumulative distribution and probability density function in the diffusion model. *Behav. Res. Methods* 36, 702–716.

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

Vervaeck, K., and Boer, L. (1980). Sequential effects in two-choice reaction time: subjective expectancy and automatic after-effect at short response-stimulus intervals. *Acta Psychol. (Amst.)* 44, 175–190.

Wilder, M., Jones, M., and Mozer, M. (2009). Sequential effects reflect parallel learning of multiple environmental regularities. *Adv. Neural Inf. Process. Syst.* 22, 2053–2061.

Yu, A., and Cohen, J. (2009). Sequential effects: superstition or rational behavior. *Adv. Neural Inf. Process. Syst.* 21, 1873–1880.

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: 19 February 2012; paper pending published: 06 March 2012; accepted: 08 June 2012; published online: 16 July 2012.

*Citation: Goldfarb S, Wong-Lin K, Schwemmer M, Leonard NE and Holmes P (2012) Can post-error dynamics explain sequential reaction time patterns? Front. Psychology **3**:213. doi: 10.3389/fpsyg.2012.00213*

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

APPENDIX

In this section, we derive the mean reaction time for the drift diffusion model (DDM) conditioned on hitting either the upper z_u or lower $-z_l$ boundaries, and for a general initial condition $x_0 \in (-z_l, z_u)$.

Suppose that $x(t)$ is the position of a Brownian particle at time t. The dynamics of the movement of this particle are governed by the drift diffusion equation:

$$dx = \mu dt + \sigma dW \tag{A1}$$

$$x(0) = x_0, \tag{A2}$$

in which μ is the deterministic drift of the particle, x_0 is the starting position, and σdW are independent white noise (Weiner) increments of r.m.s strength σ. We assume that the particle is allowed to move until it hits either an upper boundary $x(T) = z_u$ or a lower boundary $x(T) = -z_l$ where T is the hitting time. In this case, the joint densities of the hitting time for boundaries at z_u and $-z_l$ are given by

$$g(t, x(T) = z_u) = \frac{\pi\sigma^2}{(z_u + z_l)^2} e^{\frac{\mu}{\sigma^2}(z_u - x_0)} \sum_{n=1}^{\infty} n e^{-\alpha_n t} \sin\left(\frac{n\pi(z_u - x_0)}{z_u + z_l}\right), \quad t \geq 0, \tag{A3}$$

$$g(t, x(T) = -z_l) = \frac{\pi\sigma^2}{(z_u + z_l)^2} e^{-\frac{\mu}{\sigma^2}(z_l + x_0)} \sum_{n=1}^{\infty} n e^{-\alpha_n t} \sin\left(\frac{n\pi(z_l + x_0)}{z_u + z_l}\right), \quad t \geq 0, \tag{A4}$$

where $\alpha_n = \frac{1}{2}\left[\frac{\mu^2}{\sigma^2} + \left(\frac{n\pi\sigma}{z_u + z_l}\right)^2\right]$ (cf. Feller, 1968; Ratcliff, 1978; Ratcliff and Smith, 2004).

To obtain the conditional densities, one must divide the above equations by the probability of hitting that particular boundary, i.e., $g(t|x(T) = z_u) = g(t, x(T) = z_u)/P[x(T) = z_u]$. These probabilities are (Feller, 1968)

$$P[x(T) = z_u] = \frac{e^{-2\frac{\mu x_0}{\sigma^2}} - e^{2\frac{\mu z_l}{\sigma^2}}}{e^{-2\frac{\mu z_u}{\sigma^2}} - e^{2\frac{\mu z_l}{\sigma^2}}}, \tag{A5}$$

$$P[x(T) = -z_l] = \frac{e^{-2\frac{\mu z_u}{\sigma^2}} - e^{-2\frac{\mu x_0}{\sigma^2}}}{e^{-2\frac{\mu z_u}{\sigma^2}} - e^{2\frac{\mu z_l}{\sigma^2}}}. \tag{A6}$$

Thus, the mean reaction time conditioned on hitting the upper boundary is given by

$$\begin{aligned} \langle T \rangle_{|z_u} &= \int_0^{\infty} t g(t|x(T) = z_u)\, dt \\ &= \frac{1}{P[x(T) = z_u]} \frac{\pi\sigma^2}{(z_u + z_l)^2} e^{\frac{\mu}{\sigma^2}(z_u - x_0)} \sum_{n=1}^{\infty} \frac{n \sin\left(\frac{n\pi(z_u - x_0)}{z_u + z_l}\right)}{\alpha_n^2}. \end{aligned} \tag{A7}$$

Fortunately, a closed-form expression exists for the sum of the infinite series (Prudnikov et al., 1986; Tuerlinckx, 2004):

$$\sum_{n=1}^{\infty} \frac{n \sin(ny)}{\left(C^2 + D^2 n^2\right)^2} = \frac{1}{D^2}\left[\frac{\pi y}{4\frac{C}{D}} \frac{\cosh\left(\left(\pi - y\right)\frac{C}{D}\right)}{\sinh\left(\pi\frac{C}{D}\right)} - \frac{\pi^2}{4\frac{C}{D}} \frac{\sinh\left(y\frac{C}{D}\right)}{\sinh^2\left(\pi\frac{C}{D}\right)}\right]. \tag{A8}$$

We set $C^2 = \mu^2/2\sigma^2$, $D^2 = (\pi\sigma)^2/2(z_u + z_l)^2$, and $y = (z_u - x_0)$. After some algebra, we arrive at a closed form for the mean decision time conditioned on hitting the upper boundary:

$$\langle T \rangle_{|z_u} = \frac{1}{P[x(T) = z_u]} \frac{1}{\mu} e^{\frac{\mu(z_u - x_0)}{\sigma^2}} \left[\frac{(z_u - x_0)\cosh\left(\frac{\mu(z_l + x_0)}{\sigma^2}\right)\sinh\left(\frac{\mu(z_u + z_l)}{\sigma^2}\right) - (z_u + z_l)\sinh\left(\frac{\mu(z_u - x_0)}{\sigma^2}\right)}{\sinh^2\left(\frac{\mu(z_u + z_l)}{\sigma^2}\right)}\right]. \tag{A9}$$

In a similar fashion we obtain the mean decision time conditioned on hitting the lower boundary:

$$\langle T\rangle_{|-z_l} = \frac{1}{P\left[x\left(T\right)=-z_l\right]}\frac{1}{\mu}e^{\frac{-\mu(z_l+x_0)}{\sigma^2}}\left[\frac{(z_l+x_0)\cosh\left(\frac{\mu(z_u-x_0)}{\sigma^2}\right)\sinh\left(\frac{\mu(z_u+z_l)}{\sigma^2}\right)-(z_u+z_l)\sinh\left(\frac{\mu(z_l+x_0)}{\sigma^2}\right)}{\sinh^2\left(\frac{\mu(z_u+z_l)}{\sigma^2}\right)}\right]. \tag{A10}$$

frontiers in
NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 12 June 2012
doi: 10.3389/fnins.2012.00079

Using time-varying evidence to test models of decision dynamics: bounded diffusion vs. the leaky competing accumulator model

Konstantinos Tsetsos[1]*†, Juan Gao[2]†, James L. McClelland[2] and Marius Usher[3]*

[1] *Department of Experimental Psychology, University of Oxford, Oxford, UK*
[2] *Department of Psychology, Center for Mind, Brain and Computation, Stanford University, Stanford, CA, USA*
[3] *School of Psychology and Sagol School of Neuroscience, Tel-Aviv University, Tel-Aviv, Israel*

Edited by:
David Albert Lagnado, University College London, UK

Reviewed by:
Alexander C. Huk, The University of Texas at Austin, USA
Eric-Jan Wagenmakers, University of Amsterdam, Netherlands
Joseph G. Johnson, Miami University, USA

***Correspondence:**
Konstantinos Tsetsos, Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford, OX1 3UD, UK.
e-mail: konstantinos.tsetsos@psy.ox.ac.uk;
Marius Usher, Department of Social Sciences, School of Psychology and Sagol School of Neuroscience, Tel-Aviv University, Tel-Aviv 69978, Israel.
e-mail: marius@post.tau.ac.il

†*Konstantinos Tsetsos and Juan Gao have contributed equally to this work.*

When people make decisions, do they give equal weight to evidence arriving at different times? A recent study (Kiani et al., 2008) using brief motion pulses (superimposed on a random moving dot display) reported a *primacy* effect: pulses presented early in a motion observation period had a stronger impact than pulses presented later. This observation was interpreted as supporting the *bounded diffusion* (BD) model and ruling out models in which evidence accumulation is subject to leakage or decay of early-arriving information. We use motion pulses and other manipulations of the timing of the perceptual evidence in new experiments and simulations that support the leaky competing accumulator (LCA) model as an alternative to the BD model. While the LCA does include leakage, we show that it can exhibit primacy as a result of competition between alternatives (implemented via mutual inhibition), when the inhibition is strong relative to the leak. Our experiments replicate the primacy effect when participants must be prepared to respond quickly at the end of a motion observation period. With less time pressure, however, the primacy effect is much weaker. For 2 (out of 10) participants, a primacy bias observed in trials where the motion observation period is short becomes weaker or reverses (becoming a *recency* effect) as the observation period lengthens. Our simulation studies show that primacy is equally consistent with the LCA or with BD. The transition from primacy-to-recency can also be captured by the LCA but not by BD. Individual differences and relations between the LCA and other models are discussed.

Keywords: bounded diffusion, LCA, perceptual choice, non-stationary evidence, order effects

INTRODUCTION

The process of decision making has been the subject of intensive recent investigations in both experimental psychology (Usher and McClelland, 2001; Ratcliff and Smith, 2004; Brown and Heathcote, 2005; Bogacz et al., 2006; Ratcliff and McKoon, 2008; van Ravenzwaaij et al., 2012) and neuroscience (Huk and Shadlen, 2005; Gold and Shadlen, 2007; Ratcliff et al., 2007; Wong et al., 2007; Wang, 2008; Ditterich, 2010; Rorie et al., 2010). A central idea emerging from these investigations is that decision makers take multiple samples of noisy evidence and integrate them over time until the integrated evidence reaches a decision boundary. The time to reach the bound determines the reaction time (Gold and Shadlen, 2001, 2002; Roitman and Shadlen, 2002). Some of these decision making models generate optimal decisions in the sense that they achieve the shortest possible mean reaction time for a fixed error-rate (Wald, 1946; Gold and Shadlen, 2001, 2002; Bogacz et al., 2006). In addition, neurophysiological studies have reported that when monkeys make decisions about the direction of motion in a noisy moving dots display, neurons in several visual-motor integration areas (e.g., the lateral intraparietal cortex, LIP) show ramping activity consistent with the integration of evidence (Hanes and Schall, 1996; Gold and Shadlen, 2000, 2001; Horwitz and Newsome, 2001; Shadlen and Newsome, 2001).

A number of computational models that can account for both the behavioral and physiological choice data have been developed. These models not only account for the accuracy of participants' responses, but also for details of the distributions of response times and their dependence on experiment conditions such as difficulty levels and speed-accuracy instructions (Ratcliff and McKoon, 2008).

The starting point for a wide range of decision making research is the *drift-diffusion* model (Ratcliff, 1978; Ratcliff and Rouder, 1998; Ratcliff and McKoon, 2008). In this model, the difference in evidence supporting each of two decision-alternatives is accumulated linearly over time, without loss or distortion. Here we consider a variant of this model that is often used to address neurophysiological data (Mazurek et al., 2003; **Figure 1B**). This model is represented as a process in which accumulators integrate the difference in the momentary evidence for the two alternatives via a combination of feed-forward excitation and inhibition, such that positive evidence for one alternative is negative evidence for the other.

In recent years, several researchers have proposed decision making models that do not adhere to the perfect integration of the

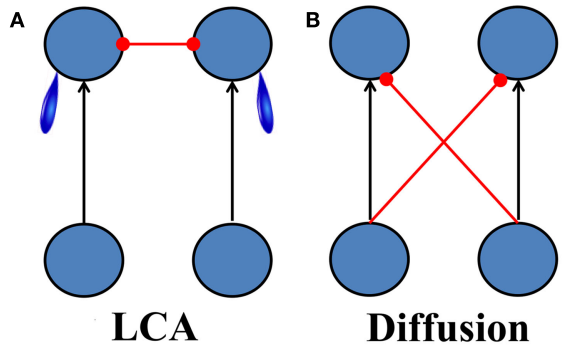


FIGURE 1 | Architecture of the two-choice reaction-time models. **(A)** The leaky competing accumulator model (Usher and McClelland, 2001), **(B)** Mazurek et al. (2003) model. Arrows and filled circles indicate excitatory and inhibitory connections respectively. Blue tears indicate leakage.

drift-diffusion model. These models include variations based on the *Decision Field Theory* (Busemeyer and Townsend, 1993; Roe et al., 2001; Johnson and Busemeyer, 2005), the neurophysiologically grounded attractor network model (Wang, 2002; Wong and Wang, 2006) and the *leaky competing accumulator* model (LCA; Usher and McClelland, 2001; Bogacz et al., 2007). To varying degrees, all of these models draw on inspiration from principles of neural computation and attempt to capture ways in which decision making deviates from perfect optimality. For example, these models incorporate the possibility of leakage or decay of information, as well as mutual inhibition between the representations of the decision-alternatives, and both the attractor and LCA models incorporate non-linearities that can affect information integration.

In the present work we focus primarily on the LCA (**Figure 1A**). In this model, as in the model of Mazurek et al. (2003), accumulators representing the available alternatives accumulate noisy evidence over time, but in this case, there is no feed-forward inhibition. Instead, accumulated evidence is subject to leak, and the accumulators compete with each other through mutual inhibition. The LCA has been successful in capturing a number of features of human decision making data (Usher and McClelland, 2001, 2004; Bogacz et al., 2007; Gao et al., 2011; Tsetsos et al., 2011). This model is intermediate in complexity between the other models; it introduces a lower bound on activation, unlike the decision field theory, but it lacks additional features that are present in the attractor model, including an activity dependent gating of special channels that change its leakage characteristics. We retain the lower bound at 0 because it has important implications for aspects of the dynamics of decision making that have already received support in another recent study (Tsetsos et al., 2011). As we shall see, this lower bound will also play a role in understanding the findings we will present in the present article. The greater simplicity of the LCA compared to the attractor model (Wang, 2002) makes it more tractable for analysis, and this is one of the prime reasons for our focus on the LCA. We are open, however, to the possibility that the added features of the attractor model may be important, and we will return to this class of models in the Section "Discussion."

Research on decision making often employs what is called the *free-response* paradigm, which sets up decision-time under the control of the observer. In this paradigm, a stimulus is presented on each trial, and participants are assumed to integrate evidence until they reach a decision bound. All of the models under consideration assume that this bound represents a criterion amount of accumulated evidence. However, the models differ in their handling of decision making in *time-controlled* paradigms, in which evidence is presented for a period of time controlled by the experimenter, and in which the overt response is prompted by a cue called a *go cue*. When difficult stimuli are used in such experiments, stimulus sensitivity (measured by d') is 0 with very short evidence accumulation times, then rises to a finite asymptotic level after about 1 s, remaining constant even if more integration time is allowed (Wickelgren, 1977; Usher and McClelland, 2001; Kiani et al., 2008). The LCA and the diffusion model have different ways of addressing this finding. In the LCA and related models, evidence accumulation is assumed to continue until the end of the evidence evaluation period, at which point the decision maker is thought to choose the alternative associated with the most active accumulator. The fact that accuracy levels off is attributed to an imbalance between leak and inhibition, as discussed in more detail below. In contrast, in the Mazurek et al. version of the drift-diffusion model, decision sensitivity can increase without bound as integration time increases, since there is no loss or distortion in evidence accumulation; the model predicts that the signal to noise ratio should increase with $\sqrt{t}$. To address the fact that performance levels off in time-controlled paradigms, Mazurek et al. (2003) proposed that, just as in free-response paradigms, participants employ a decision bound in time-controlled situations, such that evidence integration stops when the boundary is reached, even though stimulus input continues and the response must be withheld until a cue to respond is presented (see also Ratcliff, 2006). Because of the presence of this decision bound, even in time-controlled situations, we call this model the *bounded diffusion* (BD) model in the remainder of this article.

In a recent paper (Kiani et al., 2008), the authors proposed a way to determine whether the leveling off of accuracy in time-controlled paradigms is more consistent with the presence of a bound, or alternatively with leaky integration. The paper considered the BD model and what they referred to as the leaky accumulation model, a variant of the LCA in which leakage is stronger than inhibition (henceforth called the *leak-dominant LCA*). The leaky accumulation model predicts that late information is more important (a pattern called *recency*) since early information has more time to leak away. This contrasts with the BD model, which predicts that early information is more important (a pattern called *primacy*) because late information is more likely to arrive after the bound is reached and therefore to be ignored.

Two pieces of evidence were shown to support the primacy pattern in the experiment. The first was based on the reverse correlation technique. The reverse correlation analysis is applied to experimental trials in which the evidence (in the form of dot motion) is completely random. Trials are grouped according to the observed response choice between the two available alternatives, which we will label A and B. The analysis examines the averaged input signal in the time course of the entire trial in the

two groups. If the analysis reveals no difference between the A and B groups at some time points, it means that inputs at those time points do not contribute to the outcome. On the other hand, if the analysis reveals a large difference between the two groups of trials at some time points, it means inputs at those time points are contributing to determining the response. When this analysis was applied to model simulations, it confirmed that the BD model predicts a primacy pattern while the leaky accumulation model predicts a recency pattern (**Figure 2A**). The same analysis based on behavioral data demonstrated a primacy pattern (**Figure 2B**). The second source of support for the primacy pattern was based on a pulse perturbation study, using 200 ms motion pulses that influenced monkey's choices in the direction of the pulse. The size of the pulse effect was largest when the pulse was applied early in the trial, and decreased when pulses occurred later, consistent with BD rather than leaky accumulation.

In the present paper, we further examine the temporal weighting of evidence in experiments and in the LCA and BD models. Our examination is motivated by both empirical and model-based observations presented in Usher and McClelland (2001). On the empirical side, the result of the perturbation study in Kiani et al. (2008) stands in contrast with experimental findings reported in Usher and McClelland, 2001; Experiment 3). In that experiment, participants viewed a stream of interleaved S's and H's and reported after the end of the sequence which letter was predominant. While most of the trials contained sequences with a majority of either S or H, some of the trials contained equal numbers of S's and H's. Within the latter type of trials, one of the letters sometimes predominated early in the trial, with the other letter predominating later. Out of the six subjects, two showed a primacy bias, favoring the letter that predominated early in the sequence; two showed a recency bias, favoring the latter that predominated late; and two showed approximate balance, or little bias in either direction.

On the theoretical side, the LCA was able to account for all three types of behavior. While the model shows a recency pattern when leak is stronger than inhibition, it shows a primacy pattern when inhibition is stronger than leak, and it shows equal weight of early and late information when the strength of leak and inhibition are equal. All else being equal, balanced leak, and inhibition lead to greater accuracy, and indeed, the experimental data indicated the expected relationship between accuracy on trials when the number of S's and H's were different, and the degree of bias (either toward primacy or recency) exhibited on trials when the number of S's and H's was the same. Specifically, greater imbalance when the number of S's and H's was the same

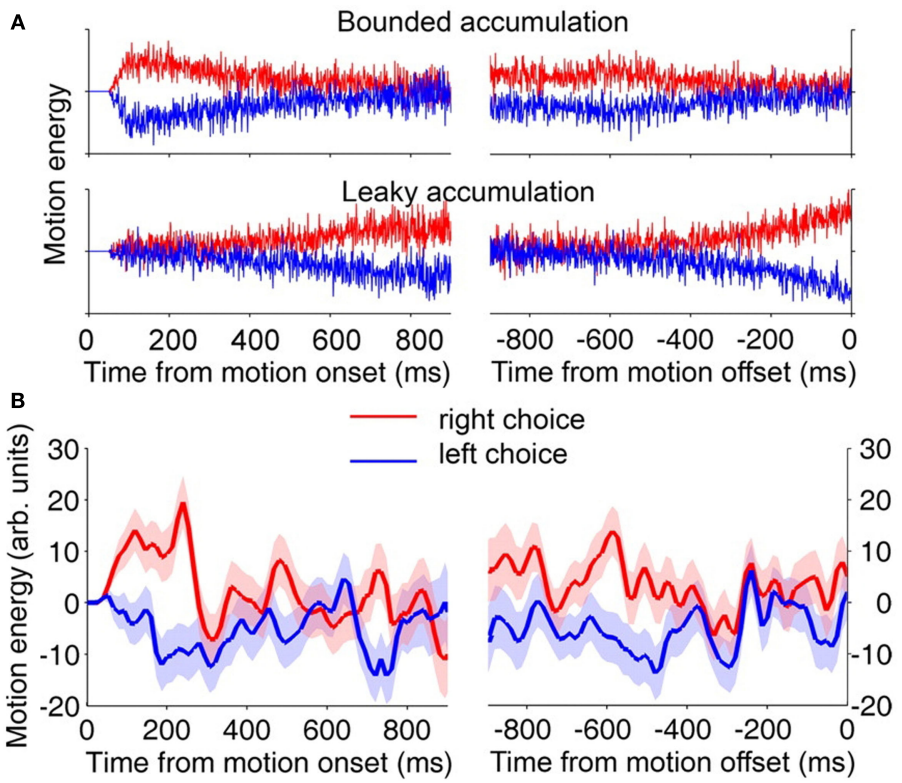


FIGURE 2 | Reverse correlation analysis (reproduced from Kiani et al., 2008). (A) Expected separation of motion energy profiles for rightward (red) and leftward (blue) choices for bounded diffusion (top) and leak-dominant LCA (bottom). Late information is more critical in the leak-dominant LCA model while early information is more critical in the bounded diffusion model. **(B)** Left, signals aligned with motion onset, right signals aligned with motion offset. One can observe that in the data (Panel B), the difference between the evidence that favors the response (red) and the one that opposes it (blue) is larger at the beginning of the choice interval.

was associated with lower accuracy when the number of S's and H's was different.

The present study seeks to examine these empirical and model-based considerations further. On the empirical side, there are many differences between the experiments of Usher and McClelland (2001) and Kiani et al. (2008). Among other things, Usher and McClelland's study involved six relatively unpracticed human participants who were not placed under strong time pressure. Kiani et al. used two highly practiced rhesus macaque monkey participants who received a go cue (on half of their experimental trials) coincident with the end of the stimulus presentation period, requiring them to respond within 500 ms. Several questions naturally arise: Would different patterns have been observed in the Kiani et al. study if it had been conducted on humans? Would individual differences have emerged had a larger number of participants been tested? Does extensive practice, or the need to be prepared to respond quickly, alter the tendency to observe a pattern of primacy vs. recency? The present research attempts to address these issues by using a paradigm quite similar to that of Kiani et al. (2008), employing highly practiced human participants, and manipulating the time pressure to respond across experiments. While our studies still use relatively small numbers of participants, we will see that there are indeed considerable individual differences within the set of participants.

Another goal of our research is to further explore the primacy pattern seen in some participants in both the Usher and McClelland (2001) and Kiani et al. (2008) studies. We will examine whether the LCA can capture the primacy pattern as well as the BD model does, and whether it can also capture other aspects of performance that are challenging to the BD model. As we will see, the LCA can exhibit primacy on some trials and recency on others, using the same parameter values. That is, it can exhibit a primacy effect when the length of the evidence accumulation interval is short, while exhibiting a recency effect when the length of the evidence accumulation interval is long. Our study will allow us to examine whether such a pattern can be observed in human participants.

We begin by reviewing an analysis of the LCA presented in Usher and McClelland (2001), extending this analysis by further examining the model using the same reverse correlation analysis as in Kiani et al. We then discuss the primacy-to-recency shift that can occur in the LCA model under certain ranges of its parameter values. Following this, we report two experimental studies with human observers[1]. In the first we place participants under high time pressure, using procedures similar to Kiani et al. and we find similar primacy patterns. In the second, we relax the time pressure by lengthening the response window and introducing longer trials, and we find the primacy bias diminishes significantly. We see individual differences in both studies, with one participant in the second study showing recency for short evidence integration periods and primacy for long integration periods. As the moving dots paradigm is a central one to the neuroscience of decision making (Burr and Santoro, 2001; Shadlen and Newsome, 2001; Kiani et al., 2008), we use moving dot stimuli in our study. Our use of temporally manipulated stimuli builds on the pioneering efforts of Huk and Shadlen (2005) and Wong et al. (2007) as well as the study of Kiani et al. (2008).

MATERIALS AND METHODS

EXPERIMENTAL METHODS

Moving dot stimuli

The moving dot stimuli were created following the method described in Kiani et al. (2008). The motion stimulus consisted of circular dots of radius 2 pixels, moving horizontally at a speed of 5°/s. Total dot density was 16.7 dots per degree squared per second. The stimulus was viewed through a circular aperture of radius 5°. The coherence of the motion stimulus varied from trial to trial and within trials as specified below.

Dots were randomly divided into three sets. One set of dots was displayed per frame, which lasted 13.33 ms. Each set of dots appeared on the monitor once every frame-triple, each of which contains three frames, spanning 50 ms. On every displayed frame, each dot had a (1 – coherence) probability of being redrawn at random coordinates within the circular aperture. Those not redrawn at random would be redrawn to move horizontally 5°/s in the direction specified for the trial. At 0% coherence, every dot would be redrawn randomly on every frame.

Experiment 1A

In this experiment, 80% of trials with duration 300 ms or greater contained a "pulse," or momentary change of coherence level. A pulse consisted of a ±3.2% change in coherence level for 200 ms, or four frame-triples. The motion pulse could originate between 100 ms after the beginning of the stimulus and 200 ms before it ended. See Appendix for detailed information about the pulse.

Observers. Three participants (CS, MT, and SC; two male, one female) with normal or corrected-to-normal vision were tested. Participants CS, MT, and SC performed 32, 46, and 34 sessions respectively. Ordinarily, successive sessions were separated by less than 5 days, but there were some exceptions (this was also the case for experiments 1B and 2A,B). We excluded initial sessions while participants' performance stabilized, excluding 5 sessions for CS, 14 for MT, and 12 for SC, leaving 27 sessions for CS, 32 sessions for MT and 22 sessions for SC that were treated as test sessions included in our analysis.

Procedure. In each session, participants completed 9 blocks of 100 trials. A self-paced break occurred between blocks to allow rest. Each trial began with a fixation cross at the center of the screen. The moving dots stimulus was displayed 1000 ms later. Coherence values employed were 6.4, 12.8, 25.6, and 51.2%. Stimulus duration followed an exponential distribution taking values from 100 to 1750 with an increment of 50 ms. Stimulus termination occurred simultaneously with an auditory go signal. In order to earn points, participants had to respond by pressing the correct key on the computer keyboard within a 300 ms response window following the go cue.

Visual and auditory feedback was used to indicate to the participant whether the response occurred within the specified

[1]The data set is available at: http://www.stanford.edu/group/pdplab/projects/Frontiers2012/.

response interval, and (if so) whether it was correct. If participants responded within the response window and chose correctly, they heard a pleasant noise and saw the number of total points they earned (which increased from the previous value by 1) in a box at the position of the fixation. Incorrect, early, or too late responses earned no points and were followed by an "X," "Early," and "Too Slow" signs in the box together with an error, early, or late tone. The total time allotted for feedback of any type was 1 s. After the feedback time had elapsed, the fixation point appeared and the next trial began.

Experiment 1B

This experiment was carried out in order to obtain a more robust measure of the recency-primacy bias. Instead of applying pulses at different times of the trial as in Experiment 1A, for each coherence level we created three conditions: (i) the constant condition in which a fixed non-zero coherence was used during the entire trial, (ii) the early condition in which the coherence was one of the four values as in Experiment 1A during the first half of the trial and zero during the second half, (iii) the late condition in which the coherence was zero in the first half and non-zero in the second half. In addition, for the two weakest coherence levels only, we included a switch condition, in which the coherence value stayed constant in magnitude but the direction of motion switched in the middle of the stimulus duration. For the constant, early, and late conditions, the correct response was defined as the response supported by the stimulus. In the switch condition, one alternative was designated as correct at random on each trial.

Observers. The same three participants, CS, MT, SC, from Experiment 1A, participated in Experiment 1B and completed 14, 19, and 12 sessions respectively. One session was excluded for CS and MT due to a programming error[2], and nine more sessions were excluded for MT due to unstable performance (See Excluded Sessions in Experiment 1B in Appendix). This resulted in 13, 9, and 12 analyzed sessions for CS, MT, SC respectively.

Procedure. General features of the procedure were the same as in Experiment 1A. Coherence values were 6.4, 12.8, 25.6, and 51.2%, except in the switch condition where only 6.4 and 12.8% were used. Stimulus duration followed an exponential distribution from 150 to 1750 ms with an increment of 50 ms. As in Experiment 1A the response window was 300 ms.

Experiment 2A

In this experiment, we relaxed the time pressure by using a longer response window after the go cue and by using more long trials.

Observers. Four participants (one male, three female) with normal or corrected-to-normal vision were tested repeatedly in 1-h sessions over several weeks. We obtained 16, 19, 11, and 25 sessions for participants DG, LK, WW, and MM respectively. All sessions were included in our analyses.

[2]Due to a programming error, the direction of motion in the first half of each switch trial was treated as correct in the first session for participants MT and CS.

Procedure. The procedure of the experiment was the same as Experiment 1B, except for two changes. First, the response window after the go cue was extended from 300 ms to 1 s. As in previous experiments, if a response was made outside of the response window, no points were awarded even if the response was correct. Second, we employed a flat distribution of trial durations over the range of 150–1750 ms with an increment of 100 ms.

Experiment 2B

Experiment 2B was the same as Experiment 2A except that: (i) there were only early, late, and constant conditions (no switch) in this experiment, (ii) the stimulus duration was sampled from a longer range (150–2350 ms, increment of 200 ms), and (iii) an adaptive procedure was used to maintain accuracy at an approximately constant level across subjects. This was done by using a baseline coherence level b, which was adaptively changed from block to block, decreasing b by amount δ when the overall accuracy in that block was above 80% or incrementing it by δ when accuracy fell below 65%. Three coherence levels were used, equal to b, $2b$, and $4b$. In the first session, the baseline coherence was initially set to 12% and δ was set to 1.6%; for later sessions, the initial value of b was determined based on the last block from the previous session, and δ was set to 0.86% (this value changes the average coherence by 2%). For example, if in a given block in session 2 or later, the coherence levels were 5, 10, and 20%, and performance fell below 65% correct, the resulting coherence levels would be set to 5.86, 11.72, and 23.44%.

Observers. Three participants with normal or corrected-to-normal vision were tested in 5 (AP) or 10 (CB, SY) 1-h sessions over several weeks. We intended to run each participant for 10 sessions, treating the first three as practice and for stabilization of coherence levels, and analyzing the results from the remaining seven sessions. However, participant AP stopped participating after five sessions. Rather than exclude the participant completely, we excluded only the first session of this participant, leaving four sessions for inclusion in the analysis.

COMPUTATIONAL METHODS

The LCA and BD models were simulated as two-layered neural networks illustrated in **Figures 1A,B** respectively. The simulation of the LCA model was based on the following finite difference equations[3]:

$$\Delta x_1 = I_1 - k x_1 - \beta x_2 + I_0 + N(0,\sigma); \quad (1)$$
$$\Delta x_2 = I_2 - k x_2 - \beta x_1 + I_0 + N(0,\sigma),$$

subject to a lower bound on activation at 0:

$$x_1(t+1) = \max(0, x_1(t) + \Delta x_1);$$
$$x_2(t+1) = \max(0, x_2(t) + \Delta x_2).$$

[3]These equations correspond to discrete versions of the differential Equations $dx_1 = dt\left[I'_1 - k'x_1 - \beta' x_2 + I'_0\right] + N\left(0,\sigma'\right)\sqrt{dt}$; $dx_2 = dt\left[I'_2 - k'x_2 - \beta' x_1 + I'_0\right] + N\left(0,\sigma'\right)\sqrt{dt}$ with the following correspondences with the parameters in the finite difference equations: $I_1 = I'_1\,dt$, $I_2 = I'_2\,dt$, $I_0 = I'_0\,dt$, $k = k'\,dt$, $\beta = \beta'\,dt$, and $\sigma = \sigma'\sqrt{dt}$. In the simulations, $dt = 0.0035$ s (3.5 ms) and the reported parameter values are those in the finite difference equations.

In Eq. 1, Δ represents a change or increment in the adjacent variable, I_0 is a baseline input, k and β stand for the leak and the lateral inhibition and $N(0, \sigma)$ stands for normally distributed noise of standard deviation σ. The output of the max function is equal to its second argument when this is positive and is equal to 0 otherwise. This max function introduces non-linearity to the system that prevents x_1 or x_2 from becoming negative.

In time-controlled paradigms such as the one used here and in Kiani et al. (2008), in which a decision is called for by presenting a go cue, the model assigns the decision to the most active accumulator a short time after the go cue occurs as discussed further below.

The simulation of the BD model was based on

$$\Delta x_1 = I_1 + N(0,\sigma); \quad \Delta x_2 = I_2 + N(0,\sigma), \tag{2}$$

The decision variables are $y_1 = x_1 - x_2$, and $y_2 = x_2 - x_1$.

In BD, information integration is subject to a bound, even in time-controlled paradigms. When the activation of one of the accumulators, y_1 or y_2, corresponding to the difference between the integrated evidence for the two alternatives, in Eq. 2, reaches the bound, the race ends and the more active unit at that time wins the trial. If the bound is not reached, the model assigns the decision to the most active accumulator a short time after the go cue occurs, as in the LCA.

Shared parameters

The noise strength was set at $\sigma = 0.1$ in both models. The inputs to the units were $I_1 = c \times s$, and $I_2 = 0$, where c stands for the coherence level and where sensitivity, s, is a free parameter fitted for each model. All simulations employed an integration time step of 3.5 ms.

The experiments we will report involve presenting a visual stimulus at some time $t = 0$ and then presenting a response signal or "go cue" at a variable time post stimulus onset. Responses are considered to be triggered by the go cue. Thus, the time between the stimulus onset and the presentation of the go cue – the *go cue delay* – could be taken as the duration of the information integration period. In relating both models to experimental data, however, we included a "dead-time" parameter, T_0, to allow for the possibility that the presentation of an imperative signal to respond terminates evidence integration before all the evidence actually presented up to that time has been integrated. Previous research has established that evidence accumulation in area LIP lags behind the actual presentation of the visual evidence by about 200 ms (Mazurek et al., 2003; Rorie et al., 2010). If the go cue can terminate evidence accumulation with a shorter lag, $T_g < 200$ ms, then the total time available for evidence integration would be equal to the go cue delay less the difference between T_g and 200. The parameter T_0 represents this difference $(200 - T_g)$ and is assumed to be greater than or equal to 0.

Model specific parameters: bounded diffusion

In addition to the parameters already mentioned, the BD model had one additional parameter, the position of the decision bound, A. The value of A was assumed to take a single fixed value for each participant, independently of the coherence level of the stimulus or the trial duration, since all levels of both variables were randomly intermixed and therefore unpredictable from trial to trial.

Model specific parameters: LCA

The LCA model was implemented with two additional free parameters that were optimized to fit the data, namely the leak and inhibition strengths k and β. The LCA also includes a parameter representing the common input to the two accumulators, I_0, which was set at $I_0 = 0.2$ in fitting the model to all participants. This parameter determines how likely it is that the activation bound of zero is reached by the losing accumulator in the LCA. The particular value was chosen on the basis of exploratory simulations so that this boundary is often but not always reached on longer trials, and was not otherwise adjusted in fitting data from individual participants in our experiments. This parameter had different values in the simulation studies; the values are explicitly reported in the relevant sections below.

Simulation protocol

According to the protocol of experiment 1B (see Experimental Methods), there were four levels of motion coherencies ($c = 6.4$, 12.8, 25.6, 51.2%) and four different timing conditions (*constant*, *early*, *late*, and *switch*). Since we have less data for the switch condition, which occurred only with the two lowest coherences, we fitted the models based on the constant, early, and late conditions and used the optimized parameters to predict the choice preference in the switch condition. Assuming that unit one is supported by the stimulus and unit two is not supported, the inputs in the three conditions were assigned in the following way. In the constant condition $I_1 = c \times s, I_2 = 0$ throughout the entire trial. In the early condition $I_1 = c \times s, I_2 = 0$ for the first half of the trial and $I_1 = 0$, $I_2 = 0$ for the second half. In the late condition, $I_1 = 0$, $I_2 = 0$ for the first half and $I_1 = c \times s$, $I_2 = 0$ for the second half. The durations of the simulation trials were sampled from an exponential distribution with a mean of $\mu = 243$ simulation time-steps, or 850 ms. The minimum duration was set at 43 time-steps (150 ms) and the maximum at 500 time-steps (1750 ms). The trials were grouped in quartiles according to stimulus durations, resulting in 48 conditions (4 coherencies × 3 conditions × 4 durations).

Optimization procedure

The best fitting parameters of the models were obtained by an optimization procedure performed on the 48 (4 coherencies × 3 timing conditions × 4 durations) mean accuracy scores of each participant. For presentation purposes we averaged the experimental data and the fits across the four coherency levels. Assuming that the correct responses follow a binomial distribution, we can compute the likelihood of a model given the $N = 48$ experimental conditions as: $L = \prod_i^N \binom{n_i}{y_i} p_i^{y_i} (1 - p_i)^{n_i - y_i}$, where $N = 48$ is the number of data points, n_i is the number of trials for the i-th data point, y_i is the corresponding number of correct responses and p_i the probability of correct response predicted by the model. The cost function we minimized was the negative logarithm of L, i.e., $-\mathrm{LL} = -\log_e(L)$. For optimization we used the SUBPLEX minimization routine (Bogacz and Cohen, 2004), which extends the multi-dimensional simplex algorithm in order to better handle noisy functions for simulation-based models. For each subject

and each model we ran the optimization 200 times with starting points randomly sampled from uniform distributions within a parameter-specific range. At that stage, each predicted data point was generated from 1000 simulated trials. We re-evaluated each of these 200 fits by running more iterations of the model with the best fitting parameters (10,000 simulated trials per data point). At the final refinement stage, the parameters of the best fit (after the re-evaluation of the 200 parameter sets) were used as the starting point of one last run of the SIMPLEX routine, using 2000 simulated trials per data point.

In order to compare the quantitative fits of the two models we used the Bayesian information criterion (BIC), which takes into account both the goodness of fit and the complexity of the model. The BIC penalizes the extra free parameters much more strongly than other similar measures such as the Akaike information criterion. The BIC is computed as: $-2\mathrm{LL} + P1\mathrm{n}(N)$, where P is the number of the free parameters of the model, N the number of data points and LL is as defined above. For **Figure 7** and for the calculation of BIC values, the models were run with the best fitting parameters for 100,000 simulated trials per data point.

RESULTS

We start with a computational investigation showing that the LCA model can capture all three of the patterns seen in Experiment 2 of Usher and McClelland (2001), namely primacy, recency and perfect balance. We also demonstrate that the LCA model with moderate inhibition dominance predicts a transition from primacy-to-recency as the duration of the trial increases. Following the computational investigations, we present the experimental results.

CONTRASTING BOUNDED DIFFUSION AND LEAKY INTEGRATION: A SIMULATION STUDY

For binary choices, the LCA is a stochastic two-dimensional system described by two variables x_1 and x_2, each corresponding to the accumulated evidence for one of the two alternatives. Each accumulator is updated at every simulation time step according to Eq. 1 presented in Section "Materials and Methods," and reproduced here for convenience:

$$\Delta x_1 = I_1 - kx_1 - \beta x_2 + I_0 + N(0,\sigma); \quad (1)$$
$$\Delta x_2 = I_2 - kx_2 - \beta x_1 + I_0 + N(0,\sigma).$$

As noted in the Section "Materials and Methods," the values of x_1 and x_2 were subject to a lower bound on activation at 0.

When x_1 and x_2 are both positive, the LCA dynamics stay in the linear regime. Since decisions are based only on which of the two decision variables is more active, we only need to examine the difference between them: $x = x_1 - x_2$. In this case, LCA is reduced to the Ornstein–Uhlenbeck (OU) diffusion process (Busemeyer and Townsend, 1993; Usher and McClelland, 2001):

$$\Delta x = I - (k-\beta)x + N\left(0, \sqrt{2}\sigma\right) \quad (3)$$

where $I = I_1 - I_2$. When leak exceeds inhibition, the activation difference x is characterized by leaky accumulating dynamics. Both the mean and the standard deviation of x stop changing once the net leak [equal to $(k-\beta)x$] in Eq. 3 becomes equal in magnitude to the input term I. The left column in **Figure 3** demonstrates how the distribution of x evolves with time. The resulting accuracy, which corresponds to the area of the distribution to the right of the vertical neutral line, therefore also levels off at an asymptotic value. Since evidence that arrives early has a longer time to leak away than the information that arrives late, late information overweighs early information under these circumstances, causing the recency effect.

On the other hand, when inhibition dominates leak in the full model, $k < \beta$, the quantity $(k-\beta)$ in Eq. 3 becomes negative; taking this together with the minus sign in front of the $(k-\beta)x$ term, we see that net effect of leak and inhibition becomes self excitation. In that case, any difference between the two decision variables will grow and explode with time. See **Figure 3**, middle column. Since early evidence has more time to grow than late information, early evidence overweighs late information in determining decisions, causing primacy. Although the mean and the standard deviation of the distribution in this condition both grow without bound as time increases, the resulting choice probability, determined by the ratio between the two, evolves and levels off with time in the same way in this condition as in the leak-dominant condition (see Usher and McClelland, 2001; Gao et al., 2011 for more details). Finally, when leak and inhibition are in perfect balance, $k = \beta$, neither leak nor self excitation occurs. The $(k-\beta)x$ term disappears from Eq. 3, and the model behaves as the drift-diffusion model (Bogacz et al., 2006; this case is not illustrated in the figure).

Non-linearity comes into play in the inhibition-dominant regime. According to the linear version of the LCA in Equation 3, the self excitation drives the evidence difference, x, to infinity with time. However, in the full LCA model, including the non-linearity at 0, once the losing unit's activation reaches 0, it stops going further down and stops sending any inhibitory signal. The activity of the winning unit will be driven only by its leak and by its input (I_1 or I_2 depending on which unit is the winner). Therefore its activity, as well as the difference between its activity, levels off as further time passes. **Figure 3**, right column demonstrates the dynamics of the evidence difference variable x in this situation. Although the detailed dynamic of x in the non-linear model differs from that in the linear version, the choice probability distributions for the two models are very similar. This is because the non-linearity takes effect only after some time has passed. By this time, the amplification of early signals has already exerted its influence on the outcome (Usher and McClelland, 2001). Therefore, in the inhibition-dominant regime, both the full non-linear LCA and the linearized LCA produce a primacy pattern.

To illustrate the recency and primacy effects exhibited by the leak and inhibition-dominant LCA we performed the same reverse correlation analysis as in **Figure 2**, comparing leak-dominant and inhibition-dominant LCA with the BD model (**Figure 4**). Both alternatives (left/right) received noisy input for 200 simulation time-steps (Gaussian values with zero mean and SD of 0.1). BD was simulated with $A = 0.8$, inhibition-dominant LCA with $k = 0.05$, $\beta = 0.095$, $I_0 = 0.1$ and leak-dominant LCA with $k = 0.05$, $\beta = 0.025$, $I_0 = 0.1$. Larger differences between the left choice activity curve (blue) and the right choice activity curve

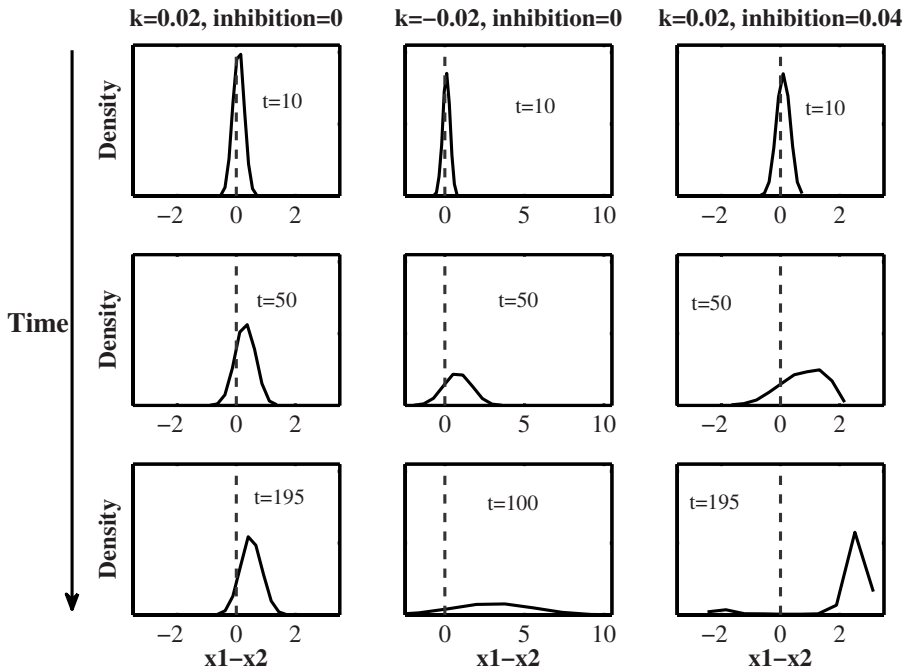


FIGURE 3 | Time evolution of the decision variable $x = x_1 - x_2$ in three different leak-inhibition conditions. Adapted with permission from Usher and McClelland (2001).

(red) at the beginning of the trial indicates primacy, while larger differences at the end indicates recency. **Figure 4** demonstrates that although the leak-dominant LCA (**Figure 4C**) results in recency, the inhibition-dominant LCA results in primacy (**Figure 4B**). The behavioral results reported in Kiani et al. (2008), although inconsistent with the leak-dominant LCA, are thus shown to be consistent with either the BD model or the inhibition-dominant LCA.

The rich dynamics of the LCA model also allows it, with certain settings of its parameters, to produce a transition between primacy and recency. In **Figure 5A**, we demonstrate such a case with the following values of the leak, inhibition and baseline parameters: $k = 0.172$, $\beta = 0.748$, $I_0 = 0.095$, and $\sigma = 0.1$. We simulated a switch trial, where the motion coherence stays constant in magnitude throughout the trial but the direction switches in the middle. Inputs are $I_1 = 0.026, I_2 = 0$ for the first half of the trial and $I_1 = 0$, $I_2 = 0.026$ for the second half. We plot the probability of choices supported by the early half of the trial. A value above 0.5 implies early information determines the final decision more often than late information, i.e., primacy, and a value below 0.5 implies early information determines decisions less often than late information, i.e., recency. Each data point is based on simulations of 30,000 trials, and five durations were used consisting of 71, 157, 243, 329, and 414 time-steps. One can see a transition from primacy to recency as stimulus duration increases.

In order to explain how the transition results from the LCA, we show activations of the two accumulators in a typical trial in **Figure 5B**. The red curve stands for the alternative supported in the first half of the trial, and the blue curve for the one supported in the second half. When stimulus duration is short (top panel), the accumulator associated with the red curve wins because the input during the first half of the trial leads it to suppress the other alternative, which does not have a chance to recover after the evidence reverses. At the time of the switch, the early-supported (red) accumulator is sending strong inhibition to the other accumulator (blue curve). Although the blue accumulator is supported by the stimulus input in the second half of the trial, its activation grows very slowly, rising only after the red accumulator's activation has sufficiently decayed. This takes long enough so that the blue accumulator does not have a chance to win out. When stimulus duration is long (bottom panel, solid lines), the activity of the blue accumulator reaches zero well before the switch and stays pinned at this value. Following that, the activity of the red curve levels off; it no longer receives any inhibition from the other accumulator, but its activation levels off due to the effect of the leak. Therefore, although the first half of the trial in this case is much longer than that in the short duration scenario, the activity levels of the two accumulators are similar at the time of the switch. After the switch, the two curves evolve with time in the same manner as they do in the top panel. However, in this case, the activation of the red accumulator has more time to decay. The activation of the blue accumulator has more time to grow and its activation eventually comes to surpass the activation of the red alternative. Note that this transition from primacy to recency is caused by the interplay between the non-linearity at zero and the greater weight to early evidence caused by inhibition dominance. It does not occur in the linear case (dashed lines, lower panel of **Figure 5B**), nor does it occur with a high level of inhibition dominance.

In summary, primacy bias is consistent with both the BD and the inhibition-dominant LCA. However, LCA is also consistent with recency or balanced weighting of early vs. late evidence. A

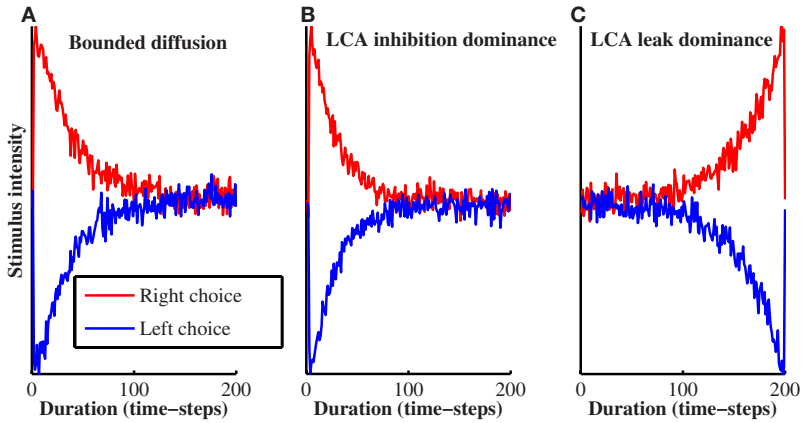


FIGURE 4 | Reverse correlations for the bounded diffusion (A) and the leaky competing accumulators in inhibition-dominance (B), leak-dominance (C) models, showing the average input of the winning (red) and losing (blue) units, in zero-coherence simulated trials.

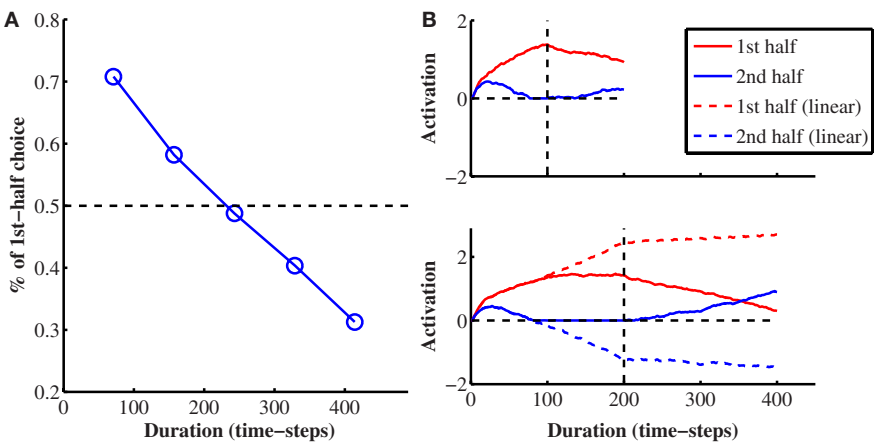


FIGURE 5 | The transition from primacy-to-recency as stimulus duration increases. (A) Probability of choices toward the alternative supported in the first half of the trial in the switch condition. See text for parameter values. **(B)** Activity trajectories of the two accumulators when stimulus duration is short (top) and long (bottom). Red denotes the alternative supported in the first half of the trial, while blue denotes the alternative supported in the second half. In the bottom panel, we also plotted out the simulation results using the linear LCA (dashed lines).

distinctive signature of the non-linear inhibition-dominant LCA is the transition from primacy at short durations to recency at long durations with some parameter settings. In the following section, we report the experimental findings of our studies, considering whether they exhibit features consistent with the greater flexibility of the LCA.

EXPERIMENT 1A

The experiment followed the procedures used in Kiani et al. (2008), as described in Section "Materials and Methods." Observers were asked to determine the predominant direction of moving dots. While some dots were moving randomly, some were moving coherently either to the left or to the right. As in Kiani et al., we used four coherence levels and exponentially distributed stimulus durations in the range 150–1750 ms. Participants were trained to respond within a window of 300 ms following onset of the go cue in order to earn points. The critical manipulation of the evidence was applied in a subset of trials (80% of the trials with durations 300 ms or longer), in the form of a 200 ms "motion pulse" corresponding to a change in coherence of ±3.2%.

All of the observers learned to respond within the 300 ms response window and their accuracy increased with motion coherence according to a sigmoidal function (results from participant CS are shown in **Figure 6A**). As in Kiani et al. (2008), the pulse resulted in a shift of the psychometric curves. A logistic regression was performed to measure the size of the horizontal shift in units of coherence (see also Equation 4 in Kiani et al., 2008). Of special interest is that the effect size of the shift dropped as the pulse was applied later in the trial, indicating that early information has a larger effect on choices. To quantify this, we considered trials with durations of 700 ms or more, and divided the trials into three quantiles according to the time of the pulse (**Figure 6C**).

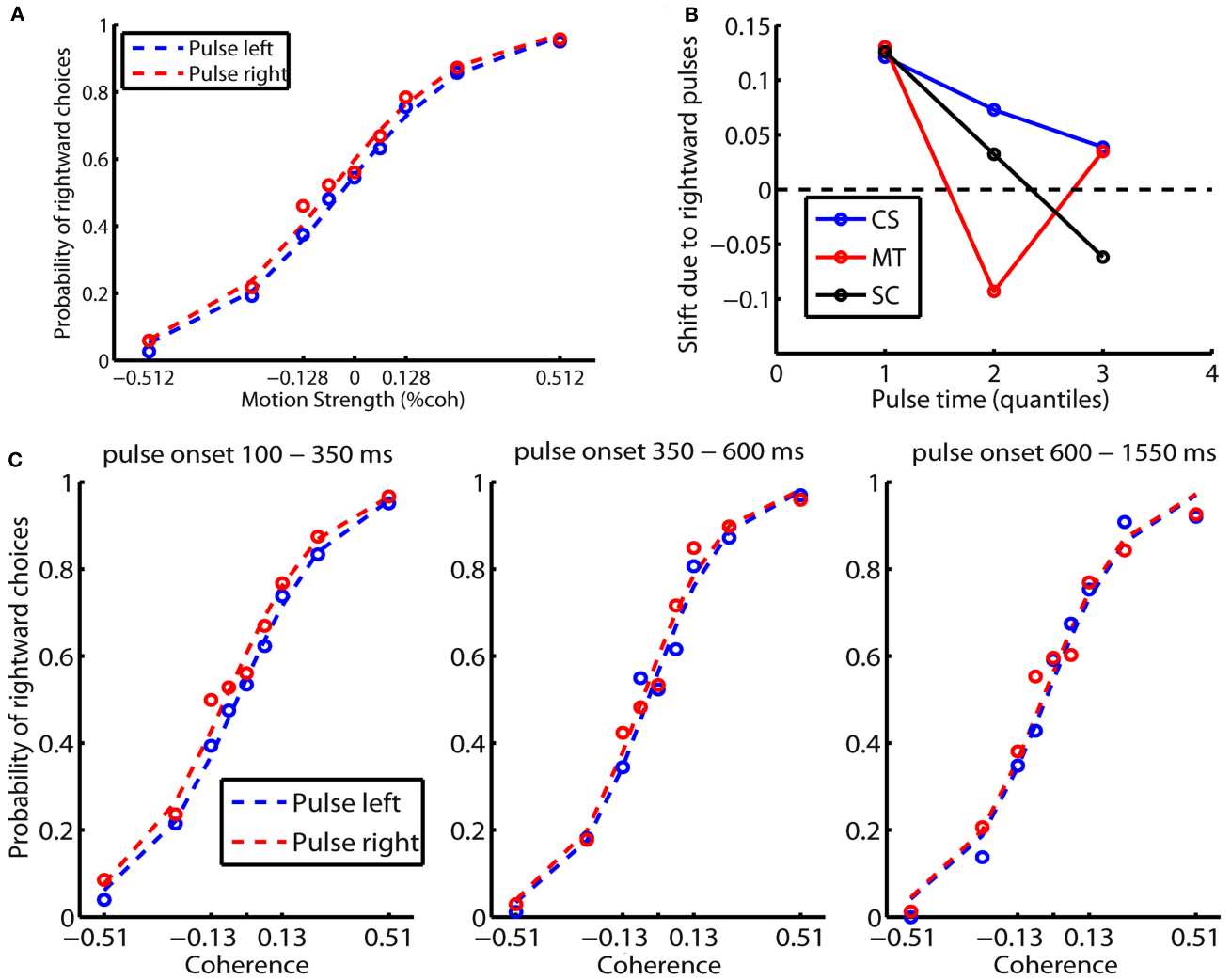


FIGURE 6 | Results of experiment 1A with pulse perturbations. (A) The pulse results in a shift in the observer's psychometric function. As an example, the result of CS is shown here. The percentage of rightward choices is plotted against rightward coherence levels. The red curve denotes rightward pulse and the blue curve denotes leftward pulse. **(B,C)** The effect size of the shift varies with the time of the pulse. **(C)** The psychometric function of participant CS as the pulse comes early, intermediate and late in the trial. **(B)** The effect size of the pulse in the unit of the equivalent motion coherence as a function of pulse time for all of the three observers.

Figure 6B shows that this primacy pattern was present in all participants, with some variability as the effect of the pulse weakens in later quantiles.

EXPERIMENT 1B

Because the effect size in Experiment 1A is very small and therefore difficult to quantify, Experiment 1B was carried out in order to obtain a more robust measure of the temporal weighting profile. To do so, for each coherence and duration combination we created four conditions: (i) the *constant* condition, in which coherence stays fixed throughout the entire trial, (ii) the *early* condition, in which coherence is a fixed non-zero value during the first half of the trial and zero during the second half, (iii) the *late* condition, in which the coherence is zero in the first half of the trial and is a fixed non-zero value during the second half, and (iv) the *switch* condition, in which the coherence stays constant in magnitude but the direction of motion switches in the middle of the trial. The switch condition occurred only with the two low coherence levels to minimize the possibility of participants noticing the switch in motion direction. It is expected that the constant condition will result in higher choice accuracy, as it contains twice as much information as the early/late conditions. There are two critical tests. The first one is the accuracy level in the early condition relative to that in the late condition; and the second is the choice preference toward the alternative supported in the first half relative to that in the second half in the switch condition. A primacy pattern means higher accuracy in the early condition than in the late condition, and more choices toward the alternative supported in the first half

of the trial. Recency means the opposite. The observations are shown in **Figure 7**.

In **Figure 7A**, the accuracy averaged across coherence levels is displayed as a function of stimulus duration for the constant (blue), early (black), and late (red) conditions for the three observers. The results are also fitted by the LCA (left panel) and the BD (right panel) models. In all observers, accuracy increases with stimulus duration and the accuracy in the constant condition is higher than in the early and late conditions. More importantly, accuracy in the early condition is higher than in the late condition, implying a primacy effect. The size of the accuracy difference in the two conditions, however, varies among the three observers. It is very large in one of them (MT), who completely neglected late evidence except in the shortest lag condition, but is smaller in the other two. In SC, this difference also declines as stimulus duration lengthens. The interaction between the recency-primacy pattern and stimulus duration was consistent with the non-linear LCA model, but it provides a challenge to the BD model as shown below.

Quantitative measures of goodness of fit are shown for the LCA and BD models in **Table 1**. We used BIC, which takes the number of degrees of freedom into account, to measure the goodness of fit. BD and LCA fit the data of CS and MT equally well, while LCA provides a better fit to the data from SC – the participant who showed an interaction between the primacy effect and stimulus duration.

In **Figure 7B**, we plotted choice probabilities toward the alternative supported in the first half of the trial. A value above 0.5 means primacy, while a value below 0.5 means recency. Consistent with the results in the early/late conditions, the switch condition also reveals a clear primacy in CS and MT, and this effect is particularly strong in MT. For SC, we see a primacy pattern when stimulus duration is short, and it disappears and even reverses to a recency pattern as the stimulus duration lengthens. Due to its smaller data size, we did not use the switch condition in model fitting. Rather, we adopted the parameters from the fitting of the constant/early/late condition and plotted the model predictions in the switch condition (solid lines in **Figure 7B**). Again, both models fit the first two participants about equally well, but BD does not fit the data of SC as well as LCA does.

EXPERIMENTS 2A AND 2B

Both versions of Experiment 1 replicate the primacy bias reported by Kiani et al. (2008). Since the results of Experiment 1B and the data fitting we conducted showed that it was not possible in two of the three participants to discriminate the two models using fits of the data, we chose in our second set of experiments to focus on the detection of the qualitative pattern of data that can discriminate the models (**Figure 5**). While this pattern only arises at a particular set of LCA parameters, it is special because it goes against what a BD model *can* predict. In particular, we wished to examine whether

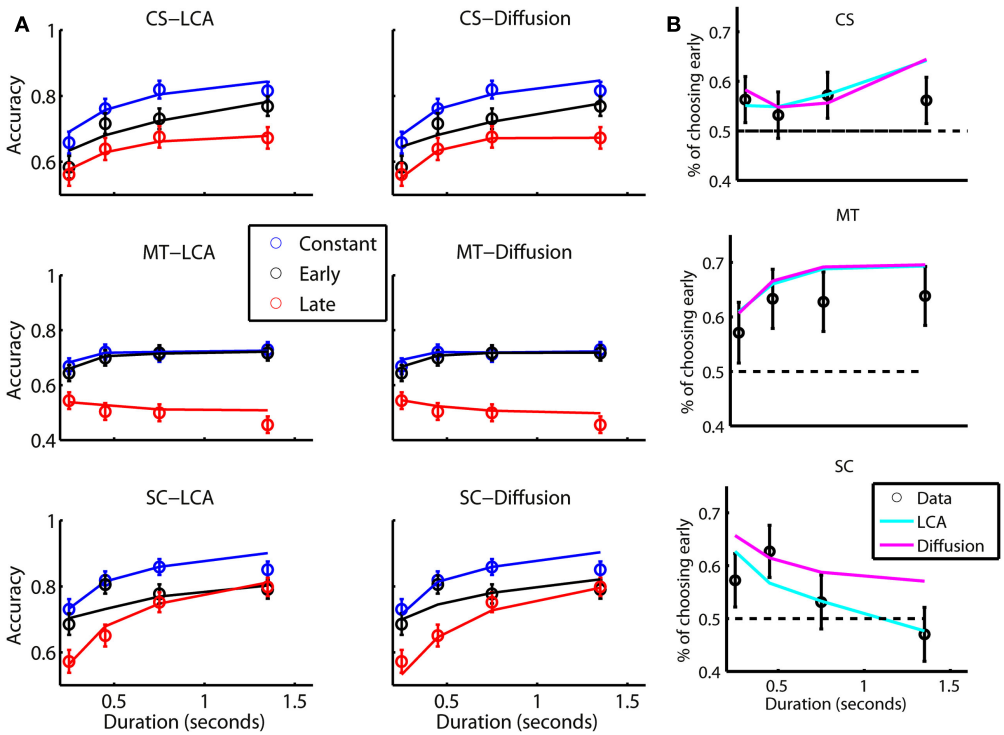


FIGURE 7 | Results of experiment 1B. (A) Accuracy as a function of stimulus duration in the constant, early and late conditions. Left: Data (symbols) and the leaky competing accumulator fit (lines). Error bars correspond to 95% CI. Right: Data and bounded diffusion fit. **(B)** Predictions of the leaky competing accumulator (cyan) and bounded diffusion (magenta) in the switch condition. Parameters of the models are from the fitting in Panel A. Proportion of choices supported early in the trial was plotted against stimulus duration. Error bars correspond to 95% CI. Larger error bars are due to smaller sample size in the switch condition.

Table 1 | Bayesian information criterion values and model parameters for the LCA and BD models for the three subjects in experiment 1B.

Participant	BIC values				Parameters						
	C-E-L		SW		LCA				BD		
	LCA	BD	LCA	BD	β	*K*	*s*	T_0	*A*	*s*	T_0
CS	355.5	355.0	99.3	97.6	0.037	0.034	0.05	12.61	2.50	0.06	17.09
MT	349.5	348.4	78.5	78.7	0.129	0.089	0.06	15.72	0.69	0.06	13.53
SC	350.7	397.7	92.9	155.7	0.028	0.029	0.08	20.19	5.76	0.08	27.33

C-E-L stands for constant-early-late conditions and SW for the switch condition. T_0 values correspond to simulation time-steps.

any of the observers show a transition from primacy-to-recency, which is a signature prediction of the non-linear LCA model and is a challenge to the BD model.

A further goal of our second experiment is to examine if the primacy bias observed in Experiment 1 can be reversed or attenuated. Although the primacy bias seems to be a robust observation (Kiani et al., 2008), it is possible that it may be task-dependent. The time pressure in Experiment 1, is very high, to an extent that is similar to, and perhaps even more extreme than that in Kiani et al. (see text footnote 5). Under such circumstances, decision makers presumably need to be ready to make a prompt response when the go cue comes; this could promote either a lower decision bound for the BD model or stronger lateral inhibition in the LCA.

In order to investigate this question, we relaxed the time pressure in our remaining experiments. First, we relaxed the response window after the go cue from 300 ms to 1 s. This allowed observers enough time to prepare their response after the go cue. Second, we used uniformly distributed stimulus durations instead of exponentially distributed durations. This way, long stimulus durations are equally likely as short stimulus durations (see Discussion for further consideration of this issue).

As in Experiment 1, each participant was tested for several sessions to provide statistical power (see Materials and Methods). In total seven observers were tested with this procedure. The first four participants were tested in Experiment 2A with stimulus durations of 150–1750 ms. After noticing that their accuracy levels differed dramatically, we adapted the difficulty level individually and employed a wider range of stimulus durations (150–2350 ms) for another three participants.

The results are summarized in **Figure 8A**. The average *primacy* score, defined as the average accuracy level in the early condition minus that in the late condition, drops dramatically from Experiment 1B to Experiment 2A and 2B. Since there is no significant difference in procedure or results between the participants in 2A and 2B, we collapse these two groups into one, and refer to this as Experiment 2. The primacy score was significantly larger in Experiment 1B than in Experiment 2 [11 vs. 2%; $t(8) = 2.98$; $p < 0.02$]. While all the observers in the Experiment 1B showed the primacy effect, there was considerable variation among observers in the second group. We therefore conducted a subject-by-subject ANOVA on the main effect of early vs. late and on the interaction between the size of this effect and the stimulus duration. To carry out this analysis, we divided the data of each observer into mini-sessions or *quasi-subjects* that corresponded to all of the session-by-coherence combinations. Each such quasi-subject contributed an equal number of trials to the relevant dependent variables of duration and condition (early vs. late), factoring out the common variability related to fatigue, practice, or performance levels. We thus subjected the mini-session data to a repeated-measure (4×2) ANOVA, with 4 levels of trial duration and 2 levels of timing within trial (early vs. late). The ANOVA results are summarized in **Table 2**.

Table 2 revealed that only two of the seven observers (LK and CB) showed a significant main effect of primacy. More interestingly, participant WW showed a significant interaction between temporal weighting and duration (**Figure 8B**). WW's decisions were mainly driven by early information when stimulus duration was short, while they were driven by late information when stimulus duration was long. This transition from primacy-to-recency is a signature of the non-linear LCA model and it is not consistent with the BD model. Please refer to the Appendix for detailed individual data for all seven participants (Figure A3).

DISCUSSION

Stimulated by the recent study of Kiani et al. (2008), we have examined the temporal weighting of evidence in decision making using a time-controlled protocol. In both of the tested monkeys, Kiani et al. found a primacy bias – early information was more important in decision making – and they proposed the BD model as the mechanistic basis for this observation. According to this model, observers make a decision when a decision bound is reached and ignore any information afterward. In Experiment 1 we examined two types of evidence manipulations: brief motion pulses (or perturbations; see also Huk and Shadlen, 2005) and larger within trial evidence changes at the middle of the stimulus duration. The two methods gave similar results, indicating primacy, though the effect of the latter manipulation was more robust. In our first pair of experiments (1A and 1B), we used a procedure with high time pressure, similar to Kiani et al. In the second pair of experiments (2A and 2B), we relaxed the time pressure by allowing slower responses after the go cue and by using relatively more long trials. Experiments 1A and 1B replicated the primacy bias reported by Kiani et al. while in Experiments 2A and 2B the primacy bias significantly diminished. With some participants, we also found that primacy bias drops, or even transitions to recency (with a stronger weight to *late* evidence relative to early evidence) as stimulus duration lengthens. We showed that the LCA model can account for the

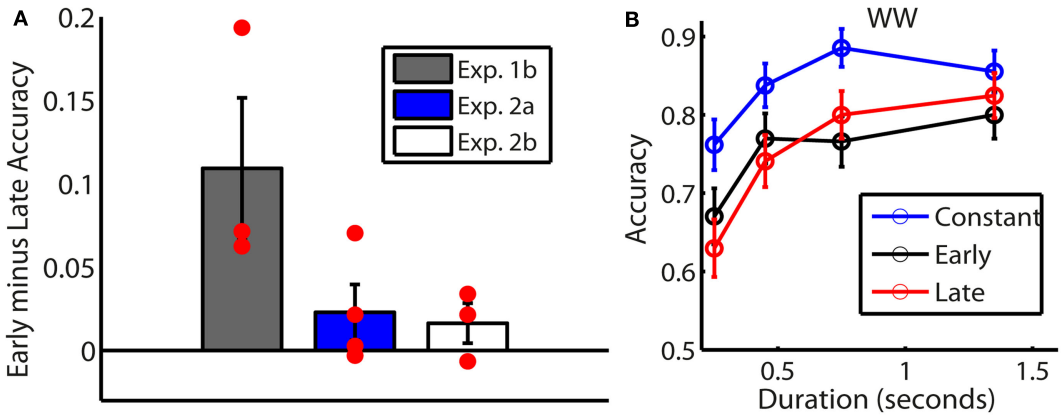


FIGURE 8 | Results of experiment 2. (A) Primacy scores of all the participants in experiments 2A and 2B in comparison to those in Experiment 1B. The primacy score is calculated as the accuracy in the early minus that in the late condition, averaged across all coherences and all stimulus durations. Red circles correspond to individual data. Error bars correspond to 1 SE. **(B)** Accuracy as a function of stimulus duration for participant WW in Experiment 2A. Constant, early, and late conditions are shown in blue, black, and red respectively. Error bars correspond to 95% CI.

primacy bias as well as the BD model, and that it can also capture the transition from primacy to recency, a pattern that poses a challenge to the BD model.

The LCA model does not assume the presence of a decision bound in the time-controlled paradigm. In this model, accuracy levels off due to the imbalance between the leak and the inhibition, and the time scale of this process is determined by the *absolute value* of the difference between the strength of the leak and the strength of the inhibition. The *sign* of this difference, although it does not affect the overall time-accuracy profile, has a profound effect on the relative weight of early vs. late evidence (Usher and McClelland, 2001; Gao et al., 2011). Unlike in the leak-dominant LCA, which gives a higher weight to late evidence, the inhibition-dominant LCA gives a higher weight to early evidence. Thus, this framework, as well as the attractor model (Wang, 2002)[4], provides an alternative to the BD model's account of the primacy pattern. The LCA and related models are also consistent with aspects of the results of an earlier perturbation study by Huk and Shadlen (2005). In this study, the effect of a transient change in evidence on activity in putative evidence accumulation neurons in area LIP is higher when applied early during the observation interval, and becomes very weak near the end (Huk and Shadlen, 2005; Figure 10B). The authors attempted to fit these results using the BD model and noted that it did not explain the very weak impact of later pulses on the neuron's responses (p. 3027). These authors suggested the attractor model of Wang (2002) as one mechanism that could account for the residual effect. In **Figure 9**, we present an informal simulation showing that the inhibition-dominant LCA can also capture the pattern Huk and Shadlen (2005) found in their data. Like neurons in LIP, the accumulators in the LCA are highly sensitive to motion pulses occurring early in a stimulus presentation period, and this effect becomes progressively weaker as integration time continues.

[4]The attractor model was not directly simulated in relation to the temporal weighting of evidence, but we expect it to have similar predictions as the inhibition-dominant *LCA*, as both have *unstable Ornstein–Uhlenbeck dynamics*.

The main result of Experiment 2 was a reduction in the primacy bias, compared to Experiment 1. This difference in the temporal weighting of evidence can be understood in relation to two procedural differences between the two experiments. The first change is that the response window was relaxed from 300 to 1000 ms. With a 300 ms response window, participants must be prepared to respond very quickly once the go cue comes. Under the BD model, this time pressure could lead them to adopt a lower decision bound, so that they will be ready to respond when the go cue occurs. Similarly, under the LCA, this time pressure could encourage adjusting the strength of lateral inhibition, since stronger inhibition helps to encourage a difference in the activation of the two accumulators, which may facilitate faster responding (Gao et al., 2011; Gao and McClelland, in preparation). In any case, time pressure may be one factor contributing to the strong primacy pattern observed in our Experiment 1 and in Kiani et al. (2008)[5].

The second experimental change is that we used uniformly distributed stimulus durations rather than exponentially distributed durations. The reason Kiani et al. (2008) used exponentially distributed stimulus durations was to ensure that observers have no information about the time when the go cue would appear. This choice, however, results in much more frequent short trials than long trials. This factor could encourage participants to ensure they are ready to respond early in the trial, a factor that could further encourage a primacy bias. The empirical findings of our study suggest two potential reasons why Kiani et al. found only primacy while the study of Usher and McClelland (2001) found all

[5]We note that in Kiani et al. (2008), a delay period was included in half of the trials after the stimulus offset. However, trials with and without delays were mixed randomly within blocks, making it necessary for the animal to be ready to respond promptly at the termination of the stimulus, which was very brief on many trials. The response window was 500 ms in Kiani et al. as compared with only 300 ms in our Experiment 1. We conducted a small experiment with a 500 ms response window and found that the primacy bias was not distinguishable in the 500 ms and the 300 ms conditions.

Table 2 | Results of ANOVA examining the effect of timing within trial (early vs. late) and its interaction with trial duration for all participants in experiments 2A and 2B.

Subject	Main effect of early/late	Interaction: duration × early/late
DG	$F(1, 63) = 2.715$; $p = 0.104$	$F(3, 63) = 1.353$; $p = 0.259$
LK	$F(1, 75) = 40.527$; $\mathbf{p < 0.001}$	$F(3, 75) = 1.297$; $p = 0.276$
MM	$F(1, 99) = 0.721$; $p = 0.398$	$F(3, 99) = 0.192$; $p = 0.662$
WW	$F(1, 43) = 0.062$; $p = 0.805$	$F(3, 43) = 3.410$; $\mathbf{p = 0.020}$
AP	$F(1, 11) = 0.026$; $p = 0.876$	$F(3, 11) = 0.563$; $p = 0.643$
CB	$F(1, 20) = 6.678$; $\mathbf{p = 0.018}$	$F(3, 20) = 1.365$; $p = 0.262$
SY	$F(1, 20) = 1.815$; $p = 0.193$	$F(3, 20) = 0.55$; $p = 0.650$

The bold fonts highlight the places where the differences were statistically significant

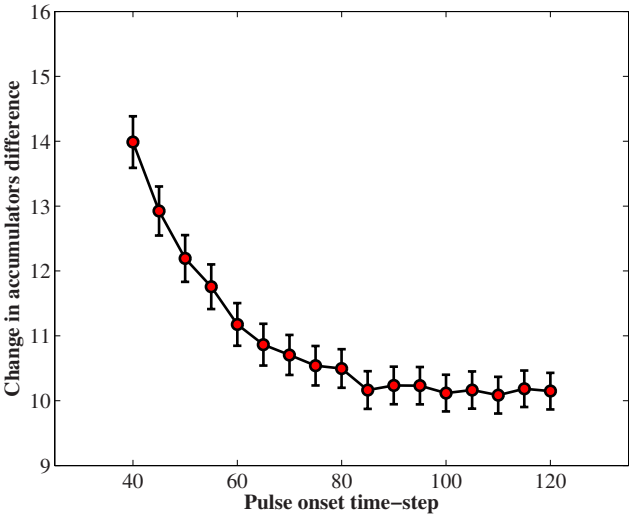


FIGURE 9 | The effect of a short pulse on the activation states of two leaky competing accumulators, at different times in the trial. Each trial lasted for 200 time-steps. The two accumulators received Gaussian inputs with mean and standard deviation both equal to 1. The pulse was inserted for 40 time-steps and increased the mean input to the target accumulator by 1 unit. In the *y*-axis we plot the change that the pulse conferred to the difference between the target and the non-target accumulator. The change is calculated by subtracting the difference between the accumulators' activity 40 time-steps after the offset of the pulse minus the activity difference at the onset of the pulse. The effect diminishes as the time of the pulse onset increases. The leaky competing accumulator model was simulated with inhibition three times larger than the leakage ($\beta = 0.15$, $k = 0.05$). Error bars correspond to 95% CI.

three patterns of primacy, recency, and balanced integration. Like in Experiment 2, participants in the Usher and McClelland study were not presented with predominantly short stimuli, or a short deadline. Our findings also suggest that time pressure, exerted by a narrow response window and/or by more short trials, is one of the factors determining the relative importance of information at different time points.

The results of these experiments also show important individual differences (see also Usher and McClelland, 2001). We were particularly interested in examining whether observers show a transition from primacy, when stimulus duration is short, to recency, when stimulus duration is long. This signature prediction of the inhibition-dominant LCA is challenging for the BD model. Such a transition was found in the performance of subject WW in Experiment 2A, and a similar pattern was found in observer SC in Experiment 1B. Despite detecting the predicted signature of the non-linear LCA, we believe that any conclusions at this stage should be tentative, since they are only supported by the data from 2 of 10 participants.

Further experimentation with additional observers and experimental protocols will be needed to more thoroughly examine the relative merits of the BD and LCA models and to delineate in more detail the conditions under which recency as well as primacy patterns might be obtained. This is important because a number of other experimental paradigms have shown recency patterns (Pietsch and Vickers, 1997; Usher and McClelland, 2001; Newell et al., 2009). Note also that here we only examined temporal weighting of perceptual evidence in a time-controlled paradigm. Although more challenging (since one cannot plan a mid-point evidence change when RT is under subject control), the examination of temporal evidence is also possible in the free-response paradigm. Recently, Zhou et al. (2009) have developed a sophisticated perturbation protocol that can distinguish between a number of competing choice-RT-models in conditions of high signal-to-nose (low error-rate). Future work with such perturbation protocols as well as with balanced or non-balanced evidence switches (e.g., 40% left vs. 60% right) are vital to fully understand the details of the mechanisms of decision making, as are investigations that collect enough data per participant to reliably explore individual differences.

One additional factor that may explain the difference in temporal profile obtained in this study and that in Kiani et al. (2008), compared to studies that showed recency effects is the degree of practice. Practice is quite extensive in our studies as well as in the Kiani study. One possibility, suggested by Brown and Heathcote (2005), is that practice increases the efficiency of evidence accumulation by reducing the effective leak. This factor could play a role in the comparison between our Experiment 1 and 2 as well, since participants in Experiment 1 had more practice, on average, than those in Experiment 2.

Kiani et al. (2008) proposed that bounded integration is a universal decision principle that applies not only to self-paced decisions but also to tasks in which the duration of evidence accumulation is controlled by the experimenter. The results we report here, taken together with other studies showing recency effects, suggest that this conclusion should be reconsidered. Interestingly, one of the motivations suggested by Kiani and colleagues against leaky integration was the idea that leaky integration might be maladaptive in that it discards some of the evidence. While this may be true in some conditions, it is also true that placing a bound on information integration also disregards important decision-relevant information[6]. It might be supposed that unbounded

[6]We do not argue against the idea that decision boundaries are sometimes used even when the stimulus duration is experimentally controlled (Ratcliff, 2006). However,

integration (achieved in the drift-diffusion model without a bound or by a linear version of the LCA with a perfect balance between leak and inhibition) would always be the best policy, but this may ignore important contingencies that could make a recency vs. a primacy strategy more adaptive. These contingencies include the need to be ready to respond quickly and the need to be sensitive to a change in evidence as well as other factors.

We propose that the mechanism in play in the non-linear inhibition-dominant LCA has the advantage of prioritizing early information in a flexible and *reversible* manner. Interestingly, while the non-linearity reduces the optimality of the model in choices between two alternatives, it has the advantage of making the mechanism more optimal and robust when there is a larger number of alternatives (Bogacz et al., 2007). In other work in our labs, this mechanism is supported by data showing that responses triggered by a go cue are faster for correct than incorrect choices (Gao and McClelland, in preparation) and also by decision biases in favor of alternatives whose evidence is temporally anti-correlated with evidence for other alternatives (Tsetsos et al., 2011). Yet other work indicates that some participants exhibit the bimodal decision states like these exhibited by the inhibition-dominant LCA (as illustrated in **Figure 3**, right column; Lachter et al., 2011).

we suspect that such boundaries should be under subject control, and reflect a variety of experimental demands (such as speed-accuracy trade-offs) and contingencies (such as information about expected stimulus durations). Additionally, the bound should be soft rather than rigid.

In closing, we suggest that the principles that are at play in the LCA – leaky integration and lateral inhibition – may generalize beyond the domain of evidence based decisions that we have focused on here. These principles, inspired by known properties of neural systems (Usher and McClelland, 2001), are also found in the attractor model of Wang (2002), and in models based on Decision Field Theory, an approach that has been successfully applied to various aspects of preference based decisions, such as risky choice (Busemeyer and Townsend, 1993; Johnson and Busemeyer, 2005), and to several distinctive characteristics of performance in multi-attribute, multi-alternative decisions (Roe et al., 2001; Usher and McClelland, 2004; Tsetsos et al., 2010).

AUTHOR CONTRIBUTIONS

Juan Gao and James L. McClelland designed and performed the experiments. Marius Usher, Konstantinos Tsetsos, Juan Gao, and James L. McClelland developed theoretical ideas. Konstantinos Tsetsos, Marius Usher, and Juan Gao analyzed the data and conducted model simulations. Marius Usher, Konstantinos Tsetsos, Juan Gao, and James L. McClelland wrote the paper.

ACKNOWLEDGMENTS

This research was supported by Air Force Research Laboratory Grant FA9550-07-1-0537. We thank the reviewers for helpful comments and we thank Jenica Law for proofreading the manuscript.

REFERENCES

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., and Cohen, J. D. (2006). The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. *Psychol. Rev.* 113, 700–765.

Bogacz, R., and Cohen, J. D. (2004). Parameterization of connectionist models. *Behav. Res. Methods* 36, 732–741.

Bogacz, R., Usher, M., Zhang, J. X., 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.

Brown, S. D., and Heathcote, A. (2005). Practice increases the efficiency of evidence accumulation in perceptual choice. *J. Exp. Psychol. Hum. Percept. Perform.* 31, 289–298.

Burr, D. C., and Santoro, L. (2001). Temporal integration of optic flow, measured by contrast and coherence thresholds. *Vision Res.* 41, 1891–1899.

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

Ditterich, J. (2010). A comparison between mechanisms of multi-alternative perceptual decision making: ability to explain human behavior, predictions for neurophysiology, and relationship with decision theory. *Front. Neurosci.* 4:184. doi:10.3389/fnins.2010.00184

Gao, J., Tortell, R., and McClelland, J. L. (2011). Dynamic integration of reward and stimulus information in perceptual decision-making. *PLoS ONE* 6, e16749. doi:10.1371/journal.pone.0016749

Gold, J. I., and Shadlen, M. N. (2000). Representation of a perceptual decision in developing oculomotor commands. *Nature* 404, 390–394.

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

Gold, J. I., and Shadlen, M. N. (2002). Banburismus and the brain: decoding the relationship between sensory stimuli, decisions, and reward. *Neuron* 36, 299–308.

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

Hanes, D. P., and Schall, J. D. (1996). Neural control of voluntary movement initiation. *Science* 274, 427–430.

Horwitz, G. D., and Newsome, W. T. (2001). Target selection for saccadic eye movements: prelude activity in the superior colliculus during a direction-discrimination task. *J. Neurophysiol.* 86, 2543–2558.

Huk, A. C., and Shadlen, M. N. (2005). Neural activity in macaque parietal cortex reflects temporal integration of visual motion signals during perceptual decision making. *J. Neurosci.* 25, 10420–10436.

Johnson, J. G., and Busemeyer, J. R. (2005). A dynamic, computational model of preference reversal phenomena. *Psychol. Rev.* 112, 841–861.

Kiani, R., Hanks, T. D., and Shadlen, M. N. (2008). Bounded integration in parietal cortex underlies decisions even when viewing duration is dictated by the environment. *J. Neurosci.* 28, 3017–3029.

Lachter, J., Corrado, G. S., Johnston, J. C., and McClelland, L. L. (2011). "Distribution of confidence ratings for a simple perceptual task," in *the 52nd Annual Meeting of the Psychonomic Society*, Seattle.

Mazurek, M. E., Roitman, J. D., Ditterich, J., and Shadlen, M. N. (2003). A role for neural integrators in perceptual decision making. *Cereb. Cortex* 13, 1257–1269.

Newell, B. R., Wong, K. Y., Cheung, J. C. H., and Rakow, T. (2009). Think, blink or sleep on it? The impact of modes of thought on complex decision making. *Q. J. Exp. Psychol.* 62, 707–732.

Pietsch, A. J., and Vickers, D. (1997). Memory capacity and intelligence: novel techniques for evaluating rival models of a fundamental information-processing mechanism. *J. Gen. Psychol.* 199, 229–339.

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

Ratcliff, R. (2006). Modeling response signal and response time data. *Cogn. Psychol.* 53, 195–237.

Ratcliff, R., Hasegawa, Y. T., Hasegawa, R. P., Smith, P. L., and Segraves, M. A. (2007). Dual diffusion model for single-cell recording data from the superior colliculus in a brightness-discrimination task. *J. Neurophysiol.* 97, 1756–1774.

Ratcliff, R., and McKoon, G. (2008). The diffusion decision model: theory and data for two-choice decision tasks. *Neural Comput.* 20, 873–922.

Ratcliff, R., and Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychol. Sci.* 9, 347–356.

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

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.

Roitman, J. D., and Shadlen, M. N. (2002). Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *J. Neurosci.* 22, 9475–9489.

Rorie, A. E., Gao, J., McClelland, J. L., and Newsome, W. T. (2010). Integration of sensory and reward information during perceptual decision-making in lateral intraparietal cortex (LIP) of the macaque monkey. *PLoS ONE* 5, e9308. doi:10.1371/journal.pone.0009308

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.

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

Tsetsos, K., Usher, M., and McClelland, J. L. (2011). Testing multi-alternative decision models with non-stationary evidence. *Front. Neurosci.* 5:63. doi:10.3389/fnins.2011.00063

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

Usher, M., and McClelland, J. L. (2004). Loss aversion and inhibition in dynamical models of multialternative choice. *Psychol. Rev.* 111, 757–769.

van Ravenzwaaij, D., van der Maas, H. L. J., and Wagenmakers, E.-J. (2012). Optimal decision making in neural inhibition models. *Psychol. Rev.* 119, 201–215.

Wald, A. (1946). Differentiation under the expectation sign in the fundamental identity of sequential analysis. *Ann. Math. Stat.* 17, 493–497.

Wang, X. J. (2002). Probabilistic decision making by slow reverberation in cortical circuits. *Neuron* 36, 955–968.

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

Wickelgren, W. A. (1977). Speed-accuracy tradeoff and information processing dynamics. *Acta Psychol. (Amst.)* 41, 67–85.

Wong, K. F., Huk, A. C., Shadlen, M. N., and Wang, X. J. (2007). Neural circuit dynamics underlying accumulation of time-varying evidence during perceptual decision making. *Front. Comput. Neurosci.* 1:6. doi:10.3389/neuro.10.006.2007

Wong, K. F., and Wang, X. J. (2006). A recurrent network mechanism of time integration in perceptual decisions. *J. Neurosci.* 26, 1314.

Zhou, X., Wong-Lin, K.-F., and Holmes, P. (2009). Time-varying perturbations can distinguish among integrate-to-threshold models for perceptual decision-making in reaction time tasks. *Neural Comput.* 21, 2336–2362.

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 December 2011; accepted: 11 May 2012; published online: 12 June 2012.

Citation: Tsetsos K, Gao J, McClelland JL and Usher M (2012) Using time-varying evidence to test models of decision dynamics: bounded diffusion vs. the leaky competing accumulator model. Front. Neurosci. ***6****:79. doi: 10.3389/fnins.2012.00079*

This article was submitted to Frontiers in Decision Neuroscience, a specialty of Frontiers in Neuroscience.

APPENDIX

PERTURBATION PROTOCOL IN EXPERIMENT 1A

In experiment 1A, a momentary change (or pulse) in the motion coherence was introduced in 80% of the trials with duration longer than 300 ms. The motion pulse could be inserted between 100 ms after the beginning of the stimulus and 200 ms before it ended. **Figure A1** illustrates the perturbation protocol: T ms after the stimulus onset ($100\,\text{ms} < T < T_{\text{overall}} - 200\,\text{ms}$) the motion coherence which previously equaled to a, increased by p (i.e., coherence level during the pulse was $a + p$, with $p = \pm 3.2\%$). The duration of the pulse was $\Delta T = 200$ ms.

EXCLUDED SESSIONS IN EXPERIMENT 1B

For participants CS and MT the first session was discarded due to a programming mistake (see text footnote 2). Participants CS and SC had high and stable mean accuracy for all sessions (SD of accuracy was 2.2 and 3% respectively) and therefore we used 13 (after excluding the first session) and 12 sessions correspondingly. For participant MT the performance was unstable during the first 10 sessions (see **Figure A2**). These sessions were not included in the analysis, resulting in 9 analyzed sessions (the SD of accuracy for the first 10 sessions was 5%; after excluding these sessions SD was 1.7% for the remaining 9 sessions).

INDIVIDUAL RESULTS FROM EXPERIMENT 2A AND 2B

In **Figure A3** the results of all participants (4 for Experiment 2A and 3 for Experiment 2B) are presented. **Table 2** in the main text shows the statistical analysis performed on each subject regarding the direction of the timing effect (primacy/recency) and its interaction with the trial duration. Participants LK and CB showed a significant primacy while participant WW showed a significant interaction between primacy/recency and trial duration. This interaction is uniquely predicted by the LCA model (see Experiment 1A). The patterns exhibited by the other participants do not discriminate between the models.

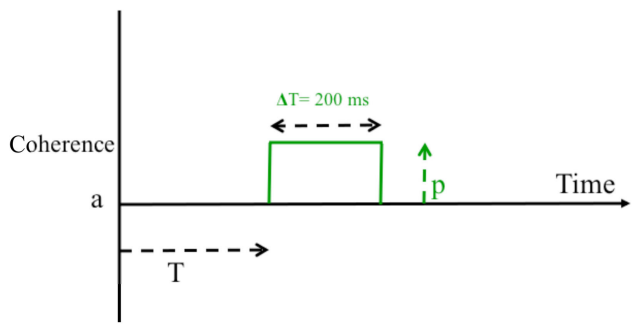


FIGURE A1 | Illustration of perturbation protocol in Experiment 1B. After T ms the motion coherence level changes from a to $a + p$ for a duration of 200 ms.

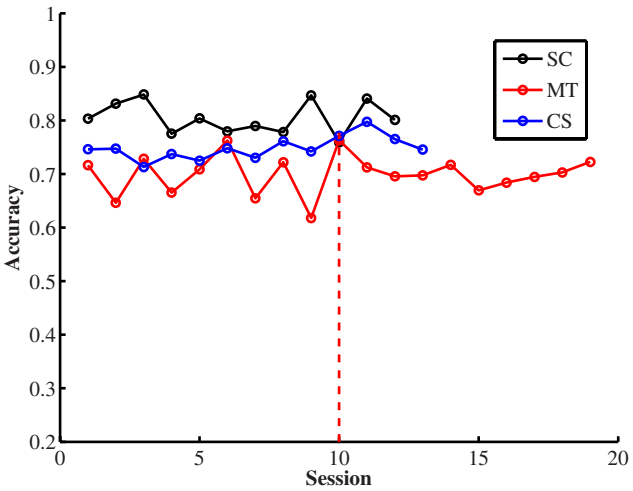


FIGURE A2 | Mean accuracy as a function of session for the three subjects in Experiment 1B. For SC and CS, whose performance was relatively high and stable, all sessions were maintained for the analysis. For subject MT the sessions 1–10 (up to the red vertical line) were eliminated because the participant's accuracy was unstable.

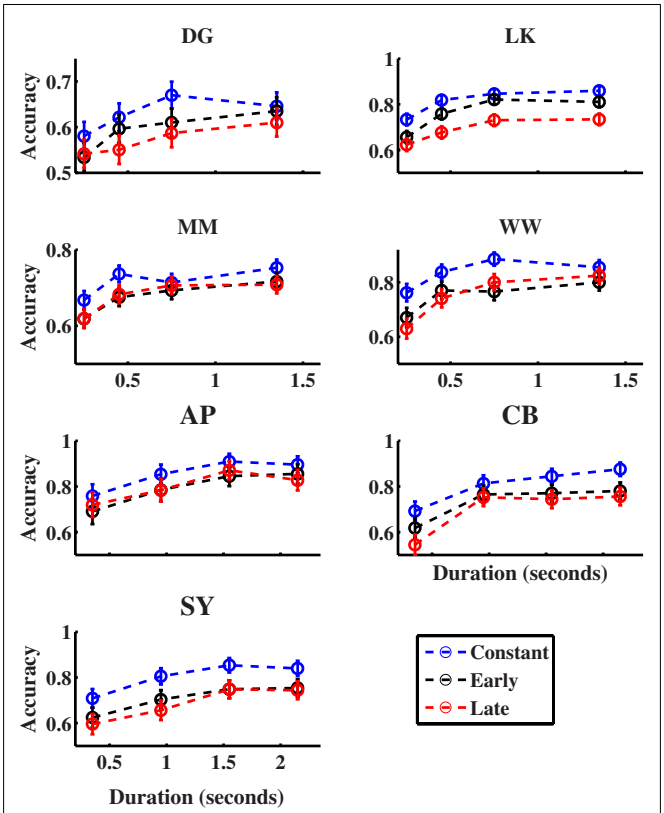


FIGURE A3 | Accuracy as a function of stimulus duration and condition in Experiment 2A (participants DG, LK, MM, and WW) and Experiment 2B (adaptive condition; participants AP, CB, and SY). Error bars correspond to 95% CI.

frontiers in
PSYCHOLOGY

REVIEW ARTICLE
published: 01 August 2012
doi: 10.3389/fpsyg.2012.00263

The effects of evidence bounds on decision-making: theoretical and empirical developments

Jiaxiang Zhang*

Cognition and Brain Sciences Unit, Medical Research Council, Cambridge, UK

Edited by:
Konstantinos Tsetsos, Oxford University, UK

Reviewed by:
Andrew Heathcote, The Newcastle Cognition Lab, Australia
Philip Smith, University of Melbourne, Australia
Andrei Teodorescu, Tel-Aviv University, Israel

****Correspondence:***
Jiaxiang Zhang, Cognition and Brain Sciences Unit, Medical Research Council, 15 Chaucer Road, Cambridge CB2 7EF, UK.
e-mail: jiaxiang.zhang@mrc-cbu.cam.ac.uk

Converging findings from behavioral, neurophysiological, and neuroimaging studies suggest an integration-to-boundary mechanism governing decision formation and choice selection. This mechanism is supported by sequential sampling models of choice decisions, which can implement statistically optimal decision strategies for selecting between multiple alternative options on the basis of sensory evidence. This review focuses on recent developments in understanding the evidence boundary, an important component of decision-making raised by experimental findings and models. The article starts by reviewing the neurobiology of perceptual decisions and several influential sequential sampling models, in particular the drift-diffusion model, the Ornstein–Uhlenbeck model and the leaky-competing-accumulator model. In the second part, the article examines how the boundary may affect a model's dynamics and performance and to what extent it may improve a model's fits to experimental data. In the third part, the article examines recent findings that support the presence and site of boundaries in the brain. The article considers two questions: (1) whether the boundary is a spontaneous property of neural integrators, or is controlled by dedicated neural circuits; (2) if the boundary is variable, what could be the driving factors behind boundary changes? The review brings together studies using different experimental methods in seeking answers to these questions, highlights psychological and physiological factors that may be associated with the boundary and its changes, and further considers the evidence boundary as a generic mechanism to guide complex behavior.

Keywords: decision, boundary, integration, modeling

NEURAL MECHANISMS OF PERCEPTUAL DECISIONS

Making decisions on the basis of sensory information is a frequent and critical element of human lives. Imagine you are driving toward a traffic light in clear weather. You can easily decide to stop or accelerate depending on the color of the traffic light ahead. When driving in foggy weather, however, since the scene is less visible, it is more difficult to distinguish between the red and green light. You may need longer to make the correct decision, and may sometimes even make a mistake.

This type of process is often referred to as perceptual decision-making (Newsome et al., 1989; Gold and Shadlen, 2001, 2007; Heekeren et al., 2008), which requires one to discriminate sensory attributes from either stationary or dynamic stimuli – such as an illumination with different colors (Yellott, 1971), a geometric shape with different orientations (Swensson, 1972), or a pixel array with different brightness (Ratcliff and Rouder, 1998) – and map the subjective perception onto multiple alternative responses. Laboratory studies of the decision process often employ one of two forced-choice paradigms. In the time-controlled (TC) paradigm, subjects are required to give their response immediately after a decision time set by the experimenter (Yellott, 1971; Swensson, 1972; Dosher, 1976, 1984). In the information-controlled (IC) paradigm, subjects are allowed to respond freely whenever they feel confident, from which subjects' response times (RTs) can be measured as a second dependent variable (Luce, 1986). The neural mechanisms of perceptual decisions have been extensively studied using a prototypical random dot motion (RDM) discrimination task (Britten et al., 1993; Shadlen and Newsome, 2001; Roitman and Shadlen, 2002; Palmer et al., 2005; Churchland et al., 2008; Kiani et al., 2008). The RDM stimulus consists of a dynamic field of moving dots, a proportion of which move coherently in one direction, while the other dots move randomly (**Figure 1**). The task is to decide the direction of coherent motion and respond with an eye movement or a button press. Its difficulty can be manipulated by varying the strength of motion coherence.

Single-unit recordings in trained monkeys performing the RDM task indicate that the formation of perceptual decisions involve distinct neural processes across different brain regions. First, neuronal activity in motion sensitive areas (MT/V5; Maunsell and Van Essen, 1983; Born and Bradley, 2005; Zeki, 2007) are closely related to the statistics of the RDM stimulus (i.e., the motion coherence; Newsome and Pare, 1988; Salzman et al., 1990, 1992; Ditterich et al., 2003), but only weakly correlate with behavioral responses (Britten et al., 1992, 1993, 1996), suggesting that sensory neurons encode noisy, transient, and stimulus dependent evidence to support an alternative (Gold and Shadlen, 2001, 2007). Second, neurons in the lateral intraparietal (LIP) area respond with ramp-like changes, and the rate of change depends on the level of

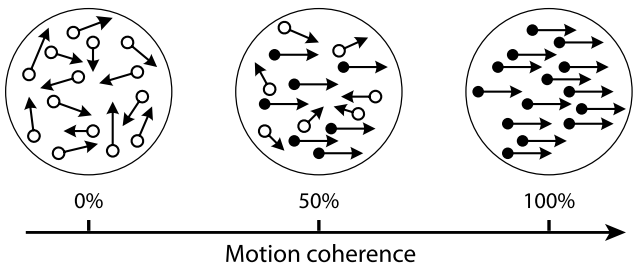


FIGURE 1 | Schematic diagram of the RDM stimulus with different motion coherence levels. In each frame a proportion of the dots (solid dots) are repositioned with fixed spatial offset, indicating the coherent motion direction, and the rest of the dots (open dots) are repositioned randomly. More detailed specification of the stimulus is available in Britten et al. (1992).

motion coherence (Shadlen and Newsome, 2001; Roitman and Shadlen, 2002). Unlike the MT neurons that respond transiently to visual stimuli, the LIP neurons gradually build up or attenuate their activity even if the visual stimuli remain ambiguous (i.e., 0% coherence). This activity pattern starts shortly after the stimulus onset and terminates before a saccadic response. Importantly, around ∼80 ms before a response, there is no obvious variability in firing rates of LIP neurons when responses are made under different motion coherence levels, and neural activity correlates only with the direction of eye movement (i.e., the decision). These findings suggest that LIP neurons integrate sensory evidence up to a decision boundary[1] prior to a response (Mazurek et al., 2003; Huk and Shadlen, 2005; Hanks et al., 2006). Similar activity patterns have also been observed in other brain regions, including frontal eye fields (FEF; Schall, 2002), superior colliculus (SC; Basso and Wurtz, 1998), and dorsolateral prefrontal cortex (DLPFC; Kim and Shadlen, 1999). Taken together, these studies suggest a generic *integration-to-boundary* mechanism manifested in different brain regions for perceptual decisions. That is, certain neuronal populations integrate sensory information over time, and a response is committed to when the accumulated evidence reaches a decision boundary (Schall and Thompson, 1999; Gold and Shadlen, 2001, 2007; Heekeren et al., 2008).

The integration-to-boundary mechanism receives further support from psychological models of choice decisions that have been developed over the last half-century, namely sequential sampling models (Wald, 1947; Lehmann, 1959; Stone, 1960; Link, 1975; Link and Heath, 1975; Townsend and Ashby, 1983; Luce, 1986; Ratcliff and Smith, 2004; Smith and Ratcliff, 2004; Bogacz et al., 2006; Barnard, 2007). Sequential sampling models assume that evidence supporting alternatives are represented by a sequence of noisy observations over time. A process essential to reduce the noise in evidence is to integrate momentary observations over time and make a decision on the basis of the accumulated evidence. The sequential sampling models provide a detailed account of behavioral performance on choice tasks, including RT distributions, response accuracy, and relationships between the two (e.g., the speed–accuracy tradeoff). These models have been widely used as a mechanistic framework for isolating the decision process from sensory inputs or motor outputs.

A key prediction of almost all sequential sampling models is the presence of evidence boundaries, which limit the quantity of evidence available for making a decision. This article reviews recent theoretical and experimental developments in understanding the functions and mechanisms of the evidence boundary. The focus on the boundary mechanisms in general, rather than on particular decision models, is primarily due to its empirical relevance and importance. First, both experimental data and psychological models imply that the evidence boundary does not depend solely on sensory evidence, but can be internally set and controlled by a decision-maker. This unique characteristic of the boundary raises two important questions: (1) how can the evidence boundary influence decision performance? (2) How is the boundary implemented and adapted in neural circuits? Answers to such questions may provide insight into high-level cognitive control that subserves decision-making processes. Second, although the presence of the boundary is consistently supported by the neurophysiological (Mazurek et al., 2003; Huk and Shadlen, 2005; Hanks et al., 2006; Kiani et al., 2008) and neuroimaging (Ploran et al., 2007; Heekeren et al., 2008; Kayser et al., 2010a,b) data, only recently have researchers begun to investigate the function and effects of the evidence boundary. The understanding of its neural mechanisms is still insufficient.

The article is organized as follows: Section "Models of Decision-Making" reviews the decision-making problem and three representative sequential sampling models: the drift-diffusion model (DDM; Ratcliff, 1978), the Ornstein–Uhlenbeck (OU) model (Busemeyer and Townsend, 1993), and the leaky-competing-accumulator (LCA) model (Usher and McClelland, 2001). Section "Theoretical Considerations of Evidence Boundaries" examines the effects of the evidence boundary on the three models. This section discusses how the boundary may affect the models' dynamics and fits to experimental data, and to what extent the boundary may affect the performance of these models. Section "Neural Implementation of Decision Boundary" and "Effects of Boundary Changes" review recent experimental findings that reveal possible neural underpinnings and behavioral influences on the decision boundary. Finally, Section "Discussion" offers some concluding remarks.

MODELS OF DECISION-MAKING

THE DECISION PROBLEM AND THE OPTIMAL DECISION-MAKING THEORIES

Perceptual decision-making can be formalized as a problem of statistical inference (Gold and Shadlen, 2001, 2007). Let us consider a decision task with a choice between N ($N \geq 2$) alternatives, each supported by a population of sensory neurons exclusively selective to a choice (e.g., motion sensitive neurons in area MT/V5). Stimuli drive the N populations of sensory neurons to generate noisy evidence streams $I_i(t)$ at time t, with mean μ_i and variance σ_i^2 ($i = 1, 2, 3, \ldots, N$). The goal of the decision process (e.g., reflected

[1]The term "decision boundary" is referred to the type of evidence boundary that directly affects the termination of the decision. The tem "evidence boundary" is referred to all types of boundaries that limit the accumulation process. See Section "Theoretical Considerations of Evidence Boundaries" for a detailed discussion.

in activity of LIP neurons) is to identify which sensory population has the highest mean activity based on the evidence $I_i(t)$. This article mainly considers three representative models under this framework, as a more complete survey on sequential sampling models is available elsewhere (Ratcliff and Smith, 2004; Smith and Ratcliff, 2004; Bogacz et al., 2006).

Statistically optimal strategies exit for solving the decision problem with two alternatives ($N = 2$), which would achieve the lowest error rates (ER; the probability of making an incorrect choice in a block of trials) and the shortest RT compared with all other decision-making strategies. This optimality criterion can be divided into two sub-criteria (Bogacz et al., 2006): (1) the strategy yielding the lowest ER for any fixed amount of evidence, and (2) the strategy yielding the fastest response for any given ER. The two criteria correspond with the optimal conditions of the TC and IC paradigms, respectively. The optimal strategy for the TC paradigm, i.e., the lowest ER for fixed RT, is provided by the Neyman–Pearson test (NPT; Neyman and Pearson, 1933). The optimal strategy for the IC paradigm, i.e., the fastest RT for a given ER, is provided by the sequential probability ratio test (SPRT; Wald, 1947; Wald and Wolfowitz, 1948; Barnard, 2007). For multiple alternative decision tasks ($N > 2$), asymptotically optimal strategies are also available for the TC (Mcmillen and Holmes, 2006) and IC paradigms (Draglia et al., 1999; Dragalin et al., 2000).

Decision strategies that meet the optimal criteria above require linear integration of evidence over time, which, as reviewed below, can be implemented by many accumulator models on different level of abstraction (the implementation of optimal strategies for multiple alternative decisions requires models with additional complexity to those discussed here, see Bogacz and Gurney, 2007; Zhang and Bogacz, 2010b). Models that can accomplish optimal strategies have been shown to provide better explanations of experimental data than other, non-optimal, models (Ratcliff and Smith, 2004). This leads us to an ecologically motivated assumption that the brain may implement strategies for optimizing the speed and accuracy of decision-making, and hence optimal decision theories may offer a normative benchmark to generate experimental predictions and link behaviors to neural circuits for decision-making (Bogacz, 2007).

The perspective that the brain implements optimal decision-making relies on precise and circumspect definitions of the decision problem and criteria for optimality *per se*. For the simple decision problem with time-invariant evidence, linear integration is the optimal strategy in the sense of its speed and accuracy (see van Ravenzwaaij et al., 2012 for a discussion on other possible definitions of optimality). For tasks with time-varying signal-to-noise ratio within each trial (Huk and Shadlen, 2005; Tsetsos et al., 2011), linear integration may no longer be optimal. Intuitively, if the statistics and regularities of the time-varying evidence (i.e., when more reliable evidence arrives) are known, a decision strategy that exploits such knowledge and gives greater weight to more reliable evidence would have better performance than linear integration strategy (Papoulis, 1977). Whether humans are biased toward early or late evidence, or if their weights of evidence vary with practice (Brown and Heathcote, 2005b), or if their decision strategies are flexibly adapted (Brown et al., 2005), is still not fully understood and merits further investigation.

DRIFT-DIFFUSION MODEL

The DDM was proposed for two-alternative forced-choice (2AFC) tasks (Stone, 1960; Ratcliff, 1978). Mathematically, the DDM can be thought of as a standard Wiener process with external drift (Wiener, 1923), and is equivalent to a continuous limit of the random walk models (Estes, 1955; Laming, 1968; Link, 1975; Link and Heath, 1975; Luce, 1986). The model implies a single integrator that integrates the momentary difference between two sensory streams $[I_1(t) - I_2(t)]$ supporting two alternatives (**Figure 2A**). The dynamics of the DDM can be characterized by a stochastic differential equation:

$$dX(t) = \mu\, dt + \sigma dW(t). \quad (1)$$

Here $dX(t)$ denotes the increment of the accumulated evidence $X(t)$ over a small unit of time dt. The sign of $dX(t)$ implies that the momentary evidence at time t supports the first $[dX(t) > 0]$ or the second $[dX(t) < 0]$ alternative. μ is the drift rate of integration, representing the mean evidence difference $(\mu_1 - \mu_2)$ per unit of time. If $\sigma_1 = \sigma_2$. The magnitude of μ is determined by the quality of the stimulus (the drift rate may be also determined by the allocation of attention, see Schmiedek et al., 2007). For example, for the RDM task, μ would represent the coherence level of the RDM stimulus: a large μ implies high motion coherence and an easy task, while a small μ implies low motion coherence and a high-level of difficulty in distinguishing between two coherent motion directions. The second term $\sigma dW(t)$ denotes Gaussian noise with mean 0 and variance $\sigma^2 dt$. The DDM can be applied to either IC or TC paradigms. In the IC paradigm, decision time is unrestricted and two decision boundaries are introduced to indicate termination states (see Boundary Mechanisms). Once $X(t)$ reaches a boundary, a corresponding choice is made. The predicted RT is equal to the duration of the integration, plus a non-decision time, corresponding to other cognitive processes unrelated to evidence integration (e.g., sensory encoding or response execution).

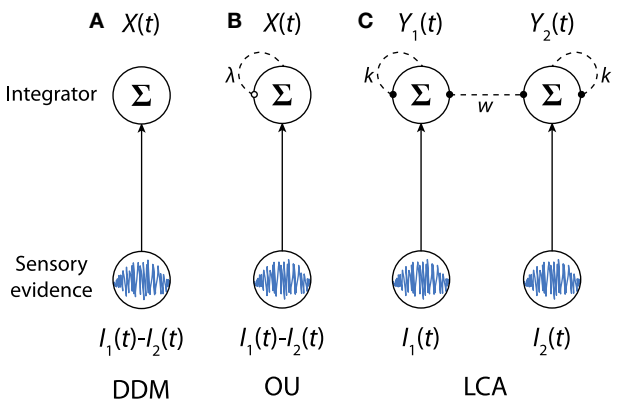


FIGURE 2 | The sequential sampling models for 2AFC tasks: (A) the DDM, (B) the OU model, (C) the LCA model. Arrows denote excitatory connections. Dashed lines with solid circle end denote inhibitory connections. For the OU model, the dashed line with an open circle end denotes the effect of the growth-decay parameter. For each model, the bottom nodes denote sensory evidence, and the top notes denote neural integrators. Model parameters are defined in Eqs 1–3.

For the TC paradigm, which requires subjects to respond at the experimenter-determined decision time T_c, the model selects an alternative by locating the ultimate integrator state $X(T_c)$ and selecting the first alternative if $X(Tc) > 0$, or the second alternative if $X(T_c) < 0$.

Several extensions of the DDM have been proposed since its original introduction, allowing model parameters to vary across trials. First, between-trial variability in the starting point of the integrator $X(0)$ was introduced to account for premature sampling (Laming, 1968), which predicts faster errors than correct responses. Second, between-trial variability in the drift rate was introduced to account for slower errors when compared to correct responses (Ratcliff, 1978). The additional sources of parameter viabilities have been shown to improve fits to experimental data (Ratcliff et al., 1999).

The DDM have been applied to a number of cognitive tasks, including memory retrieval (Ratcliff, 1978), lexical decisions (Ratcliff et al., 2004a; Wagenmakers et al., 2008), letter identification (Ratcliff and Rouder, 2000), and visual discrimination including the brightness discrimination (Ratcliff, 2002; Ratcliff et al., 2003b) and the RDM task (Palmer et al., 2005). In all its applications, the model has successfully accounted for response accuracies and RT distributions observed from individual subjects (Ratcliff and Rouder, 1998; Ratcliff and Smith, 2004; Ratcliff and McKoon, 2008). More importantly, the simple DDM without between-trial parameter variability has been shown to implement the statistically optimal strategies for choosing between two alternatives (the NPT and the SPRT) in both TC and IC paradigms (Wald, 1947; Edwards, 1965; Gold and Shadlen, 2001, 2007; Bogacz et al., 2006), and hence the DDM is often used as a benchmark to compare the performance of other decision models. For the extended version of the DDM, previous studies suggest that the DDM with variable drift rate may still be the optimal model in the TC paradigm but the DDM with variable starting point is not optimal compared to other models (Bogacz et al., 2006). However a strict proof of the optimality of the DDM with between-trial visibilities is still not available yet.

One limitation of the DDM is that it was initially designed for binary choice tasks. Recent studies have attempted to extend the DDM to account for N-alternative forced-choice (NAFC) tasks ($N > 2$). One approach has been suggested by Niwa and Ditterich (2008). For a RDM task with three alternatives (i.e., three possible motion directions), Niwa and Ditterich (2008) modeled three integrators supporting the three alternatives rather than using a single integrator. The three integrators compete against each other in a race toward a common decision boundary and a response is determined by the winning integrator. Crucially, each integrator not only integrates sensory evidence supporting its preferred choice in a diffusion process, but also receives weighted feed-forward inhibition from evidence supporting the other two alternatives (Ditterich, 2010; see also Mazurek et al., 2003 for a similar approach). Churchland et al. (2008) proposed a slightly different approach for modeling a RDM task with four possible motion directions orthogonal to each other. Their hypothesis was that discriminating between two opposite motion directions (e.g., upper-left and lower-right) is independent of sensory evidence supporting the other two orthogonal directions (e.g., lower-left and upper-right). As a result, any sensory evidence supporting the two alternatives neighboring the true alternative was assumed to have a zero mean. The model nicely predicts a feature of their behavioral data that the probability for choosing the alternative directly opposing the true alternative is higher than that for the two alternatives neighboring the true alternative (Churchland et al., 2008). Leite and Ratcliff (2010) examined a family of models with multiple integrators in NAFC tasks with different number of alternatives ($N = 2, 3, 4$). Their results suggest that the models with independent integrators (i.e., no mutual inhibition) and zero to moderate decay produce qualitatively good fits to the RT distributions.

ORNSTEIN–UHLENBECK MODEL

Similar to the DDM, the OU model has been proposed for 2AFC tasks (Busemeyer and Townsend, 1993), and has been applied to a variety of choice tasks to account for response accuracies and RT distributions (Heath, 1992; Diederich, 1995, 1997; Smith, 1995; Busemeyer, 2002). The OU model is identical to the DDM except that it includes a first-order filter that varies the change rate of an integrator (Busemeyer et al., 2006; **Figure 2B**). More precisely, the model is equivalent to a one-dimensional OU process (Uhlenbeck and Ornstein, 1930) and its dynamics can be described by the following differential equation:

$$dX(t) = [\mu + \lambda X(t)]\, dt + \sigma dW(t). \quad (2)$$

The drift rate μ and the noise term $\sigma dW(t)$ have the same definitions as in Eq. 1 (see "Drift-diffusion model" above). The model contains a linear coefficient λ, a growth-decay parameter. As a result the rate of change of $X(t)$ depends not only on the mean drift rate, but also on the current state of the integrator.

The growth-decay parameter brings some interesting properties to the OU model. First, in the TC paradigm, the response accuracy of the OU model reaches an asymptote for a large decision time T_c. Note that the same prediction can be made from the DDM by introducing variability in drift rate across trials (Ratcliff et al., 1999), and that therefore theoretically the two models can account for behavioral data equally well (but, see Ratcliff and Smith, 2004). However, recent studies suggest that the two models are distinguishable by introducing temporal uncertainty to the stimulus (Huk and Shadlen, 2005; Kiani et al., 2008; Zhou et al., 2009). Second, the value of λ can account for the serial position effects observed in decision-making tasks (Wallsten and Barton, 1982; Busemeyer and Townsend, 1993; Usher and McClelland, 2001). For $\lambda < 0$, the linear term $\lambda X(t)$ inhibits the integrator and the evolution of $X(t)$ tends toward a stable attractor $-\mu/\lambda$. Because evidence presented earlier in a trial decays over time, the choice mainly depends on the evidence later in the trial (a recency effect). In contrast, for $\lambda > 0$, the evolution of $X(t)$ is repelled from the unstable fixed point $-\mu/\lambda$, and the speed of repulsion is proportional to the distance between the current stage $X(t)$ and $-\mu/\lambda$. Therefore after $X(t)$ has been driven to one side or other of the fixed point, subsequent evidence has little effect on the final choice due to repulsion (a primacy effect). For $\lambda = 0$, the OU model reduces to the DDM and hence implements the optimal decision strategy.

LEAKY-COMPETING-ACCUMULATOR MODEL

The LCA model was proposed by Usher and McClelland (2001). Unlike the DDM and the OU model which integrate the relative evidence for one alternative compared with another, the LCA model assumes that evidence supporting different alternatives is integrated by separate integrators (**Figure 2C**). Therefore the LCA model can be naturally extended to account for decision tasks with multiple alternatives (Usher and McClelland, 2004; Mcmillen and Holmes, 2006; Tsetsos et al., 2011). Each integrator in the LCA model is leaky, as accumulated information continuously decays, and receives mutual inhibition from other integrators. For 2AFC tasks, the dynamics of the two integrators $Y_1(t)$ and $Y_2(t)$ can be described by:

$$\begin{cases} dY_1(t) = \left(\mu_1 - ky_1(t) - wy_2(t)\right)dt + \sigma dW_1(t) \\ dY_2(t) = \left(\mu_2 - ky_2(t) - wy_1(t)\right)dt + \sigma dW_2(t) \end{cases}. \quad (3)$$

Here k ($k \geq 0$) denotes the rate of decay, and w ($w \geq 0$) denotes the weight of mutual inhibition from the other integrator. In the absence of sensory evidence ($\mu_1 = \mu_2 = 0$), the two integrators will converge to zero due to the effect of decay. The additional mutual inhibition means that the integrators are not independent, as each integrator can access the evidence that supports other alternatives. The LCA model can be applied to both IC and TC paradigms. In the IC paradigm, the first integrator that reaches a decision boundary renders its preferred choice. In the TC paradigm, the decision is determined by identifying which integrator has higher activity at a decision time T_c. The model in Eq. 3 is a simplified linear version of the LCA model and the integrators' values are unconstrained. In their original publication, Usher and McClelland (2001) assumed that the integrators' stages are transformed by using a threshold-linear activation function, which prevents any integrator having negative values (Brown and Holmes, 2001; Brown et al., 2005). This non-linearity is motivated by the fact that activities of neural integrators can never be negative (see Boundary Mechanisms).

The LCA model is closely related to other sequential sampling models. For $w = k = 0$ (no decay or inhibition), the LCA model is equivalent to a model with independent integrators, which resembles a continuous version of the accumulator or counter models (Pike, 1966; Vickers, 1970). For 2AFC tasks, the LCA model can be reduced to an OU model if both decay and inhibition are large relative to the noise strength σ (Bogacz et al., 2006, 2007). The relative difference between w and k determines the growth-decay parameter λ in the reduced OU model ($\lambda = w - k$). That is, if the inhibition is larger than the decay ($w > k$), the LCA model can be reduced to an OU model with $\lambda > 0$. In contrast, if the inhibition is smaller than the decay ($w < k$), the LCA model can be reduced to the OU model with $\lambda < 0$. Therefore, similar to the OU model with $\lambda \neq 0$, the LCA model with unbalanced inhibition and decay ($w \neq k$) can account for primacy and recency effects (Usher and McClelland, 2001). For balanced decay and inhibition ($w = k$), the LCA model can be approximated by the DDM and hence implements the optimal decision strategy.

Because the LCA model can mimic the DDM and the OU model within a certain parameter range, the LCA model retains the strength of the simpler models to account for detailed aspects of behavioral data from 2AFC tasks. The LCA model has also been successfully applied to perceptual decision tasks with multiple alternatives (Usher and McClelland, 2001; Tsetsos et al., 2011), and value-based decisions, in which the decisions are settled on subjective preferences, rather than perceptual information (Usher and McClelland, 2004; Usher et al., 2008).

DECISION-MAKING MODELS AT DIFFERENT LEVELS OF COMPLEXITY

The sequential sampling models do an excellent job of accounting for the variability of responses and RTs in various decision tasks. Over decades researchers have tended to extend existing models to account for more systematic effects (e.g., RT differences between correct and error responses) or more biologically realistic constraints (e.g., the mutual inhibition and decay in the LCA model). These attempts led to an increase of model complexity and number of model parameters, which, in practice, makes such models difficult to apply to experimental data. There are several previous attempts to simplify existing models. For example, Wagenmakers et al. (2007) proposed a simplified version of the DDM by assuming that there is no between-trial variability, and a further simplified DDM proposed by Grasman et al. (2009) additionally assumes the starting point of the integrator is not biased toward any alternative. These simplified models can directly estimate the DDM parameters from analytical solutions without a parameter-fitting procedure.

More recently, Brown and Heathcote (2008) proposed a linear ballistic accumulator (LBA) model of choice decisions (see Brown and Heathcote, 2005a for a non-linear version of the model). The LBA model has been applied to many choice tasks including perceptual discrimination (Forstmann et al., 2008, 2010a,b; Ho et al., 2009), absolute identification (Brown and Heathcote, 2008), lexical decisions (Donkin and Heathcote, 2009), and saccadic eye movements (Ludwig et al., 2009; Farrell et al., 2010). Similar to the LCA model, the LBA model assumes each integrator integrates evidence supporting one alternative and hence can be applied to NAFC tasks, but with two major simplifications. First, the integrators are independent (no mutual inhibition) and have no leakage (no decay). Second, the integration process within each trial is linear and deterministic (i.e., ballistic), omitting the within-trial variability in momentary evidence. These two assumptions greatly simplify the model dynamics and hence the LBA model has analytical solutions for RT distributions and response accuracies for NAFC tasks. This is a significant advantage in terms of computational complexity as one can estimate the model parameters without using Monte Carlo simulations. However, the strong assumptions inevitably introduce limitations. Because the integration process is assumed to be linear and deterministic, the LBA model cannot distinguish evidence arriving at different times over a trial, and hence it is not straightforward to apply the LBA model when accounting for primacy and recency effects, or any task paradigms that deliberately introduce temporal uncertainty within a trial (Usher and McClelland, 2001; Huk and Shadlen, 2005; Tsetsos et al., 2011).

Decision-making models can be used to isolate decision components (e.g., boundary and drift rates), from which estimated model parameters can infer experimental data collected from different sources, such as fMRI or EEG/MEG signals. This model-based approach provides an invaluable way of linking latent

decision processes predicted by the accumulator models with their implementations in large neural populations, and not surprisingly has attracted increasing interest over the last few years (Philiastides et al., 2006; Philiastides and Sajda, 2007; Forstmann et al., 2008, 2010b; Ho et al., 2009; Ratcliff et al., 2009; Kayser et al., 2010a,b; Wenzlaff et al., 2011). It is worth noting that all models can be used for this purpose, although simpler models are often employed due to less computational complexity.

However, models at a highly abstract level (e.g., the DDM and the LBA model) are not sufficient to address some more fundamental questions of decision-making, such as the neural mechanism of slow ramping activity in LIP neurons during RDM tasks, or the mechanisms of decay and inhibition in neural integrators. The answers to these questions require more detailed models at the level of single neurons (the LCA model provides a middle ground in neural plausibility between single neuron models and the DDM). Wang (2002) proposed a biophysically based spiking neuron model for perceptual decision-making. For the RDM task with two alternatives, the model assumes two LIP neural populations supporting each alternative. Instead of mutual inhibition in the LCA model, all neurons from different populations project to a common pool of inhibitory neurons, which then inhibits each population via feedback inhibitory connections. Wang (2002) proposed that evidence integration over a long timescale (on the order of several hundred milliseconds to over 1 s), as assumed by most sequential sampling models, could be realistically carried out by neural populations with recurrent excitatory connections mediated by NMDA receptors at a very short timescale (on the order of less than 100 ms). This model has been demonstrated to successfully account for the activity of LIP neurons as well as behavioral performance in the RDM tasks (Wong and Wang, 2006; Wong et al., 2007), and has recently been applied to multiple alternative decision tasks (Furman and Wang, 2008). However, although the biophysical model is important for understanding the neural mechanisms of decision processes, due to the model complexity and the large number of model parameters it could be difficult to use such a specialized model as an exploratory tool for other decision tasks, or to search through the parameter space to fit the model to RT distributions. Smith and McKenzie (2011) recently proposed a simplified version of Wang's (2002) model that overcomes these difficulties. In their minimal recurrent loop model, evidence is represented by Poisson shot noise processes (Smith, 2010) and evidence integration for each alternative is represented by the superposition of Poisson processes, resembling the essential statistical features of the reverberation loops in Wang's model. The model provides a theoretical account of how diffusive-like evidence integration at an abstract level naturally emerges from the spike densities in the recurrent loops. Further, at a cost of two more free parameters, the minimal recurrent loop model can fit the RT distributions and associated choice accuracies almost equally well as the DDM (Smith and McKenzie, 2011), suggesting that the model offers a promising balance between biological plausibility and generality to predict experimental data. In summary, decision models at different levels of complexity could be useful to capture experimental data obtained from different modalities (**Figure 3**), and empirical researchers should choose an appropriate model that suits their research questions.

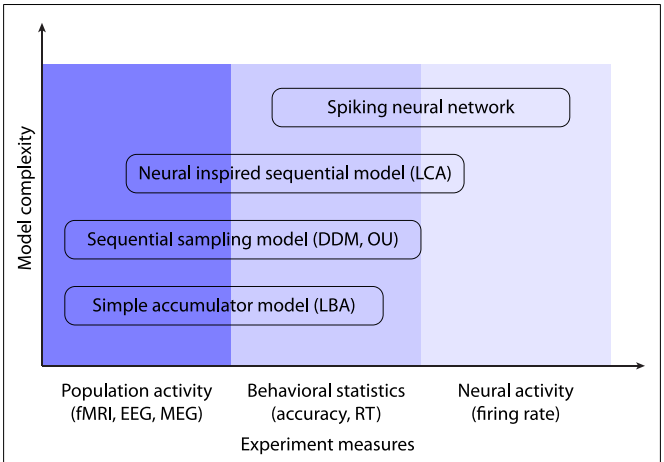


FIGURE 3 | The complexity and generality of the decision-making models. All models are capable of capturing basic behavioral statistics such as the RT and the response accuracy. The simple accumulator models and the sequential sampling models are suitable to describe the congregate activity of large neural populations (e.g., fMRI or EEG/MEG signals). The most complex model (i.e., the spiking neural network) can be used to account for dynamics of neural circuits.

THEORETICAL CONSIDERATIONS OF EVIDENCE BOUNDARIES

BOUNDARY MECHANISMS

All the sequential sampling models discussed above describe a diffusion-like evidence integration during the decision process (Brown and Holmes, 2001; Brown et al., 2005). However they need to be bundled with evidence boundaries that constrain accumulation. This section examines evidence boundaries according to two different but not mutually exclusive definitions: (1) evidence boundaries that determine the amount of accumulated evidence required to make a decision (i.e., the decision boundaries), and (2) evidence boundaries that act as barriers to the amount of accumulated evidence (**Figure 4A**).

The first type of evidence boundary, hereafter referred to as the *absorbing* boundary, provides an evidence criterion or threshold for the termination state of an integration process, and assumes a decision is made once accumulated evidence supporting one alternative reaches the boundary. The absorbing boundary is necessary for modeling tasks that require subjects to implement a self-initiated stopping rule (e.g., in the IC paradigm) and hence it has been widely used by many models in the choice RT modeling literature (Ratcliff, 1988, 2006; Gomez et al., 2007).

The second type of evidence boundary introduces biologically inspired constraints that limit the amount of accumulated evidence. Early decision models did not explicitly constrain activity of integrators (Ratcliff, 1978), which raised theoretical and practical concerns to the validity of the models. The theoretical concern is that unconstrained integrators imply a possibility of an unlimited amount of evidence being maintained by the model (**Figure 4A**). For example, in the TC paradigm, the integrator state of the DDM has infinite mean and variance as T_c approaches infinity (see Eq. 1). For the LCA model, unconstrained integrators further imply the possibility that model activation may become negative

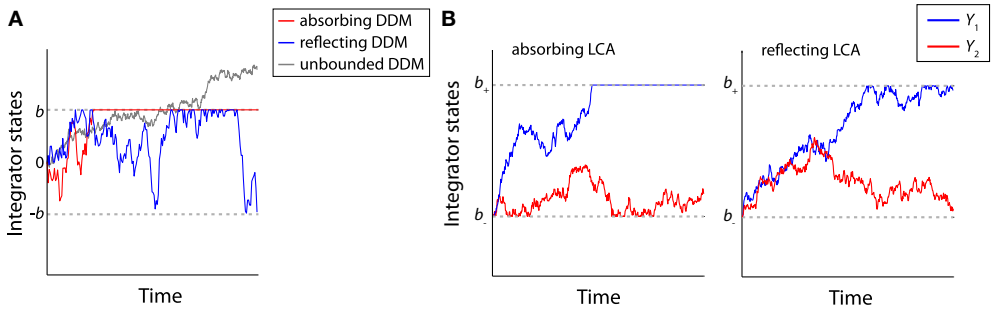


FIGURE 4 | Time course of the integrators of the DDM and LCA model with boundaries. (A) Examples of trajectories of the absorbing (red), reflecting (blue) and unbounded (gray) DDM. Two boundaries ($\pm b$) are indicated by the gray dashed lines. **(B)** Examples of trajectories of the absorbing (left panel) and reflecting (right panel) LCA models. The lower boundary (b_-) and the upper boundary (b_+) boundaries are indicated by the gray dashed lines.

due to mutual inhibition. Unlimited or negative activations are undesirable for a biologically plausible model, because neural integrators cannot exceed certain values due to intrinsic limitations of biological neurons. Their activity should also be non-negative. These constraints need to be satisfied before attempting to extend abstract models to qualitatively account for neural firing rate patterns during the decision process (Usher and McClelland, 2001; Ratcliff et al., 2003a; Huk and Shadlen, 2005; Ditterich, 2006; Purcell et al., 2010).

The practical concern is that models with unconstrained integrators may not fit experimental data well. In the TC paradigm, the ER of the DDM with an unconstrained integrator diminishes to zero for a large decision time T_c (without between-trial variability), and hence the model predicts that subjects can achieve arbitrarily small ER even for difficult tasks. Nevertheless, it is known that humans cannot achieve 100% accuracy even for large T_c (Meyer et al., 1988; Usher and McClelland, 2001). Furthermore, negative activation in the LCA model may result in abnormal model predictions. Bogacz et al. (2007) showed that in a multi-alternative decision task, if the inputs to an LCA model favor only a small subset of possible alternatives, integrators favoring irrelevant choices (i.e., those that do not receive inputs) would become negative and send uninformative positive evidence via mutual inhibition to the relevant competing integrators (i.e., those receiving inputs). As a result the LCA model without truncation of negative activation may select inferior alternatives in value-based decisions (Usher and McClelland, 2004; Usher et al., 2008), and provide qualitatively poorer fits to experimental data than the models with non-negative evidence only (Leite and Ratcliff, 2010). The same problem also exists in models with feed-forward inhibitory connections (van Ravenzwaaij et al., 2012).

One way to introduce constraints is to transform the integrator state through a non-linear activation function (Brown and Holmes, 2001; Usher and McClelland, 2001; Brown et al., 2005), or to assume high-level baseline activity for avoiding non-negative activations (van Ravenzwaaij et al., 2012). A simpler approach, without losing the explicit nature and tractability of a linear system and yet offering a good approximation of the non-linear activation functions, is to introduce explicit evidence boundaries to existing models. This type of boundary is hereafter referred to as the *reflecting boundary* (Diederich, 1995; Bogacz et al., 2007; Zhang et al., 2009; Zhang and Bogacz, 2010a; Smith and McKenzie, 2011). The reflecting boundary only constrains the maximum or minimum amount of evidence that can be presented by an integrator (much as a non-linear activation function provides cutoffs at high or low activations), but unlike the absorbing boundary, reaching a reflecting boundary does not terminate the integration process (**Figure 4A**).

Both types of boundary mechanisms have been applied to various decision models (Ratcliff, 2006; Bogacz et al., 2007; Zhang et al., 2009; Zhang and Bogacz, 2010a; Tsetsos et al., 2011; van Ravenzwaaij et al., 2012). The decision models with boundaries are hereafter referred to as *bounded*, and the models without a boundary as *unbounded*. For the DDM and the OU model, when there is no bias toward either alternative, two symmetric absorbing or reflecting boundaries ($\pm b$) can be imposed to limit the integrator's activity (**Figure 4A**). For simplicity, the terms absorbing DDM and absorbing OU model are used when the two absorbing boundaries apply to the models, and the reflecting DDM and reflecting OU model when referring to models with two reflecting boundaries. For an LCA model with multiple integrators, if one assumes that integrators cannot have arbitrarily large or negative values, then two boundary conditions need to be applied to each integrator (**Figure 4B**). First, each integrator requires one lower boundary b_- at zero to constrain the minimum activity to be non-negative (Bogacz et al., 2007). This lower boundary needs to be a reflecting boundary, since otherwise the model may not render a decision (i.e., if the lower boundary is absorbing, activities of all integrators could be fixed at the boundary). Second, each integrator requires one upper boundary b_+ ($b_+ > 0$) to limit the maximum activity. The upper boundary b_+ could be either absorbing or reflecting. The LCA model with an absorbing boundary at b_+ is referred to as the absorbing LCA model, and the model with a reflecting boundary at b_+ as the reflecting LCA model. **Table 1** summarizes the bounded decision models discussed above and their properties.

It is worth noting that models with absorbing boundaries provide a unified account for both IC and TC paradigms (Ratcliff and McKoon, 2008), because contact with absorbing boundaries induces a decision. In contrast, models with pure reflecting boundaries require an external criterion to stop (e.g., decision deadline T_c), and hence they are only for the TC paradigm but cannot account for the IC paradigm. Although the pure reflecting model

Table 1 | Properties of the sequential sampling models with and without boundaries.

		Primacy	Recency	Optimality	TC paradigm	IC paradigm
DDM	Unbound	–	–	Optimal	✓	✓
	Absorbing	✓	–	–	✓	✓
	Reflecting	–	✓	–	✓	–
OU	Unbound	$\lambda > 0$	$\lambda < 0$	$\lambda = 0$	✓	✓
	Absorbing	Various	$\lambda < 0$	$\lambda < 0$	✓	✓
	Reflecting	$\lambda > 0$	Various	$\lambda > 0$	✓	–
LCA	Unbound	$w > k$	$w < k$	$k = w$	✓	✓
	Lower-bound	$w > k$	$w < k$	Unknown	✓	✓
	Absorbing	Unknown	Unknown	$w < k$	✓	✓
	Reflecting	Unknown	Unknown	$w > k$	✓	–

The lower-bound LCA model refers to the LCA model that has only lower reflecting boundary at zero but no upper boundary.

may be criticized for its lack of generality, it is necessary to consider the models with pure reflecting boundaries together alongside models with absorbing boundaries in order to illustrate some complementary properties of the two types of boundary. First, absorbing boundaries, together with the reflecting boundaries, provide a simple solution for primacy and recency effects in different models (see Primacy and Recency Effects). Second, the two types of boundary could characterize different decision strategies in the TC paradigm (Zhang and Bogacz, 2010a). The absorbing boundary implies that subjects make their choice before the response deadline (i.e., once the absorbing boundary is reached) and withhold their decision. The reflecting boundary implies that subjects continuously hesitate between the choices even when sufficient evidence is available (i.e., when the reflecting boundary is reached) and may change their decision later. Whether subjects adopt one of the two strategies, or are able to switch between the two (see Tsetsos et al., 2012), would be an interesting question for future research.

PRIMACY AND RECENCY EFFECTS

The unbounded DDM integrates evidence independent of the current integrator state (Eq. 1), and hence the model implies that influence of sensory evidence on the final choice does not depend on the timing of its occurrence (i.e., neither primacy nor recency). One recent study suggests that the DDM can account for primacy and recency effects by introducing the two types of boundaries (Zhang et al., 2009). For the absorbing DDM, if a boundary is reached before decision time, the preferred decision is determined and only evidence occurring prior to the boundary hit contributes to the integration process, indicating a primacy effect. For the reflecting DDM, each boundary hit results in a partial loss of evidence, since the integrator does not fully integrate momentary evidence that would otherwise exceed the boundary. As a result, the momentary evidence arriving earlier is partially lost and on average a decision depends to a greater extent on later evidence, indicating a recency effect (**Figure 5A**). A further study indicates that the primacy/recency effects introduced by the two types of boundaries can coincide and interact with the effects introduced by the growth-decay parameter λ in a bounded OU model (Zhang and Bogacz, 2010a). If the boundary and λ provide the same effect, the joint primacy/recency effect of the bounded OU model is maintained. On the contrary, the joint effect of the bounded OU model is weakened or canceled if λ and the boundary present opposite effects (**Figures 5B,C**). For example, for $\lambda > 0$ (primacy effect), an OU model with absorbing boundaries (also the primacy effect) will also exhibit a strong primacy effect, but an OU model with reflecting boundaries will show a weaker effect. There is as yet no study systematically reporting primacy and recency effects in the bounded LCA model. Given the close relationship between LCA model and OU model, one may expect that the primacy/recency effects of bounded LCA model are jointly determined by the type of boundary and the value of inhibition and decay parameters. Recent studies (Tsetsos et al., 2011, 2012) demonstrates that the LCA model with only lower reflecting boundary demonstrates a strong primacy effect when the inhibition is large relative to the decay ($w > k$), and a recency effect when the inhibition is small relative to the decay ($w < k$), consistent with results obtained from the unbounded LCA model.

This section has shown that primacy and recency effects can be readily produced by evidence boundaries or their interactions with other model parameters. Nevertheless, existing experimental data is insufficient to demonstrate the strength of these effects in the way predicted by the models. An ideal paradigm to systematically investigate and differentiate these effects would be a decision task using time-varying evidence, which favors one alternative early in a trial and another alternative later in a trial. However, the interpretation of results from such an experiment would need to proceed cautiously in case of potential confounds. First, if non-stationary stimuli extends for a long period of time (as in the expanded judgment paradigm, see Pietsch and Vickers, 1997), the observed primacy/recency effects may be to some extent associated with additional attention or working memory processes. Second, if non-stationarity in the evidence is apparent to subjects, they may consciously change their decision strategy. Several studies on rapid perceptual decisions avoided these methodological problems by using carefully designed paradigms. Brown and Heathcote (2005b) presented strong prime stimuli for a very short time and used a metacontrast mask to ensure subjects did not consciously aware the non-stationarity. They

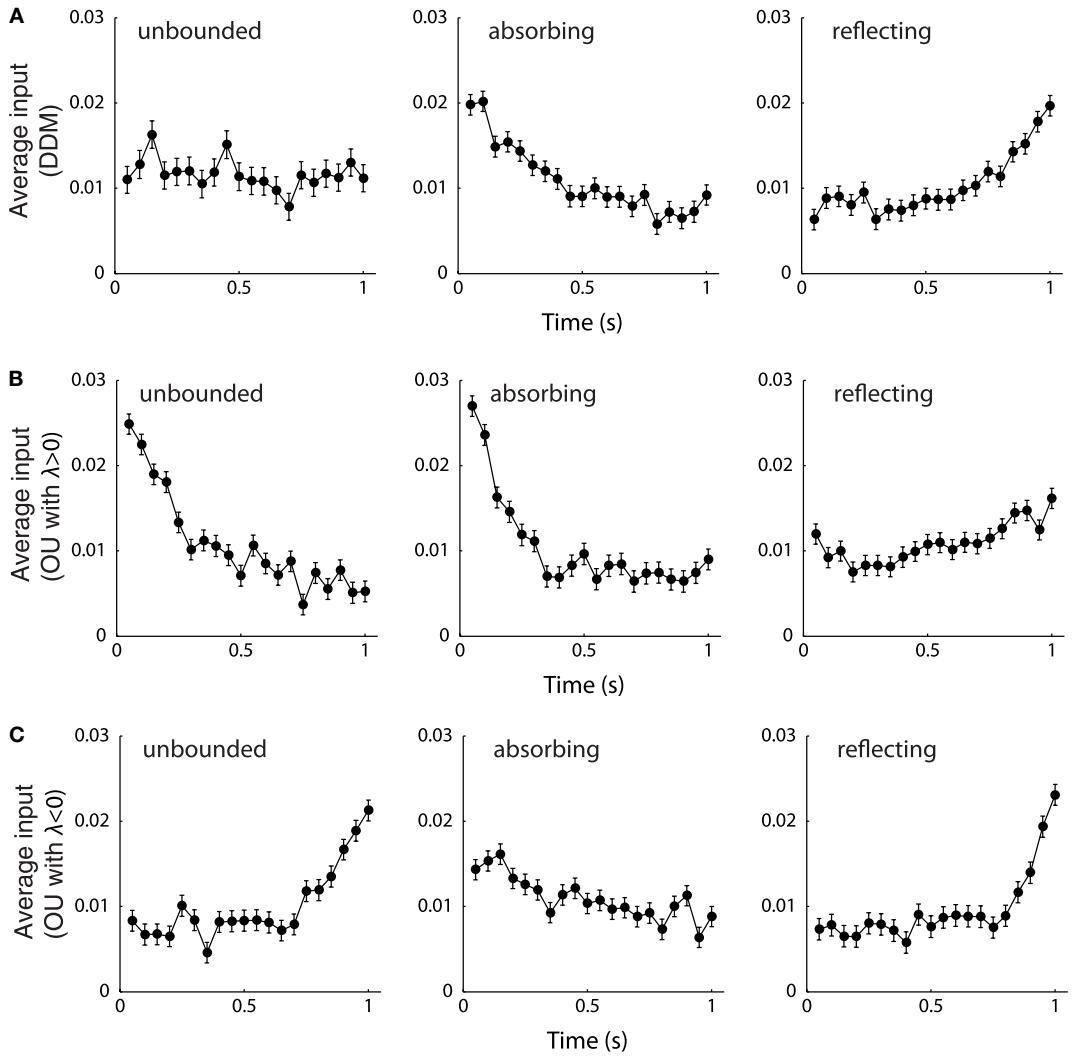


FIGURE 5 | The primacy and recency effects of the DDM and OU model. (A) The bounded and unbounded DDM. **(B)** The bounded and unbounded OU models with $\lambda > 0$. **(C)** The bounded and unbounded OU model with $\lambda < 0$. All the models were simulated with $\mu = 0.71\ s^{-1}$, $\sigma = 1\ s^{-1}$, $b = 0.47$, and $T_e = 1$ s. The growth-decay parameter of the OU models was set to $\lambda = 5.5$ **(B)** and $\lambda = -5.5$ **(C)**. In each panel, the model was simulated for 10,000 trials, and the sensory evidence from all correct trials was recorded and averaged. The data points show the means and standard errors of the sensory evidence at every time step. For $\mu > 0$, a larger averaged input indicates that the sensory evidence at that time point has, on average, a larger influence on the final choice, and a smaller averaged input indicates that the choice depends to a lesser extent on the evidence at that time. Figure modified from Zhang and Bogacz (2010a).

showed that early evidence is weighted less in a perceptual decision task (i.e., the integration is leaky), but the leakage quickly decreased with practice. In Usher and McClelland's (2001) study, primacy/recency effects were tested with fast visual streams of alternating letters lasting for only 256 ms. They randomly mixed shorter trials with non-stationary evidence and longer trials with constant evidence. Such a design encouraged subjects to estimate the entire sequence of the non-stationary evidence, because making decisions on only a fraction of early evidence would result in low performance on longer trials. Their results suggest a general recency effect with strong individual differences, although the source of the large between-subject variability has not yet been identified.

PERFORMANCE OF THE BOUNDED DECISION-MAKING MODELS

Several studies have reported significant improvements in model fit by introducing evidence boundaries. Ratcliff (2006) fitted data for the DDM and the LCA model from a categorization task in which subjects were required to decide whether the number of dots on the screen is large or small. The absorbing DDM and absorbing LCA model provide much better fits than the unbounded models, in particular for the TC paradigm with very short or long decision times. Another study showed that for a shape discrimination task (Usher and McClelland, 2001), the behavioral data is more likely to have been fitted by the bounded DDM than by the unbounded OU model (Zhang et al., 2009). Leite and Ratcliff (2010) showed that the LCA model with zero reflecting boundary produced better fits

to the RT distributions than the unbounded model in perceptual decision tasks with different number of alternatives. Zhang et al. (2009) observed that for a given set of model parameters, the ER of the absorbing and reflecting DDM are identical at any decision time. Therefore, although the two types of boundary influence the model dynamics, and weight the order of the momentary evidence in different ways, the two bounded DDMs can fit the experimental data from the TC paradigm equally well. A similar equality between absorbing and reflecting OU models has also been observed (Zhang and Bogacz, 2010a).

The successful applications of the bounded models promote us to consider how different types of evidence boundaries may affect the models' performance. For the IC paradigm, adding lower reflecting boundaries at zero generally decreases mean RT of the LCA model for a given ER, and this change is more significant for decision tasks with multiple alternatives (Bogacz et al., 2007; Leite and Ratcliff, 2010). Increasing the upper boundary in the absorbing LCA model, or the distance between the two boundaries in the absorbing DDM and absorbing OU model, leads to an increase in the mean and variance of RT distributions (Wagenmakers et al., 2005) and a decrease of ER (i.e., trading speed for accuracy, see Fast Boundary Modulation: Speed–Accuracy Tradeoff). For the TC paradigm, the bounded DDM has an asymptotic accuracy as T_c increases, which is consistent with experimental observations (Meyer et al., 1988; Usher and McClelland, 2001). Increasing boundary separation in the bounded DDM monotonically decrease the ER for a given decision time, until the boundary is sufficiently large that the integrator can barely reach the boundary before T_c, and under this condition the bounded DDM model is equivalent to the unbounded DDM (Zhang et al., 2009; Leite and Ratcliff, 2010). Interestingly, the relationship between the evidence boundary and the ER is not monotonic in the bounded OU model (Zhang and Bogacz, 2010a). For the OU model with a negative λ value, a finite absorbing boundary yields lower ER than the unbounded OU model. In contrast, a finite reflecting boundary lowers the ER for the OU model with a positive λ value (**Figure 6A**). Simulation results suggested that as T_c increases, the value of λ that yields the lowest ER decreases for the absorbing OU

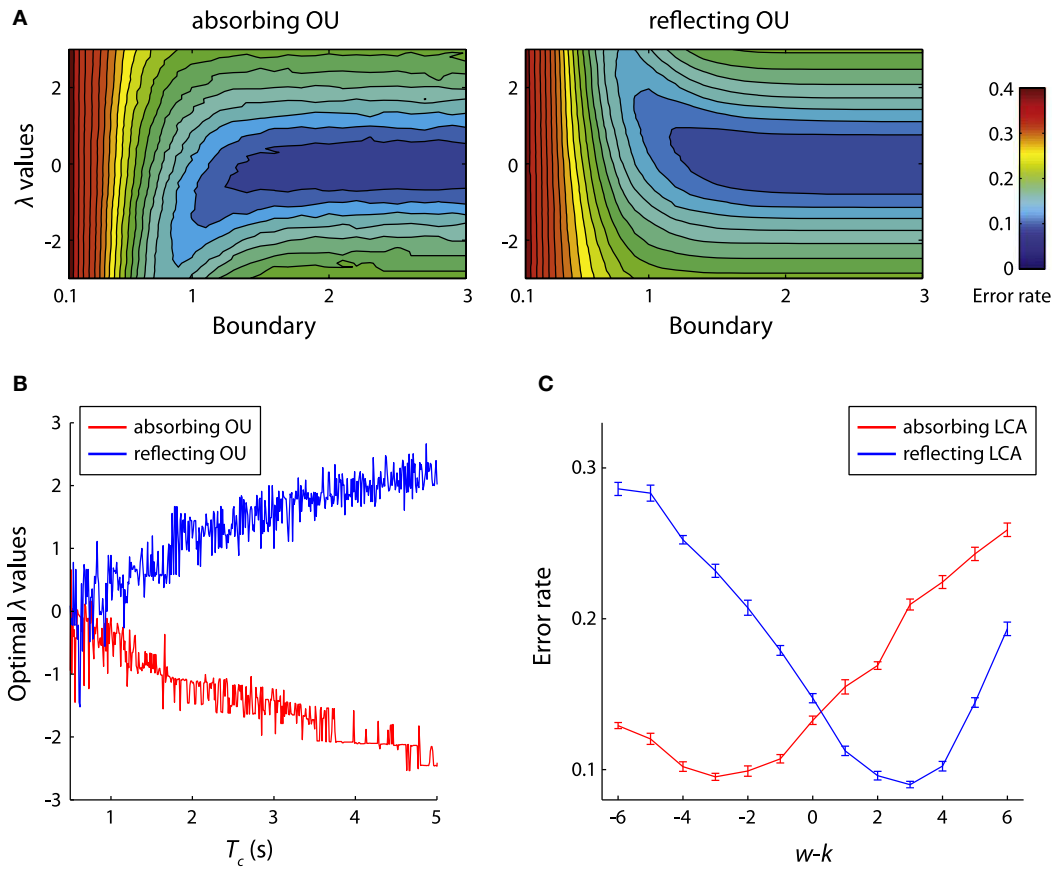


FIGURE 6 | Performance of the bounded models. (A) The error rates of the absorbing (left) and reflecting (right) OU models in the TC paradigm. The bounded OU models are simulated with the following parameters: λ in (−3, 3) with step 0.1, b in (0.1, 3) with step 0.1, $\mu = \sigma = 1\ s^{-1}$, and $T_c = 1$ s. The contour plots illustrate the mean error rates of the bounded OU models estimated from 10,000 simulations for each possible parameter combinations. Figure modified from Zhang and Bogacz (2010a). **(B)** The estimated optimal λ values of the absorbing and reflecting OU models that yield minimum error rate for different T_c varying from 0.5 to 5 s. Figure modified from Zhang and Bogacz (2010a). **(C)** The error rates of the bounded LCA model. The models were simulated with parameters: $\mu_1 = 5.41\ s^{-1}$, $\mu_2 = 4\ s^{-1}$, $\sigma = 1\ s^{-1}$, $b_+ = 1.5$, $b_- = 0$, and $T_c = 3$ s. The sum of decay and inhibition was fixed at $w + k = 6$, while their difference changed from −6 to 6.

model and increases for the reflecting OU model (**Figure 6B**). This relationship can be explained by the joint primacy/recency effects from the boundary and the λ value of the bounded OU model (see Primacy and Recency Effects). Recall that the optimal decision strategy, as suggested by the SPRT and NPT, would be to equally weight the momentary evidence received at different time points (i.e., no primacy or recency effects). The bounded OU model approximates to the optimal strategy when the primacy/recency effects introduced by the boundary and λ are balanced. That is, the absorbing OU model needs to be coupled with negative λ and the reflecting OU model needs to be coupled with positive λ. The relative strengths of the primacy/recency effects introduced by the boundary and λ values deserve further research.

The findings from one-dimensional bounded models provide clues to the understanding of performance of the bounded LCA model. Recall that the unbounded LCA model implements the optimal decision strategy when the decay and inhibition are balanced ($w = k$), i.e., when the LCA model is reduced to the DDM. Bogacz et al. (2007) showed that the balance of decay and inhibition does not optimize the performance of the bounded LCA model in the TC paradigm. Instead, by decreasing inhibition relative to decay ($w < k$) the absorbing LCA model can achieve lower ER. Conversely, the reflecting LCA model has lower ER when inhibition is larger than decay ($w > k$; **Figure 6C**). The symmetric relationship between the absorbing and reflecting LCA models is analogous to that of bounded OU models with positive and negative λ. Therefore it is possible that the bounded LCA model can be reduced to the bounded OU model for certain parameters (cf. van Ravenzwaaij et al., 2012). Bogacz et al. (2007) also suggest that by limiting the integrator stages to be non-negative, the absorbing LCA model can approximates the asymptotically optimal decision strategy (Draglia et al., 1999; Dragalin et al., 2000) for multiple alternative tasks (Bogacz and Gurney, 2007).

NEURAL IMPLEMENTATION OF DECISION BOUNDARY

How is the decision boundary realized in neural circuits? In the minimal recurrent loop model by Smith and McKenzie (2011), the decision boundary is implemented by an interaction between the recurrent loops and separate decision neurons. The decision neurons receive spiking inputs from the recurrent loops that represent the accumulated evidence. A decision is rendered as soon as the membrane potential of one decision neuron reaches a threshold. This mechanism predicts a causal link between the firing of decision neurons and overt actions. But an important question remains: where in the brain is the decision boundary implemented?

One possibility is that the decision boundary is implemented within neural integrators, namely the *local hypothesis*. Wong and Wang (2006) studied a simplified version of the biologically based model of Wang (2002) by using mean-field theory. Their analysis showed that if neural integrators are mediated by recurrent excitatory connections between spiking neurons, the dynamics of neural integrators may contain multiple stable attractor states, which act as implicit decision boundaries to terminate integration processes. This model successfully accounts for psychophysical data and LIP neural activity in RDM tasks (Wong and Wang, 2006; Wong et al., 2007). However, previous studies using the RDM task or other visual discrimination tasks have identified putative neural integrators in the FEF (Hanes and Schall, 1996; Schall and Thompson, 1999; Schall, 2002), the SC (Basso and Wurtz, 1998; Ratcliff et al., 2003a), and the DLPFC (Kim and Shadlen, 1999; Domenech and Dreher, 2010), which exhibit activity patterns similar to LIP neurons. A recent study showed that the inferior frontal sulcus is also likely to integrate evidence from multiple sensory modalities (Noppeney et al., 2010). Therefore, multiple neural integrators may coexist in different brain regions and may be simultaneously functioning during a decision process, though we do not know whether the neural integrators across different regions are independent or are more likely to interact with each other. If the local hypothesis is correct, it is yet not clear whether observed boundary crossing in one integrator region has a causal role in rendering a decision, or could merely reflect terminal integration in other integrator regions. Further experiments testing the activity of neural integrators in predefined regions under different decision tasks are necessary to confirm this hypothesis.

An alternative possibility, the *central hypothesis*, proposes that detection of boundary crossing is implemented by a central neural circuit outside integrator regions, rather than an intrinsic property of neural integrators. This hypothesis predicts that a central circuit is capable of detecting boundary crossing in integrators within different regions. One potential component of the central circuit is the basal ganglia (BG) because of its unique anatomy. First, the two BG input nuclei, the striatum and the subthalamic nucleus, receive direct inputs from multiple cortical regions including the LIP, FEF, and DLPFC (Smith et al., 1998; Hikosaka et al., 2000; Nakano et al., 2000). Second, most BG nuclei are organized in separate somatotopic areas representing different body parts, and each broad somatotopic area is further subdivided into functionally defined parallel channels, based upon specific movements of an individual body part (Alexander et al., 1986, 1990; Parent and Hazrati, 1995). Therefore the BG can access a number of information sources from the cortex and control complex motor responses, which make the BG important loci of action selection, reinforcement learning, and motor control (Karabelas and Moschovakis, 1985; Graybiel et al., 1994; Gurney et al., 2001a,b; Frank et al., 2004; Samejima et al., 2005). Lo and Wang (2006) proposed that detection of boundary crossing is implemented through a BG-SC pathway. By default the BG output nuclei send tonic inhibition (Hopkins and Niessen, 1976; Francois et al., 1984; Karabelas and Moschovakis, 1985) to downstream motor areas (e.g., the SC) to suppress any saccadic response. When the activity of a neural integrator (e.g., LIP neurons) is large enough, the striatum inhibits BG output nuclei and hence releases inhibition to the SC. The boundary crossing is then detected by burst neurons (Munoz and Wurtz, 1995) in the SC by an all-or-nothing burst signal. Bogacz and Gurney (2007) showed that the BG is necessary for the brain to implement asymptotically optimal decision strategy for NAFC tasks. Nevertheless, although Lo and Wang (2006) demonstrated that the central hypothesis can be implemented by the BG-SC circuit, the model relies on the unique burst property of the SC neurons to detect boundary crossing, which is primarily associated with eye movements. It is not clear whether the same mechanism can be applied to decision tasks requiring other response modalities (e.g., Ho et al., 2009), or tasks which require subjects

to withhold their responses before a response signal (i.e., the TC paradigm).

Taken together, although convincing data exists for the presence of neural integrators in the cortex, current findings are inconclusive regarding the neural implementation of decision boundaries. Part of the difficulty in investigating the boundary mechanism is that decision neurons may exhibit task-modulated ramping activity that is similar to neural integrators, if there exists positive feedback connections between the decision neurons and the integrators (Simen, 2012). As a result the two processes may be indistinguishable solely by the observation of ramping activity from neural recording data.

EFFECTS OF BOUNDARY CHANGES

The decision boundary is usually assumed to be under subjective control. On one hand, the decision boundary should be stable in regards to sensory evidence, enabling subjects to respond consistently when faced with similar environments or goals. The stability of the decision boundary is evident from the fact that in both IC and TC versions of the RDM tasks, LIP neurons attain the same level of activity before saccadic responses, independent of motion coherence (Shadlen and Newsome, 2001; Roitman and Shadlen, 2002). On the other hand, the decision boundary may also exhibit a certain degree of flexibility, allowing subjects to tailor their responses on demand, or accounting for changes in some internally driven factors. This section reviews psychological and physiological factors that could be modulated by changes in the decision boundary at different time scales.

FAST BOUNDARY MODULATION: SPEED–ACCURACY TRADEOFF

The change in decision boundary provides a straightforward account of the speed–accuracy tradeoff (SAT) effect that is often observed in decision-making tasks (Schouten and Bekker, 1967; Wickelgren, 1977; Luce, 1986; Franks et al., 2003; Chittka et al., 2009). For the DDM and the OU model (**Figure 7A**), decreasing the distance between two decision boundaries reduces the amount of accumulated evidence prior to a decision, leading to fast but error-prone responses. Conversely, increasing the distance between boundaries leads to slow but accurate decisions. For the LCA model or other models that have multiple integrators (e.g., the LBA model), the SAT can be manipulated by changing either the upper boundary (**Figure 7B**) or the lower baseline activity at the beginning of the trial (**Figure 7C**) (Bogacz et al., 2010b). Behavioral studies suggest that subjects can effectively trade speed for accuracy when instructed to respond as accurately as possible, or vice versa when instructed to respond as quickly as possible, and the behavioral differences between speed and accuracy instructions can be explained by a change of decision boundaries in the DDM (Palmer et al., 2005; Ratcliff, 2006; Ratcliff and McKoon, 2008). In a similar attempt to study SAT using the LBA model, Forstmann et al. (2008) observed that SAT in the RDM task can be best accounted for by a change in the decision boundary, not by changes of the drift rate or other model parameters. It has been suggested that humans can set the SAT to maximize the reward rate (producing the most correct decisions in a given period of time) by learning the optimal decision boundaries through feedback (Simen et al., 2006, 2009; Bogacz et al., 2010a; Starns and Ratcliff, 2010; Balci et al., 2011). Furthermore, impairments in the optimization of the SAT in neuropsychiatric patients with impulsive behaviors, such as attention-deficit hyperactivity disorder, has been associated with maladaptive regulation of the decision boundary in perceptual tasks (Mulder et al., 2010).

Can we consider the SAT as a signature for identifying neural correlates of decision boundaries? Several recent fMRI studies reveal brain regions associated with the SAT, including the SMA, the pre-SMA, the anterior cingulate cortex, the striatum, and the DLPFC (Forstmann et al., 2008; Ivanoff et al., 2008; van Veen et al., 2008; Blumen et al., 2011; van Maanen et al., 2011; for review, see Bogacz et al., 2010b; **Figure 8A**). Using a model-based fMRI analysis, Forstmann et al. (2008) showed that the extent of response facilitation for the speed condition in the RDM task, as quantified by a decrease of the decision boundary in the LBA model, correlated with BOLD response increase in the pre-SMA and striatum between the speed and the accuracy conditions (**Figure 8B**). Further studies suggest that the strength of structural connectivity between the two regions predicts the amount of boundary change in individual subjects (Forstmann et al., 2010a, 2011; **Figure 8C**). These results support the central hypothesis that the BG circuit is involved in controlling the decision boundary (Lo and Wang, 2006; Bogacz et al., 2010b).

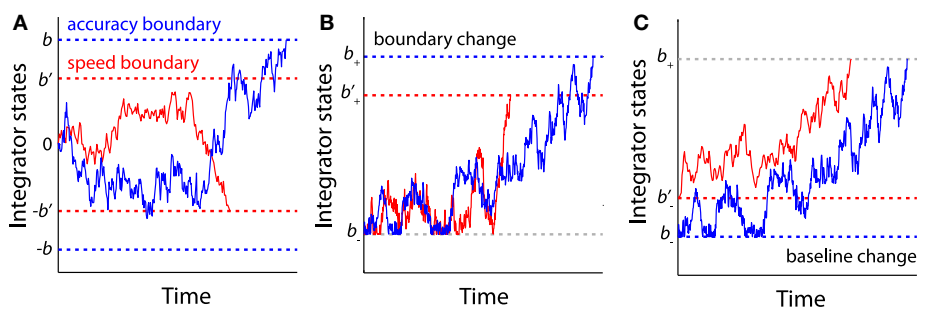


FIGURE 7 | The sequential sampling models account for SAT. **(A)** For the models with a single integrator (e.g., the DDM and the OU model), increasing the distance between two boundaries (blue boundaries $\pm b$) leads to slow but accurate decisions, while decreasing the boundary distance (red boundaries $\pm b'$) leads to fast but risky decisions. **(B)** For the models with multiple integrator (e.g., the LCA model), the SAT can be accounted for by changes in the upper boundary (b_+ and b'_+). **(C)** The SAT can also be accounted for by changes in the lower baseline activity (b_- and b'_-).

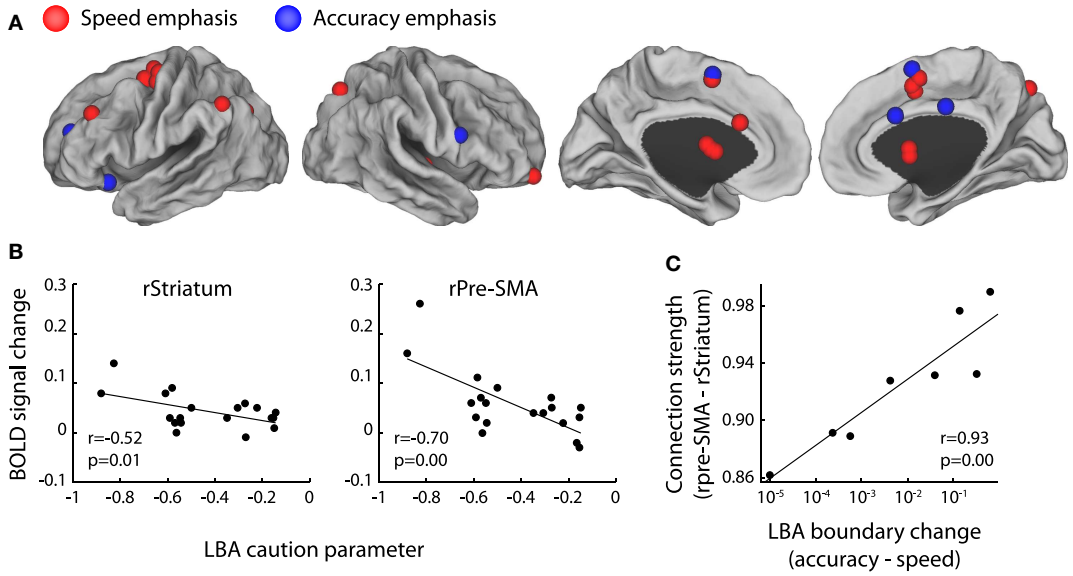


FIGURE 8 | The neural correlates of SAT. (A) Brain regions associated with the SAT are projected onto a cortical surface using Caret software (Van Essen et al., 2001). The foci represent the coordinates of the peak voxels reported by four fMRI studies (Forstmann et al., 2008; Ivanoff et al., 2008; van Veen et al., 2008; van Maanen et al., 2011). All the studies manipulated the SAT of perceptual decision tasks by speed emphasis or accuracy emphasis. The red foci illustrate increased BOLD response with speed emphasis and the blue foci illustrate increased BOLD response with accuracy emphasis. **(B)** In the RDM task, the BOLD response increases in the right Pre-SMA and the right Striatum in the speed versus the accuracy condition. These BOLD response changes are associated with decreases in the response caution parameter, which is quantified by boundary changes in the LBA model. Figure modified from Forstmann et al. (2008). **(C)** The strength of structural connections between the Pre-SMA and the Striatum in individual subjects correlate with the changes of the LBA decision boundary between the speed and the accuracy condition. Figure modified from Forstmann et al. (2010a).

Nevertheless, some concerns remain regarding the causal role of decision boundary in SAT. First, an emphasis on speed may be associated with other cognitive processes (Rinkenauer et al., 2004). For example, some studies have proposed that the integration process is coupled with an urgency signal that increases as a function of time (Churchland et al., 2008; Cisek et al., 2009). The urgency signal effectively lowers the decision boundary as time elapses (Ditterich, 2006), and the SAT can be attributed to a change in strength of the urgency signal. Second, some models predict that SAT is in fact controlled by the distance between the boundary and baseline (**Figure 7C**). Hence emphasizing speed or accuracy may modulate the decision boundary, baseline, or a combination of the two (Bogacz et al., 2010b; Simen, 2012). In particular, decreasing decision boundary is equivalent to increasing baseline activations in the LBA model. Recent fMRI studies suggest that the SAT is more likely to modulate baseline activity in the medial frontal cortex (pre-SMA and SMA), as these regions exhibit a greater BOLD response in the speed instruction compared to the accuracy instruction. Other studies suggest that SAT may modulate a decision boundary in the lateral PFC, where the speed instruction is associated with decreased BOLD responses (Ivanoff et al., 2008; Wenzlaff et al., 2011). However, it is possible that the aforementioned cortical areas do not directly change the decision boundary or baseline, but provide a control signal that modulates striatal activity (Bogacz et al., 2010b). In a recent neurophysiological study (Heitz and Schall, 2011), monkeys were trained to trade accuracy for speed in a visual search task. Fitting the behavioral data with the LBA model showed that the speed instruction can be accounted for by a decrease in the decision boundary. Interestingly, speed instruction led to an increased baseline activity as well as an increased presaccadic activity in the FEF, suggesting that the neural implementation of SAT likely involves multiple processes, rather than a single boundary or baseline change predicted by psychological models.

SLOW BOUNDARY MODULATION: PERCEPTUAL LEARNING AND AGING

It is well-known that practice can improve performance in many perceptual tasks, resulting in higher accuracy and shorter RTs (Logan, 1992; Heathcote et al., 2000). Traditional approaches usually quantify learning effects as changes in the mean accuracy or RT. Several recent studies have attempted to decompose component processes mediating perceptual learning by using sequential sampling models. Petrov et al. (2011) fitted the DDM to behavioral data from a fine motion-discrimination task and showed that learning effects across multiple training sessions are mainly associated with an increase in drift rate and a decrease in non-decision time (see also Dutilh et al., 2009). This result is consistent with previous findings that learning facilitates neural representation of task-relevant features by tuning neural selectivity in the sensory areas (Gilbert et al., 2001; Yang and Maunsell, 2004; Kourtzi and DiCarlo, 2006; Raiguel et al., 2006; Kourtzi, 2010; Zhang et al., 2010). Other studies suggest that extensive training also leads to a significant reduction in the boundary distance in the DDM (Ratcliff et al., 2006; Dutilh et al., 2009; Liu and Watanabe, 2011). Using the RDM task, Liu and Watanabe (2011) investigated the learning effect across different days and showed

that training without feedback decreases the decision boundary in the DDM and also increases drift rate. Dutilh et al. (2009) proposed that a dual process (changes in both boundary and drift rate) is necessary to account for the noticeable decrease in RT even after the improvement in accuracy saturates during training. The involvement of boundary reduction in perceptual learning is supported by experimental findings that perceptual learning may not only change sensory representation, but also enhance the decision process in intraparietal regions (Law and Gold, 2008; Zhang and Kourtzi, 2010). Further research combining a modeling approach with multiple imaging sessions over the course of training may reveal how learning and feedback modulate sensory representation and decision processes during perceptual decisions.

While training may improve the ability of subjects to make faster decisions in perceptual decision tasks and result in a lower decision boundary, one primary finding in aging is that RTs in cognitive tasks increase as people age, and this generalized slowing is sometimes coupled with impairments in accuracy (Cerella, 1985, 1991; Fisk and Warr, 1996; Salthouse, 1996). Recent studies have employed the DDM with behavioral data to identify the effects of aging in a number of choice tasks (Ratcliff et al., 2001, 2003b, 2004b, 2007; Thapar et al., 2003; Spaniol et al., 2006). A consistent observation is that slowing in older adults can be explained by two factors: an increase in the decision boundary and a prolongation of non-decision time. The decision boundary increase in aging suggests that older subjects are more cautious in making decisions compared with younger subjects (Ratcliff et al., 2006; Starns and Ratcliff, 2010). This age-dependent change in the decision boundary may be due to structural limits in pre-SMA and striatal connectivity (Forstmann et al., 2011) or functional impairments in the striatum (Kühn et al., 2011) in the aging brain. These findings are consistent with the central hypothesis that the striatum is involved in modulating decision boundaries.

DISCUSSION

This article has reviewed recent developments that shed light on the effects and mechanisms of evidence boundaries. Theoretically, boundaries shape the dynamics of decision processes in two aspects. First, the evidence boundary provides an ecological function to constrain the evidence needed for rendering a decision, since the nervous system cannot process an unlimited amount of information. Second, the evidence boundary provides a mechanistic function to determine the termination of a decision process. The necessity of the evidence boundary is not limited to a specific model, but is a common feature shared by different sequential sampling models and other accumulator models (e.g., the LBA model), independent of the model structures. Empirically, the presence of evidence boundary is evident from behavioral, neurophysiological and neuroimaging data. Existing findings suggest that evidence boundaries remains stable to changes in the external environment (e.g., sensory information), but may vary systematically with some internal factors (e.g., speed or accuracy emphasis, practice, or aging). Whether acting on its own, or interacting with other decision-related processes, boundaries play a crucial role in the formation of decisions. Therefore boundary mechanisms provide a window into understanding the cognitive processes associated with choice behavior.

Despite the increasing number of recent studies examining the evidence boundary, we are still far from a complete picture of its functions and neural implementations. Here I suggest several directions that merit further investigation. First, among decision models that implement the integration-to-boundary mechanism, it is not clear to what extent the effect of a boundary depend on the specific structure of the models. For example, if for a given dataset the DDM predicts a change in the boundary between two experimental conditions, or a correlation between the estimated boundary and cognitive assessment scores (e.g., Ratcliff et al., 2008), would we reach the same conclusion if using the LCA model or the LBA model? van Ravenzwaaij and Oberauer (2009) suggested that boundaries estimated from different sequential sampling models are generally consistent, but do not necessarily correspond with those estimated from the LBA model (cf. Donkin et al., 2011). Such discrepancies between models need be considered if researchers plan to estimate boundary changes from experimental data, or use estimated model parameters to guide subsequent neuroimaging analysis.

Psychological models conceptualize the evidence boundary as a unitary representation. The neural implementation of evidence boundaries is likely to be more sophisticated and remains to be determined (see Simen et al., 2011; Smith and McKenzie, 2011 for recent attempts to bridge the gap between the two). The existing findings favor the central hypothesis over the local hypothesis, but we do not yet fully understand the causal relationship between the activity of the BG nuclei and the changes of the boundary. Studies discussed in this article suggests that boundary changes can occur at different time scales, ranging from a few seconds during which the SAT can be effectively adapted, to a few days during which it is necessary to modulate the boundary through extensive training and feedback. Hence if a central neural circuit exists for the detection of boundary crossing, this system is likely to be affected by different underlying control signals, but we do not know how and where in the brain the control signals for boundary changes are encoded. A related question is how the evidence boundary may be affected by aging or neurodegenerative diseases. Could these long-term factors alter control signals that modulate the boundary, or directly act upon the neural circuits that implement the boundary? Answering these questions will require researchers to combine established modeling approaches with comprehensive neuroimaging protocols.

Finally, existing findings suggest that the integration-to-boundary process governs a broad range of cognitive tasks (Gold and Shadlen, 2007). An important direction for future research is to investigate the effects of boundaries in choice tasks other than perceptual decisions. One example is interval timing estimation, in which subjects produce or estimate a specific duration (Church and Deluty, 1977; Roberts, 1981; Rakitin et al., 1998; Macar et al., 1999; Allan and Gerhardt, 2001). A variant of the DDM has recently been proposed for interval timing (Simen et al., 2011). The model assumes a single integrator with variable drift rate representing elapsed time at different durations and a constant decision boundary. A fixed boundary predicted by the model is supported by experimental findings that slow cortical potentials measured in the pre-SMA/SMA, which have been interpreted as a signature of time accumulation process, show no amplitude

difference between different interval times (Elbert et al., 1991; Pfeuty et al., 2005; Kononowicz and van Rijn, 2011; Ng et al., 2011). Another example is voluntary action decision, which require subjects to make selections between actions that have no differential sensory attributes or action outcomes (Brass and Haggard, 2008; Haggard, 2008; Soon et al., 2008; Andersen and Cui, 2009; Roskies, 2010). Recent studies propose that during the formation of voluntary decisions the intention of selecting each action gradually builds up in independent integrators until the winning integrators reaches the boundary and renders the decision (Zhang et al., 2012). This hypothesis is supported by observations of a progressive rise in the readiness potential and neural activity in the medial prefrontal cortex before consciously aware of voluntary actions (Libet, 1985; Sirigu et al., 2004; Fried et al., 2011). These findings from different types of cognitive tasks suggest that the brain may encode the evidence boundary as a common currency for perceptual information, subjective intention, or individual preference (e.g., Chib et al., 2009; Krajbich et al., 2010) to guide behavioral responses, depending on the context of the task. An intriguing possibility is that evidence boundaries associated with different cognitive tasks may be mediated by the same neural implementation. This generic implementation provides a potential bridge between behavioral and neural data to regulate the formation and initiation of complex behavior.

ACKNOWLEDGMENTS

This work was supported by Medical Research Council intramural program MC_A060_5PQ30. The author thanks Laura Hughes, Anna McCarrey, Charlotte Rae, and Timothy Rittman for reading the previous version of the manuscript and useful comments.

REFERENCES

Alexander, G. E., Crutcher, M. D., and DeLong, M. R. (1990). Basal ganglia-thalamocortical circuits: parallel substrates for motor, oculomotor, "prefrontal" and "limbic" functions. *Prog. Brain Res.* 85, 119–146.

Alexander, G. E., DeLong, M. R., and Strick, P. L. (1986). Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annu. Rev. Neurosci.* 9, 357–381.

Allan, L. G., and Gerhardt, K. (2001). Temporal bisection with trial referents. *Percept. Psychophys.* 63, 524–540.

Andersen, R. A., and Cui, H. (2009). Intention, action planning, and decision making in parietal-frontal circuits. *Neuron* 63, 568–583.

Balci, F., Simen, P., Niyogi, R., Saxe, A., Hughes, J. A., Holmes, P., and Cohen, J. D. (2011). Acquisition of decision making criteria: reward rate ultimately beats accuracy. *Atten. Percept. Psychophys.* 73, 640–657.

Barnard, G. A. (2007). Sequential tests in industrial statistics. *J. R. Stat. Soc.* 8, 1–26.

Basso, M. A., and Wurtz, R. H. (1998). Modulation of neuronal activity in superior colliculus by changes in target probability. *J. Neurosci.* 18, 7519–7534.

Blumen, H. M., Gazes, Y., Habeck, C., Kumar, A., Steffener, J., Rakitin, B. C., and Stern, Y. (2011). Neural networks associated with the speed-accuracy tradeoff: evidence from the response signal method. *Behav. Brain Res.* 224, 397–402.

Bogacz, R. (2007). Optimal decision-making theories: linking neurobiology with behaviour. *Trends Cogn. Sci. (Regul. Ed.)* 11, 118–125.

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., and Cohen, J. D. (2006). The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. *Psychol. Rev.* 113, 700–765.

Bogacz, R., and Gurney, K. (2007). The basal ganglia and cortex implement optimal decision making between alternative actions. *Neural Comput.* 19, 442–477.

Bogacz, R., Hu, P. T., Holmes, P. J., and Cohen, J. D. (2010a). Do humans produce the speed-accuracy trade-off that maximizes reward rate? *Q. J. Exp. Psychol. (Hove)* 63, 863–891.

Bogacz, R., Wagenmakers, E.-J., Forstmann, B. U., and Nieuwenhuis, S. (2010b). The neural basis of the speed-accuracy tradeoff. *Trends Neurosci.* 33, 10–16.

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.

Born, R. T., and Bradley, D. C. (2005). Structure and function of visual area MT. *Annu. Rev. Neurosci.* 28, 157–189.

Brass, M., and Haggard, P. (2008). The what, when, whether model of intentional action. *Neuroscientist* 14, 319–325.

Britten, K., Shadlen, M., Newsome, W., and Movshon, J. (1992). The analysis of visual motion: a comparison of neuronal and psychophysical performance. *J. Neurosci.* 12, 4745–4765.

Britten, K. H., Newsome, W. T., Shadlen, M. N., Celebrini, S., and Movshon, J. A. (1996). A relationship between behavioral choice and the visual responses of neurons in macaque MT. *Vis. Neurosci.* 13, 87–100.

Britten, K. H., Shadlen, M. N., Newsome, W. T., and Movshon, J. A. (1993). Responses of neurons in macaque MT to stochastic motion signals. *Vis. Neurosci.* 10, 1157–1169.

Brown, E., Gao, J., Holmes, P., and Bogacz, R. (2005). Simple neural networks that optimize decisions. *Int. J. Bifurcat. Chaos* 15, 803–826.

Brown, E., and Holmes, P. (2001). Modelling a simple choice task: stochastic dynamics of mutually inhibitory neural groups. *Stochast. Dynam.* 1, 159–191.

Brown, S., and Heathcote, A. (2005a). A ballistic model of choice response time. *Psychol. Rev.* 112, 117–128.

Brown, S., and Heathcote, A. (2005b). Practice increases the efficiency of evidence accumulation in perceptual choice. *J. Exp. Psychol. Hum. Percept. Perform.* 31, 289–298.

Brown, S. D., and Heathcote, A. (2008). The simplest complete model of choice response time: linear ballistic accumulation. *Cogn. Psychol.* 57, 153–178.

Busemeyer, J. (2002). Survey of decision field theory. *Math. Soc. Sci.* 43, 345–370.

Busemeyer, J. R., Jessup, R. K., Johnson, J. G., and Townsend, J. T. (2006). Building bridges between neural models and complex decision making behaviour. *Neural Netw.* 19, 1047–1058.

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.

Cerella, J. (1985). Information processing rates in the elderly. *Psychol. Bull.* 98, 67–83.

Cerella, J. (1991). Age effects may be global, not local: comment on Fisk and Rogers (1991). *J. Exp. Psychol. Gen.* 120, 215–223.

Chib, V. S., Rangel, A., Shimojo, S., and O'Doherty, J. P. (2009). Evidence for a common representation of decision values for dissimilar goods in human ventromedial prefrontal cortex. *J. Neurosci.* 29, 12315–12320.

Chittka, L., Skorupski, P., and Raine, N. E. (2009). Speed-accuracy tradeoffs in animal decision making. *Trends Ecol. Evol. (Amst.)* 24, 400–407.

Church, R. M., and Deluty, M. Z. (1977). Bisection of temporal intervals. *J. Exp. Psychol. Anim. Behav. Process.* 3, 216–228.

Churchland, A. K., Kiani, R., and Shadlen, M. N. (2008). Decision-making with multiple alternatives. *Nat. Neurosci.* 11, 693–702.

Cisek, P., Puskas, G. A., and El-Murr, S. (2009). Decisions in changing conditions: the urgency-gating model. *J. Neurosci.* 29, 11560–11571.

Diederich, A. (1995). Intersensory facilitation of reaction time: evaluation of counter and diffusion coactivation models. *J. Math. Psychol.* 39, 197–215.

Diederich, A. (1997). Dynamic stochastic models for decision making under time constraints. *J. Math. Psychol.* 41, 260–274.

Ditterich, J. (2006). Stochastic models of decisions about motion direction: behavior and physiology. *Neural Netw.* 19, 981–1012.

Ditterich, J. (2010). A comparison between mechanisms of multi-alternative perceptual decision making: ability to explain human behavior, predictions for neurophysiology, and relationship with decision theory. *Front. Neurosci.* 4:184. doi:10.3389/fnins.2010.00184

Ditterich, J., Mazurek, M. E., and Shadlen, M. N. (2003). Microstimulation of visual cortex affects the speed of perceptual decisions. *Nat. Neurosci.* 6, 891–898.

Domenech, P., and Dreher, J.-C. (2010). Decision threshold modulation in the human brain. *J. Neurosci.* 30, 14305–14317.

Donkin, C., Brown, S., Heathcote, A., and Wagenmakers, E.-J. (2011). Diffusion versus linear ballistic accumulation: different models but the same conclusions about psychological processes? *Psychon. Bull. Rev.* 18, 61–69.

Donkin, C., and Heathcote, A. (2009). "Non-decision time effects in the lexical decision task," in *Proceedings of the 31st Annual Conference of the Cognitive Science Society,* eds N. A. Taatgen and H. van Rijn (Austin: Cognitive Science Society), 2902–2907.

Dosher, B. A. (1976). The retrieval of sentences from memory: a speed-accuracy study. *Cogn. Psychol.* 8, 291–310.

Dosher, B. A. (1984). Discriminating preexperimental (semantic) from learned (episodic) associations: a speed-accuracy study. *Cogn. Psychol.* 16, 519–555.

Dragalin, V. P., Tartakovsky, A. G., and Veeravalli, V. V. (2000). Multihypothesis sequential probability ratio tests. II. Accurate asymptotic expansions for the expected sample size. *IEEE Trans. Inf. Theory* 46, 1366–1383.

Draglia, V. P., Tartakovsky, A. G., and Veeravalli, V. V. (1999). Multihypothesis sequential probability ratio tests. I. Asymptotic optimality. *IEEE Trans. Inf. Theory* 45, 2448–2461.

Dutilh, G., Vandekerckhove, J., Tuerlinckx, F., and Wagenmakers, E.-J. (2009). A diffusion model decomposition of the practice effect. *Psychon. Bull. Rev.* 16, 1026–1036.

Edwards, W. (1965). Optimal strategies for seeking information: models for statistics, choice reaction times, and human information processing. *J. Math. Psychol.* 2, 312–329.

Elbert, T., Ulrich, R., Rockstroh, B., and Lutzenberger, W. (1991). The processing of temporal intervals reflected by CNV-like brain potentials. *Psychophysiology* 28, 648–655.

Estes, W. K. (1955). Statistical theory of spontaneous recovery and regression. *Psychol. Rev.* 62, 145–154.

Farrell, S., Ludwig, C. J. H., Ellis, L. A., and Gilchrist, I. D. (2010). Influence of environmental statistics on inhibition of saccadic return. *Proc. Natl. Acad. Sci. U.S.A.* 107, 929–934.

Fisk, J. E., and Warr, P. (1996). Age and working memory: the role of perceptual speed, the central executive, and the phonological loop. *Psychol. Aging* 11, 316–323.

Forstmann, B. U., Anwander, A., Schäfer, A., Neumann, J., Brown, S., Wagenmakers, E.-J., Bogacz, R., and Turner, R. (2010a). Corticostriatal connections predict control over speed and accuracy in perceptual decision making. *Proc. Natl. Acad. Sci. U.S.A.* 107, 15916–15920.

Forstmann, B. U., Brown, S., Dutilh, G., Neumann, J., and Wagenmakers, E.-J. (2010b). The neural substrate of prior information in perceptual decision making: a model-based analysis. *Front. Hum. Neurosci.* 4:40. doi:10.3389/fnhum.2010.00040

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., Tittgemeyer, M., Wagenmakers, E.-J., Derrfuss, J., Imperati, D., and Brown, S. (2011). The speed-accuracy tradeoff in the elderly brain: a structural model-based approach. *J. Neurosci.* 31, 17242–17249.

Francois, C., Percheron, G., and Yelnik, J. (1984). Localization of nigrostriatal, nigrothalamic and nigrotectal neurons in ventricular coordinates in macaques. *Neuroscience* 13, 61–76.

Frank, M. J., Seeberger, L. C., and O'Reilly, R. C. (2004). By carrot or by stick: cognitive reinforcement learning in parkinsonism. *Science* 306, 1940–1943.

Franks, N. R., Dornhaus, A., Fitzsimmons, J. P., and Stevens, M. (2003). Speed versus accuracy in collective decision making. *Proc. Biol. Sci.* 270, 2457–2463.

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.

Furman, M., and Wang, X.-J. (2008). Similarity effect and optimal control of multiple-choice decision making. *Neuron* 60, 1153–1168.

Gilbert, C. D., Sigman, M., and Crist, R. E. (2001). The neural basis of perceptual learning. *Neuron* 31, 681–697.

Gold, J. I., and Shadlen, M. N. (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.

Gomez, P., Ratcliff, R., and Perea, M. (2007). A model of the go/no-go task. *J. Exp. Psychol. Gen.* 136, 389–413.

Grasman, R. P. P. P., Wagenmakers, E.-J., and van der Maas, H. L. J. (2009). On the mean and variance of response times under the diffusion model with an application to parameter estimation. *J. Math. Psychol.* 53, 55–68.

Graybiel, A., Aosaki, T., Flaherty, A., and Kimura, M. (1994). The basal ganglia and adaptive motor control. *Science* 265, 1826–1831.

Gurney, K., Prescott, T. J., and Redgrave, P. (2001a). A computational model of action selection in the basal ganglia. I. A new functional anatomy. *Biol. Cybern.* 84, 401–410.

Gurney, K., Prescott, T. J., and Redgrave, P. (2001b). A computational model of action selection in the basal ganglia. II. Analysis and simulation of behaviour. *Biol. Cybern.* 84, 411–423.

Haggard, P. (2008). Human volition: towards a neuroscience of will. *Nat. Rev. Neurosci.* 9, 934–946.

Hanes, D. P., and Schall, J. D. (1996). Neural control of voluntary movement initiation. *Science* 274, 427–430.

Hanks, T. D., Ditterich, J., and Shadlen, M. N. (2006). Microstimulation of macaque area LIP affects decision-making in a motion discrimination task. *Nat. Neurosci.* 9, 682–689.

Heath, R. (1992). A general nonstationary diffusion model for two-choice decision-making. *Math. Soc. Sci.* 23, 283–309.

Heathcote, A., Brown, S., and Mewhort, D. J. K. (2000). The power law repealed: the case for an exponential law of practice. *Psychon. Bull. Rev.* 7, 185–207.

Heekeren, H. R., Marrett, S., and Ungerleider, L. G. (2008). The neural systems that mediate human perceptual decision making. *Nat. Rev. Neurosci.* 9, 467–479.

Heitz, R. P., and Schall, J. D. (2011). "Neural basis of speed-accuracy tra de-off in frontal eye field," in *Abstracts of the Society for Neuroscience Annual Meeting 2011* (Washington, DC: Society for Neuroscience).

Hikosaka, O., Takikawa, Y., and Kawagoe, R. (2000). Role of the basal ganglia in the control of purposive saccadic eye movements. *Physiol. Rev.* 80, 953–978.

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.

Hopkins, D. A., and Niessen, L. W. (1976). Substantia nigra projections to the reticular formation, superior colliculus and central gray in the rat, cat and monkey. *Neurosci. Lett.* 2, 253–259.

Huk, A. C., and Shadlen, M. N. (2005). Neural activity in macaque parietal cortex reflects temporal integration of visual motion signals during perceptual decision making. *J. Neurosci.* 25, 10420–10436.

Ivanoff, J., Branning, P., and Marois, R. (2008). fMRI evidence for a dual process account of the speed-accuracy tradeoff in decision-making. *PLoS ONE* 3, e2635. doi:10.1371/journal.pone.0002635

Karabelas, A. B., and Moschovakis, A. K. (1985). Nigral inhibitory termination on efferent neurons of the superior colliculus: an intracellular horseradish peroxidase study in the cat. *J. Comp. Neurol.* 239, 309–329.

Kayser, A. S., Buchsbaum, B. R., Erickson, D. T., and D'Esposito, M. (2010a). The functional anatomy of a perceptual decision in the human brain. *J. Neurophysiol.* 103, 1179–1194.

Kayser, A. S., Erickson, D. T., Buchsbaum, B. R., and D'Esposito, M. (2010b). Neural representations of relevant and irrelevant features in perceptual decision making. *J. Neurosci.* 30, 15778–15789.

Kiani, R., Hanks, T. D., and Shadlen, M. N. (2008). Bounded integration in parietal cortex underlies decisions even when viewing duration is dictated by the environment. *J. Neurosci.* 28, 3017–3029.

Kim, J. N., and Shadlen, M. N. (1999). Neural correlates of a decision in the dorsolateral prefrontal cortex of the macaque. *Nat. Neurosci.* 2, 176–185.

Kononowicz, T. W., and van Rijn, H. (2011). Slow potentials in time estimation: the role of temporal accumulation and habituation. *Front. Integr. Neurosci.* 5:10. doi:10.3389/fnint.2011.00048

Kourtzi, Z. (2010). Visual learning for perceptual and categorical decisions in the human brain. *Vision Res.* 50, 433–440.

Kourtzi, Z., and DiCarlo, J. J. (2006). Learning and neural plasticity in visual object recognition. *Curr. Opin. Neurobiol.* 16, 152–158.

Krajbich, I., Armel, C., and Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nat. Neurosci.* 13, 1292–1298.

Kühn, S., Schmiedek, F., Schott, B., Ratcliff, R., Heinze, H.-J., Düzel, E., Lindenberger, U., and Lövden, M. (2011). Brain areas consistently linked to individual differences

in perceptual decision-making in younger as well as older adults before and after training. *J. Cogn. Neurosci.* 23, 2147–2158.

Laming, D. R. J. (1968). *Information Theory of Choice-Reaction Times.* Oxford: Academic Press.

Law, C.-T., and Gold, J. I. (2008). Neural correlates of perceptual learning in a sensory-motor, but not a sensory, cortical area. *Nat. Neurosci.* 11, 505–513.

Lehmann, E. (1959). *Testing Statistical Hypotheses.* New York: Wiley.

Leite, F. P., and Ratcliff, R. (2010). Modeling reaction time and accuracy of multiple-alternative decisions. *Atten. Percept. Psychophys.* 72, 246–273.

Libet, B. (1985). Unconscious cerebral initiative and the role of conscious will in voluntary action. *Behav. Brain Sci.* 8, 529–539.

Link, S. W. (1975). The relative judgment theory of two choice response time. *J. Math. Psychol.* 12, 114–135.

Link, S. W., and Heath, R. A. (1975). A sequential theory of psychological discrimination. *Psychometrika* 40, 77–105.

Liu, C. C., and Watanabe, T. (2011). Accounting for speed-accuracy tradeoff in perceptual learning. *Vision Res.* 61, 107–114.

Lo, C.-C., and Wang, X.-J. (2006). Cortico-basal ganglia circuit mechanism for a decision threshold in reaction time tasks. *Nat. Neurosci.* 9, 956–963.

Logan, G. D. (1992). Shapes of reaction-time distributions and shapes of learning curves: a test of the instance theory of automaticity. *J. Exp. Psychol. Learn Mem. Cogn.* 18, 883–914.

Luce, R. D. (1986). *Response Times: Their Role in Inferring Elementary Mental Organization.* New York: Oxford University Press.

Ludwig, C. J. H., Farrell, S., Ellis, L. A., and Gilchrist, I. D. (2009). The mechanism underlying inhibition of saccadic return. *Cogn. Psychol.* 59, 180–202.

Macar, F., Vidal, F., and Casini, L. (1999). The supplementary motor area in motor and sensory timing: evidence from slow brain potential changes. *Exp. Brain Res.* 125, 271–280.

Maunsell, J. H., and Van Essen, D. C. (1983). Functional properties of neurons in middle temporal visual area of the macaque monkey. I. Selectivity for stimulus direction, speed, and orientation. *J. Neurophysiol.* 49, 1127–1147.

Mazurek, M. E., Roitman, J. D., Ditterich, J., and Shadlen, M. N. (2003). A role for neural integrators in perceptual decision making. *Cereb. Cortex* 13, 1257–1269.

Mcmillen, T., and Holmes, P. (2006). The dynamics of choice among multiple alternatives. *J. Math. Psychol.* 50, 30–57.

Meyer, D. E., Irwin, D. E., Osman, A. M., and Kounios, J. (1988). The dynamics of cognition and action: mental processes inferred from speed-accuracy decomposition. *Psychol. Rev.* 95, 183–237.

Mulder, M. J., Bos, D., Weusten, J. M. H., van Belle, J., van Dijk, S. C., Simen, P., van Engeland, H., and Durston, S. (2010). Basic impairments in regulating the speed-accuracy tradeoff predict symptoms of attention-deficit/hyperactivity disorder. *Biol. Psychiatry* 68, 1114–1119.

Munoz, D. P., and Wurtz, R. H. (1995). Saccade-related activity in monkey superior colliculus. I. Characteristics of burst and buildup cells. *J. Neurophysiol.* 73, 2313–2333.

Nakano, K., Kayahara, T., Tsutsumi, T., and Ushiro, H. (2000). Neural circuits and functional organization of the striatum. *J. Neurol.* 247, V1–V15.

Newsome, W., and Pare, E. (1988). A selective impairment of motion perception following lesions of the middle temporal visual area (MT). *J. Neurosci.* 8, 2201–2211.

Newsome, W. T., Britten, K. H., and Movshon, J. A. (1989). Neuronal correlates of a perceptual decision. *Nature* 341, 52–54.

Neyman, J., and Pearson, E. S. (1933). On the problem of the most efficient tests of statistical hypotheses. *Philos. Trans. R. Soc. Lond. A* 231, 289–337.

Ng, K. K., Tobin, S., and Penney, T. B. (2011). Temporal accumulation and decision processes in the duration bisection task revealed by contingent negative variation. *Front. Integr. Neurosci.* 5:77. doi:10.3389/fnint.2011.00077

Niwa, M., and Ditterich, J. (2008). Perceptual decisions between multiple directions of visual motion. *J. Neurosci.* 28, 4435–4445.

Noppeney, U., Ostwald, D., and Werner, S. (2010). Perceptual decisions formed by accumulation of audiovisual evidence in prefrontal cortex. *J. Neurosci.* 30, 7434–7446.

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.

Papoulis, A. (1977). *Signal Analysis.* New York: McGraw-Hill.

Parent, A., and Hazrati, L.-N. (1995). Functional anatomy of the basal ganglia. I. The cortico-basal ganglia-thalamo-cortical loop. *Brain Res. Rev.* 20, 91–127.

Petrov, A. A., Van Horn, N. M., and Ratcliff, R. (2011). Dissociable perceptual-learning mechanisms revealed by diffusion-model analysis. *Psychon. Bull. Rev.* 18, 490–497.

Pfeuty, M., Ragot, R., and Pouthas, V. (2005). Relationship between CNV and timing of an upcoming event. *Neurosci. Lett.* 382, 106–111.

Philiastides, M. G., Ratcliff, R., and Sajda, P. (2006). Neural representation of task difficulty and decision making during perceptual categorization: a timing diagram. *J. Neurosci.* 26, 8965–8975.

Philiastides, M. G., and Sajda, P. (2007). EEG-informed fMRI reveals spatiotemporal characteristics of perceptual decision making. *J. Neurosci.* 27, 13082–13091.

Pietsch, A., and Vickers, D. (1997). Memory capacity and intelligence: novel techniques for evaluating rival models of a fundamental information-processing mechanism. *J. Gen. Psychol.* 124, 229–339.

Pike, A. R. (1966). Stochastic models of choice behaviour: response probabilities and latencies of finite Markov chain systems. *Br. J. Math. Stat. Psychol.* 19, 15–32.

Ploran, E. J., Nelson, S. M., Velanova, K., Donaldson, D. I., Petersen, S. E., and Wheeler, M. E. (2007). Evidence accumulation and the moment of recognition: dissociating perceptual recognition processes using fMRI. *J. Neurosci.* 27, 11912–11924.

Purcell, B. A., Heitz, R. P., Cohen, J. Y., Schall, J. D., Logan, G. D., and Palmeri, T. J. (2010). Neurally constrained modeling of perceptual decision making. *Psychol. Rev.* 117, 1113–1143.

Raiguel, S., Vogels, R., Mysore, S. G., and Orban, G. A. (2006). Learning to see the difference specifically alters the most informative V4 neurons. *J. Neurosci.* 26, 6589–6602.

Rakitin, B. C., Gibbon, J., Penney, T. B., Malapani, C., Hinton, S. C., and Meck, W. H. (1998). Scalar expectancy theory and peak-interval timing in humans. *J. Exp. Psychol. Anim. Behav. Process.* 24, 15–33.

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

Ratcliff, R. (1988). Continuous versus discrete information processing modeling accumulation of partial information. *Psychol. Rev.* 95, 238–255.

Ratcliff, R. (2002). A diffusion model account of response time and accuracy in a brightness discrimination task: fitting real data and failing to fit fake but plausible data. *Psychon. Bull. Rev.* 9, 278–291.

Ratcliff, R. (2006). Modeling response signal and response time data. *Cogn. Psychol.* 53, 195–237.

Ratcliff, R., Cherian, A., and Segraves, M. (2003a). A comparison of macaque behavior and superior colliculus neuronal activity to predictions from models of two-choice decisions. *J. Neurophysiol.* 90, 1392–1407.

Ratcliff, R., Thapar, A., and McKoon, G. (2003b). A diffusion model analysis of the effects of aging on brightness discrimination. *Percept. Psychophys.* 65, 523–535.

Ratcliff, R., Gomez, P., and McKoon, G. (2004a). A diffusion model account of the lexical decision task. *Psychol. Rev.* 111, 159–182.

Ratcliff, R., Thapar, A., and McKoon, G. (2004b). A diffusion model analysis of the effects of aging on recognition memory. *J. Mem. Lang.* 50, 408–424.

Ratcliff, R., and McKoon, G. (2008). The diffusion decision model: theory and data for two-choice decision tasks. *Neural Comput.* 20, 873–922.

Ratcliff, R., Philiastides, M. G., and Sajda, P. (2009). Quality of evidence for perceptual decision making is indexed by trial-to-trial variability of the EEG. *Proc. Natl. Acad. Sci. U.S.A.* 106, 6539–6544.

Ratcliff, R., and Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychol. Sci.* 9, 347–356.

Ratcliff, R., and Rouder, J. N. (2000). A diffusion model account of masking in two-choice letter identification. *J. Exp. Psychol. Hum. Percept. Perform.* 26, 127–140.

Ratcliff, R., Schmiedek, F., and McKoon, G. (2008). A diffusion model explanation of the worst performance rule for reaction time and IQ. *Intelligence* 36, 10–17.

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

Ratcliff, R., Thapar, A., and McKoon, G. (2001). The effects of aging on reaction time in a signal detection task. *Psychol. Aging* 16, 323–341.

Ratcliff, R., Thapar, A., and McKoon, G. (2006). Aging, practice, and perceptual tasks: a diffusion model analysis. *Psychol. Aging* 21, 353–371.

Ratcliff, R., Thapar, A., and McKoon, G. (2007). Application of the diffusion model to two-choice tasks for adults

75–90 years old. *Psychol. Aging* 22, 56–66.

Ratcliff, R., Van Zandt, T., and McKoon, G. (1999). Connectionist and diffusion models of reaction time. *Psychol. Rev.* 106, 261–300.

Rinkenauer, G., Osman, A., Ulrich, R., Muller-Gethmann, H., and Mattes, S. (2004). On the locus of speed-accuracy trade-off in reaction time: inferences from the lateralized readiness potential. *J. Exp. Psychol. Gen.* 133, 261–282.

Roberts, S. (1981). Isolation of an internal clock. *J. Exp. Psychol. Anim. Behav. Process.* 7, 242–268.

Roitman, J. D., and Shadlen, M. N. (2002). Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *J. Neurosci.* 22, 9475–9489.

Roskies, A. L. (2010). How does neuroscience affect our conception of volition? *Annu. Rev. Neurosci.* 33, 109–130.

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychol. Rev.* 103, 403–428.

Salzman, C., Murasugi, C., Britten, K., and Newsome, W. (1992). Microstimulation in visual area MT: effects on direction discrimination performance. *J. Neurosci.* 12, 2331–2355.

Salzman, C. D., Britten, K. H., and Newsome, W. T. (1990). Cortical microstimulation influences perceptual judgements of motion direction. *Nature* 346, 174–177.

Samejima, K., Ueda, Y., Doya, K., and Kimura, M. (2005). Representation of action-specific reward values in the striatum. *Science* 310, 1337–1340.

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., and Thompson, K. G. (1999). Neural selection and control of visually guided eye movements. *Annu. Rev. Neurosci.* 22, 241–259.

Schmiedek, F., Oberauer, K., Wilhelm, O., Süss, H.-M., and Wittmann, W. W. (2007). Individual differences in components of reaction time distributions and their relations to working memory and intelligence. *J. Exp. Psychol. Gen.* 136, 414–429.

Schouten, J. F., and Bekker, J. A. M. (1967). Reaction time and accuracy. *Acta Psychol. (Amst.)* 27, 143–153.

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.

Simen, P. (2012). Evidence accumulator or decision threshold – which cortical mechanism are we observing? *Front. Psychol.* 3:183. doi:10.3389/fpsyg.2012.00183

Simen, P., Balci, F., Desouza, L., Cohen, J. D., and Holmes, P. (2011). A model of interval timing by neural integration. *J. Neurosci.* 31, 9238–9253.

Simen, P., Cohen, J. D., and Holmes, P. (2006). Rapid decision threshold modulation by reward rate in a neural network. *Neural Netw.* 19, 1013–1026.

Simen, P., Contreras, D., Buck, C., Hu, P., Holmes, P., and Cohen, J. D. (2009). Reward rate optimization in two-alternative decision making: empirical tests of theoretical predictions. *J. Exp. Psychol. Hum. Percept. Perform.* 35, 1865–1897.

Sirigu, A., Daprati, E., Ciancia, S., Giraux, P., Nighoghossian, N., Posada, A., and Haggard, P. (2004). Altered awareness of voluntary action after damage to the parietal cortex. *Nature Neurosci.* 7, 80–84.

Smith, P. L. (1995). Psychophysically principled models of visual simple reaction time. *Psychol. Rev.* 102, 567–593.

Smith, P. L. (2010). From poisson shot noise to the integrated Ornstein–Uhlenbeck process: neurally principled models of information accumulation in decision-making and response time. *J. Math. Psychol.* 54, 266–283.

Smith, P. L., and McKenzie, C. R. L. (2011). Diffusive information accumulation by minimal recurrent neural models of decision making. *Neural Comput.* 23, 2000–2031.

Smith, P. L., and Ratcliff, R. (2004). Psychology and neurobiology of simple decisions. *Trends Neurosci.* 27, 161–168.

Smith, Y., Bevan, M. D., Shink, E., and Bolam, J. P. (1998). Microcircuitry of the direct and indirect pathways of the basal ganglia. *Neuroscience* 86, 353–387.

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.

Spaniol, J., Madden, D. J., and Voss, A. (2006). A diffusion model analysis of adult age differences in episodic and semantic long-term memory retrieval. *J. Exp. Psychol. Learn Mem. Cogn.* 32, 101–117.

Starns, J. J., and Ratcliff, R. (2010). The effects of aging on the speed-accuracy compromise: boundary optimality in the diffusion model. *Psychol. Aging* 25, 377–390.

Stone, M. (1960). Models for choice-reaction time. *Psychometrika* 25, 251–260.

Swensson, R. G. (1972). The elusive tradeoff: speed vs accuracy in visual discrimination tasks. *Percept. Psychophys.* 12, 16–32.

Thapar, A., Ratcliff, R., and McKoon, G. (2003). A diffusion model analysis of the effects of aging on letter discrimination. *Psychol. Aging* 18, 415–429.

Townsend, J. T., and Ashby, F. G. (1983). *The Stochastic Modeling of Elementary Psychological Processes.* Cambridge: Cambridge University Press.

Tsetsos, K., Gao, J., McClelland, J. L., and Usher, M. (2012). Using time-varying evidence to test models of decision dynamics: bounded diffusion vs. the leaky competing accumulator model. *Front. Neurosci.* 6:79. doi:10.3389/fnins.2012.00079

Tsetsos, K., Usher, M., and McClelland, J. L. (2011). Testing multi-alternative decision models with non-stationary evidence. *Front. Neurosci.* 5:63. doi:10.3389/fnins.2011.00063

Uhlenbeck, G., and Ornstein, L. (1930). On the theory of the brownian motion. *Phys. Rev.* 36, 823–841.

Usher, M., Elhalal, A., and McClelland, J. L. (2008). "The neurodynamics of choice, value-based decisions, and preference reversal," in *The Probabilistic Mind: Prospects for Bayesian Cognitive Science,* eds N. Chater and M. Oaksford (Oxford: Oxford University Press), 277–300.

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

Usher, M., and McClelland, J. L. (2004). Loss aversion and inhibition in dynamical models of multialternative choice. *Psychol. Rev.* 111, 757–769.

Van Essen, D. C., Drury, H. A., Dickson, J., Harwell, J., Hanlon, D., and Anderson, C. H. (2001). An integrated software suite for surface-based analyses of cerebral cortex. *J. Am. Med. Inform. Assoc.* 8, 443–459.

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 Ravenzwaaij, D., and Oberauer, K. (2009). How to use the diffusion model: parameter recovery of three methods: EZ, fast-dm, and DMAT. *J. Math. Psychol.* 53, 463–473.

van Ravenzwaaij, D., van der Maas, H. L. J., and Wagenmakers, E.-J. (2012). Optimal decision making in neural inhibition models. *Psychol. Rev.* 119, 201–215.

van Veen, V., Krug, M. K., and Carter, C. S. (2008). The neural and computational basis of controlled speed-accuracy tradeoff during task performance. *J. Cogn. Neurosci.* 20, 1952–1965.

Vickers, D. (1970). Evidence for an accumulator model of psychophysical discrimination. *Ergonomics* 13, 37–58.

Wagenmakers, E.-J., Grasman, R. P. P. P., and Molenaar, P. C. M. (2005). On the relation between the mean and the variance of a diffusion model response time distribution. *J. Math. Psychol.* 49, 195–204.

Wagenmakers, E.-J., Maas, H. L. J., and Grasman, R. P. P. P. (2007). An EZ-diffusion model for response time and accuracy. *Psychon. Bull. Rev.* 14, 3–22.

Wagenmakers, E.-J., Ratcliff, R., Gomez, P., and McKoon, G. (2008). A diffusion model account of criterion shifts in the lexical decision task. *J. Mem. Lang.* 58, 140–159.

Wald, A. (1947). *Sequential Analysis.* New York: Wiley.

Wald, A., and Wolfowitz, J. (1948). Optimum character of the sequential probability ratio test. *Ann. Math. Stat.* 19, 326–339.

Wallsten, T. S., and Barton, C. (1982). Processing probabilistic multidimensional information for decisions. *J. Exp. Psychol. Learn. Mem. Cogn.* 8, 361–384.

Wang, X.-J. (2002). Probabilistic decision making by slow reverberation in cortical circuits. *Neuron* 36, 955–968.

Wenzlaff, H., Bauer, M., Maess, B., and Heekeren, H. R. (2011). Neural characterization of the speed-accuracy tradeoff in a perceptual decision-making task. *J. Neurosci.* 31, 1254–1266.

Wickelgren, W. A. (1977). Speed-accuracy tradeoff and information processing dynamics. *Acta Psychol. (Amst.)* 41, 67–85.

Wiener, N. (1923). Differential space. *J. Math. Phys.* 2, 131–174.

Wong, K.-F., Huk, A. C., Shadlen, M. N., and Wang, X.-J. (2007). Neural circuit dynamics underlying accumulation of time-varying evidence during perceptual decision making. *Front. Comput. Neurosci.* 1:6. doi:10.3389/neuro.10.006.2007

Wong, K.-F., and Wang, X.-J. (2006). A recurrent network mechanism of time integration in

perceptual decisions. *J. Neurosci.* 26, 1314–1328.

Yang, T., and Maunsell, J. H. R. (2004). The effect of perceptual learning on neuronal responses in monkey visual area V4. *J. Neurosci.* 24, 1617–1626.

Yellott, J. (1971). Correction for fast guessing and the speed-accuracy tradeoff in choice reaction time. *J. Math. Psychol.* 8, 159–199.

Zeki, S. (2007). The response properties of cells in the middle temporal area (Area MT) of owl monkey visual cortex. *Proc. R. Soc. Lond. B Biol. Sci.* 207, 239–248.

Zhang, J., and Bogacz, R. (2010a). Bounded Ornstein–Uhlenbeck models for two-choice time controlled tasks. *J. Math. Psychol.* 54, 322–333.

Zhang, J., and Bogacz, R. (2010b). Optimal decision making on the basis of evidence represented in spike trains. *Neural Comput.* 22, 1113–1148.

Zhang, J., Bogacz, R., and Holmes, P. (2009). A comparison of bounded diffusion models for choice in time controlled tasks. *J. Math. Psychol.* 53, 231–241.

Zhang, J., Hughes, L. E., and Rowe, J. B. (2012). Selection and inhibition mechanisms for human voluntary action decisions. *NeuroImage.* doi: 10.1016/j.neuroimage.2012.06.058

Zhang, J., and Kourtzi, Z. (2010). Learning-dependent plasticity with and without training in the human brain. *Proc. Natl. Acad. Sci. U.S.A.* 107, 13503–13508.

Zhang, J., Meeson, A., Welchman, A. E., and Kourtzi, Z. (2010). Learning alters the tuning of functional magnetic resonance imaging patterns for visual forms. *J. Neurosci.* 30, 14127–14133.

Zhou, X., Wong-Lin, K., and Philip, H. (2009). Time-varying perturbations can distinguish among integrate-to-threshold models for perceptual decision making in reaction time tasks. *Neural Comput.* 21, 2336–2362.

Conflict of Interest Statement: The author declares 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: 24 January 2012; accepted: 08 July 2012; published online: 01 August 2012.

Citation: Zhang J (2012) The effects of evidence bounds on decision-making: theoretical and empirical developments. Front. Psychology ***3****:263. doi: 10.3389/fpsyg.2012.00263*

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

frontiers in PSYCHOLOGY

HYPOTHESIS AND THEORY ARTICLE
published: 21 June 2012
doi: 10.3389/fpsyg.2012.00183

Evidence accumulator or decision threshold – which cortical mechanism are we observing?

Patrick Simen*

Department of Neuroscience, Oberlin College, Oberlin, OH, USA

Edited by:
Konstantinos Tsetsos, Oxford University, UK

Reviewed by:
Christian C. Luhmann, Stony Brook University, USA
Jiaxiang Zhang, Medical Research Council, UK
Andrei Teodorescu, Tel-Aviv University, Israel

****Correspondence:***
Patrick Simen, Department of Neuroscience, Oberlin College, Oberlin, OH 44074, USA.
e-mail: psimen@oberlin.edu

Most psychological models of perceptual decision making are of the accumulation-to-threshold variety. The neural basis of accumulation in parietal and prefrontal cortex is therefore a topic of great interest in neuroscience. In contrast, threshold mechanisms have received less attention, and their neural basis has usually been sought in subcortical structures. Here I analyze a model of a decision threshold that can be implemented in the same cortical areas as evidence accumulators, and whose behavior bears on two open questions in decision neuroscience: (1) When ramping activity is observed in a brain region during decision making, does it reflect evidence accumulation? (2) Are changes in speed-accuracy tradeoffs and response biases more likely to be achieved by changes in thresholds, or in accumulation rates and starting points? The analysis suggests that task-modulated ramping activity, by itself, is weak evidence that a brain area mediates evidence accumulation as opposed to threshold readout; and that signs of modulated accumulation are as likely to indicate threshold adaptation as adaptation of starting points and accumulation rates. These conclusions imply that how thresholds are modeled can dramatically impact accumulator-based interpretations of this data.

Keywords: decision, threshold, accumulator, integration, switch, reward, sequence

1. THE THRESHOLD CONCEPT IN ABSTRACT DECISION-MAKING MODELS

Simple perceptual decision making is typically thought to involve some kind of weighing of evidence. According to this story, sensory data is repeatedly sampled in order to build confidence that one decision option is correct, and the others incorrect.

Accumulating evidence is only half the battle when making a decision, however. The other half requires that a subject have a physically implemented policy governing the termination of evidence collection and the transition into action. Indeed, many of the most important results in decision theory consist of carefully designed stopping rules that terminate decisions. These policies result in hypothesis-testing procedures that are optimal or near-optimal according to some objective function. Psychological modelers are therefore naturally guided to focus on decision rules as an important source of performance adaptation in humans and other animals.

A common view in neuroscience, in contrast, is premised on the routine finding that firing rates of suspected accumulator-neurons reach a fixed threshold near the time of a behavioral response, suggesting that the critical level that causes responding or decision commitment (in monkeys at least) is close to the peak firing rate. Performance by these animals (their response times and accuracy) is therefore thought to be adapted by changing the rate, starting point, or starting time of evidence accumulation.

Once constraints on how networks of neurons can compute are taken into consideration, pressing theoretical questions arise for both views of performance adaptation. First, how can accumulating evidence produce a behavioral response when it exceeds a critical level by some small value, but not when it is just slightly below that level? Further, once a model of a threshold mechanism is proposed that answers this question, how can physiological signs of its operation be discriminated from signs of evidence accumulation? Finally, is behavioral performance adapted by modulating the activity of accumulators, thresholds, or both?

Here I formally define some basic, unavoidable physical assumptions about decision-threshold mechanisms, and I consider how these must affect interpretations of neural data. I argue that firing-rate data commonly thought to be observations of accumulators in action may instead be observations of threshold mechanisms, and that mistaken identities of this sort may be the source of an apparent conflict between findings in neuroscience (fixed thresholds) and findings in psychology (strategically controlled thresholds).

I show that this conflict eases if neural accumulators have sometimes been confused with neural threshold mechanisms. As I demonstrate with a simple decision-making model, the final-firing-rate-premise of the fixed-threshold account actually implies very little, if we allow for threshold mechanisms to send positive feedback to accumulators – a concept consistent with anatomical patterns of connectivity in parietal and prefrontal cortex. With this addition to an otherwise purely feedforward model, threshold mechanisms can serve to lift accumulators to a common final level of activation near the time of responding, even when different levels actually trigger decision commitment under different task conditions. These theoretical considerations suggest that physiological data is frequently ambiguous regarding the locus of decision-circuit control.

2. IMPLICATIONS FOR MAPPING DECISION-MAKING PROCESSES ONTO BRAIN ACTIVITY

In this section, I present a generic model of a neural threshold mechanism inspired by the individual neuron's action potential threshold. I then discuss what its implementation at the cortical population level would predict in terms of a firing-rate profile during threshold-crossing (while acknowledging that this threshold model could in fact be implemented subcortically). If the model is correct, these predictions imply that some current interpretations of single-unit recordings from monkeys performing perceptual decision-making and visual search tasks may need to be revised.

2.1. NEURAL IMPLEMENTATIONS OF ABSTRACT DECISION-MAKING MODELS REQUIRE PHYSICAL THRESHOLD MECHANISMS

Decision thresholds traditionally play an important role in psychology in explaining one of the most salient features of human perceptual categorization. This is that spending more time observing a stimulus tends to increase decision accuracy regarding which category the stimulus belongs to. Such speed-accuracy tradeoffs are easily explained as a process of accumulation to threshold: to make a decision, evidence must accumulate to a point that a decision threshold is crossed and an action is triggered. Higher thresholds imply greater accuracy (less chance of crossing the wrong threshold), but longer response times (the decision variable has a greater distance to travel). Furthermore, just as incentives are hypothesized to change category boundaries in signal detection theory (Green and Swets, 1966), top-down control strategies have been hypothesized to adapt decision thresholds in response-time tasks in order to modify speed-accuracy tradeoff functions (see Luce, 1986)[1].

This basic scheme of accumulation-to-threshold is implemented by most decision models, although even simpler, non-evidence-accumulating models can account for speed-accuracy tradeoffs (e.g., the "urgency-gating" model of Cisek et al., 2009, which responds whenever an unusually favorable sample of evidence arrives and weights evidence by a ramping urgency signal; or purely ballistic models that sample evidence only at one instant, such as the ballistic accumulator model of Brown and Heathcote, 2005, and the LATER model of saccade response times, Reddi and Carpenter, 2000). In addition, recent accumulation-based modeling approaches attempt to account for physiological evidence of fixed thresholds by adapting baseline levels of activity in competing response channels. Adapting baseline activity results in an effective change of threshold height without any change in the level of channel-activation necessary to make a response (Bogacz et al., 2010b; see also van Ravenzwaaij et al., 2011). This approach is similar but not identical to decision-threshold adaptation. Activating a response channel in these models must initiate a decision process based on accumulated noise that will ultimately culminate in a decision, even if no stimulus is present. Actual decision-threshold adaptation, in contrast, can be achieved without producing any response until a stimulus is present, allowing for top-down control to be exerted over arbitrarily long delays prior to stimulus onset.

In all cases, however, decisions must not be initiated before some critical level of evidence or some other quantity either accumulates or is momentarily sampled. Before committing to a particular decision – a period that may theoretically last an arbitrarily long time – the motor system is often assumed to receive no input from the evidence-weighing process. Thus a physical barrier must be assumed, which, once exceeded, leads inexorably to a particular outcome, but below which no response is possible. What sort of non-linear transformation of the net evidence or the "urgency to respond" can meet this specification and be implemented physically?

The simplest answer is: the same sort of transformation implemented by threshold-crossing detectors in human-engineered systems, namely *switches*. Physically implemented switches nearly always have two, related, dynamical properties – bistability and hysteresis – that define them specifically to be *latches* in engineering terminology. Bistability means that these systems are attracted to one of two stable states that are separated by an unstable equilibrium point; hysteresis means that the response of such a system to a given input depends heavily on its past output (loosely speaking, hysteresis means "stickiness" and involves a basic form of persisting memory of the past; in contrast, linear systems are non-sticky and respond to constant inputs in such a way that the system's initial conditions are forgotten at an exponential rate over time). Energy functions can be defined for such systems, consisting of two wells separated by a hump (see **Figure 1**). Any such system can then be accurately visualized as a particle bouncing around inside one or the other well under the influence of gravity, and occasionally escaping over the hump into the other well. Each "escape" is analogous to the flipping of the switch. The importance of the double-well design is that it reduces chatter, or bouncing of the switch between states as a result of noise (it "latches" into one or the other state), imposing a repulsive force away from the undefined region between ON and OFF.

It is important to note that such devices do not implement a simple step-function applied to their inputs, as suggested by a

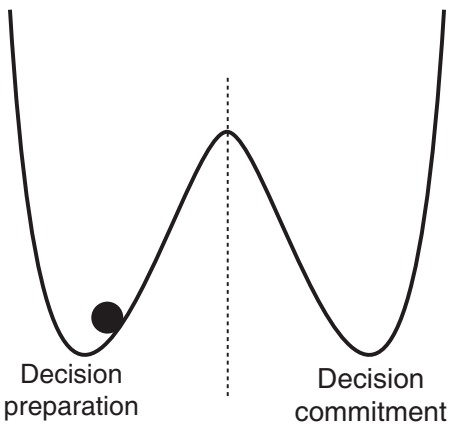


FIGURE 1 | Double-well energy potential function, with system trapped in left well. Transitions to the right well ("escapes") maybe considered transitions from an OFF state to an ON state, or a decision-preparation to a decision-committed state.

[1]Note that the use of the term "threshold" here is equivalent to that term's use in neurophysiology. It is distinct from its common use in psychology as the name of an arbitrary accuracy criterion for characterizing perceptual acuity.

common, simplified representation of switching behavior in the form of a step-function (or "Heaviside function"). Instead, their input-output relationships are not strictly functions at all: a given input level maps to more than one output level. Step-functions are nonetheless frequently used to formalize the ideal of an abstract decision threshold. This idealization is frequently very useful in both psychology and neuroscience, but it has the potential to bias the interpretation of various kinds of physiological recordings, since these are usually easier to map onto the gradual accumulation of evidence than onto the sudden, quantum leap represented by a step-function.

An analogous problem holds in the design of digital electronics: logic engineers who use idealized switch representations must factor in lower limits on the switching-speed of their basic components in their circuit designs. If they violate these constraints on assumed switching speeds, then their assumptions that a particular component will output a 1 or 0 at a given time may in turn be violated (the output may not yet have changed from its previous value, or it may be in an indeterminate level representing neither 1 nor 0). The end result is unpredictable circuit behavior (Hayes, 1993). This analogy suggests that when it comes to interpreting, for example, single-unit firing-rate data in decision making, we may be in the position of the physicists and engineers who design the transistors, rather than those who compose switching devices into large circuits. Using idealized models of switching at the sub-digital scale leads to substantial errors of prediction in electronics; making the same sort of mistake in psychology and neuroscience may lead to substantial errors of interpretation.

2.2. THE AXONAL MEMBRANE IMPLEMENTS A PHYSICAL THRESHOLD

Another constructive analogy for modeling decision-making circuits comes from a well known physical threshold device even more familiar to neuroscientists: the axonal membrane of an individual neuron. In central nervous system neurons, a high density of voltage-gated sodium-channels in the axon imparts bistability to the voltage across its membrane. This channel-density is typically highest at the axon hillock, where the axon leaves the soma, so that most action potentials are generated in this area, known as the "trigger zone" (Kandel et al., 2000). In contrast, the soma itself typically has a much higher (possibly non-existent) action potential threshold, and is classically thought to function more as a spatiotemporal integrator than as a switch.

According to the deterministic Hodgkin-Huxley equations, raising the membrane to nearly its threshold potential and then shutting off input current leads to a return to the negative resting potential – the axon-potential's low stable value – without an action potential. Physically, this corresponds to a failure to trigger a chain reaction of voltage-gated sodium-channel openings. Injecting a current that is slightly larger instead triggers this chain reaction with high probability, producing an action potential that is stereotyped in magnitude and duration under fixed conditions of temperature and chemical concentration. During the action potential, the membrane traverses a no-man's-land of positive voltages, reaches its stereotyped peak – its high stable value – and then resets to its low stable value as voltage-gated potassium channels open up. These are hallmarks of a double-well system.

The primary player in this all-or-none process is a voltage-amplifying mechanism with strong positive feedback – the voltage-gated sodium-channel population – which ranges between two stable, attracting values: all-open (1), and all-closed (0). A time-delayed version of these activation dynamics is then employed by a voltage-gated potassium-channel population to produce a shutoff switch. A central claim of this paper is that these roles may easily be played by any neural population conforming to certain assumptions, and that some cortical populations may indeed conform to them.

This important feature of individual membrane potential dynamics does more than provide a useful analogy for modeling decision-making circuits. It suggests that a simple mathematical model may be usefully employed to describe both the dynamics of the individual neuron and the dynamics of interconnected neural populations. The generation of an individual action potential in a neuron, after all, qualitatively fits the description of a typical decision process: sub-threshold, leaky integration of post-synaptic potentials in a neural soma is analogous to evidence accumulation in a decision circuit; action potential generation in the trigger zone is analogous to crossing a decision threshold.

The other major claim of this paper regards interpretation of behavioral and physiological data collected during decision making. Many researchers (e.g., Shadlen and Newsome, 2001; Purcell et al., 2010) describe their findings in ways that suggest they are observing the population-analog of a neuron's somatic membrane potential, when a better analogy in some cases may be that they are observing the analog of an *axon's* membrane potential.

2.3. "TURTLES ALL THE WAY DOWN": NEURONS AND NEURAL POPULATIONS AS THRESHOLDED LEAKY INTEGRATORS

Since so much inspiration for the present model of decision-making circuits comes from the individual neuron, I now show briefly how the dynamics of the individual neuron can, in principle, be reproduced by whole populations. **Figure 2** depicts how a population of leaky-integrate-and-fire (LIF) neurons can collectively implement a single leaky integrator whose output is the population's spike train. Suppose that many neurons project to a given receiving population and have uncorrelated spike times; that the resulting post-synaptic potentials are small due to weak synaptic connections; and that excitation is balanced by some level of inhibition. In that case, each unit in the receiving population can be modeled as a leaky integrator (**Figure 3**), whose sub-threshold membrane potential approaches some asymptotic level (see Smith, 2010). A similar model of sub-threshold membrane dynamics as a drift-diffusion process (without leak) was given in Gerstein and Mandelbrot (1964).

If the receiving population is large enough, then the receiving population's spiking output represents what the population's average membrane potential would be if its units lacked voltage-gated sodium-channels and therefore generated no action potentials. An asymptotic potential would be reached, with the level depending on the input strength and the leak. (Action potentials naturally tend to erase the record of previous potential levels in each unit after resetting.)

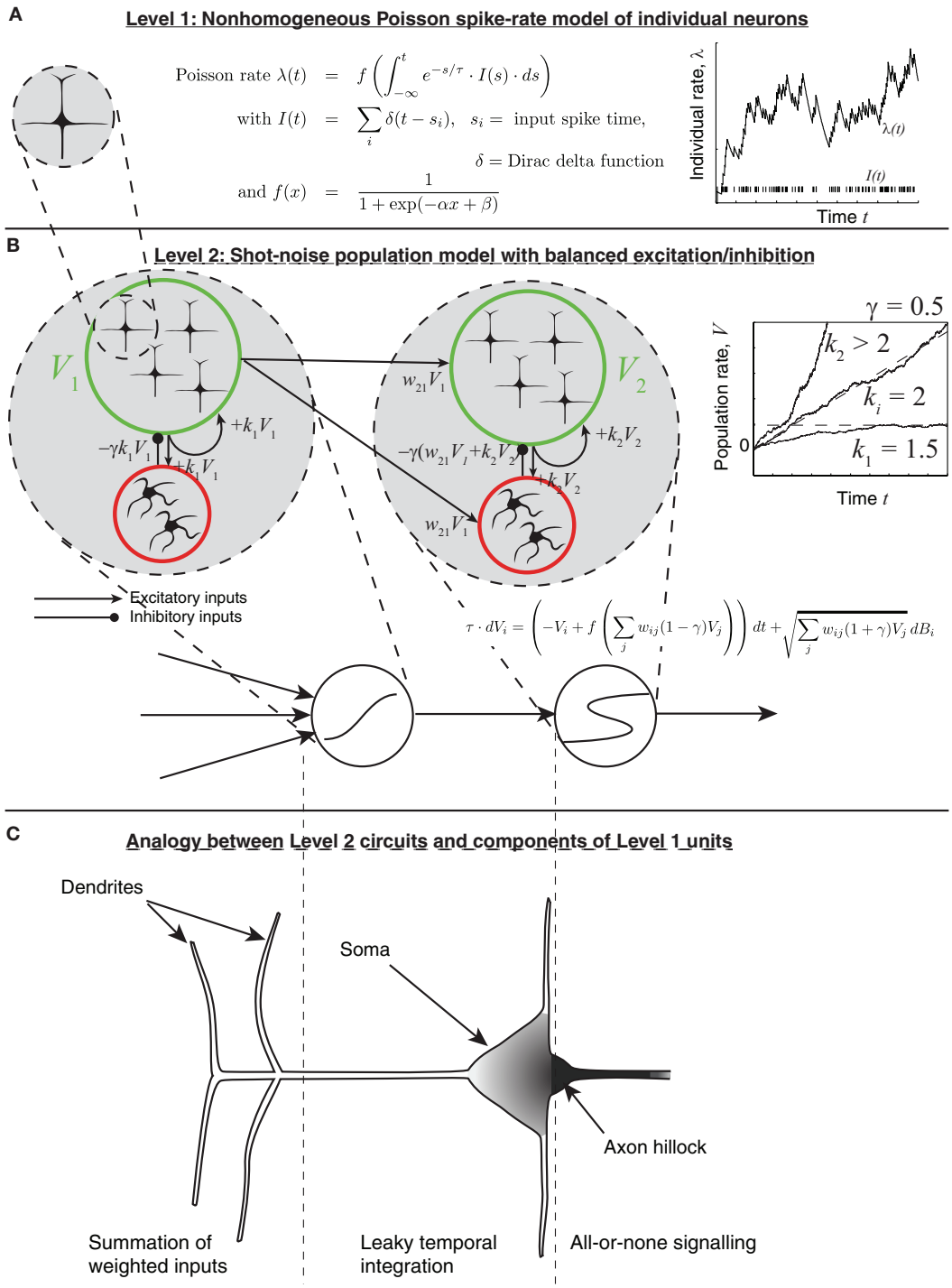


FIGURE 2 | The multiple spatial-scale modeling approach. (A) For simplicity, an individual neuron is modeled as a Poisson spike generator, with rate parameter determined by the leaky integral of weighted inputs from other units. **(B)** The population model, which allows for chaining a leaky integrator and a latch. **(C)** Correspondence of each circuit component to parts of a more classical model of a neuron. The axon hillock corresponds to a latch; the soma corresponds to a leaky integrator; shading indicates density of voltage-gated sodium-channels.

If this asymptotic level is below the action potential threshold, however, then the inter-spike times of the model will be largely memoryless. The hazard rate of a new spike is almost flat, meaning that if a spike has not occurred at some time t following the previous spike, then it has a constant probability of occurring in the next small time window,

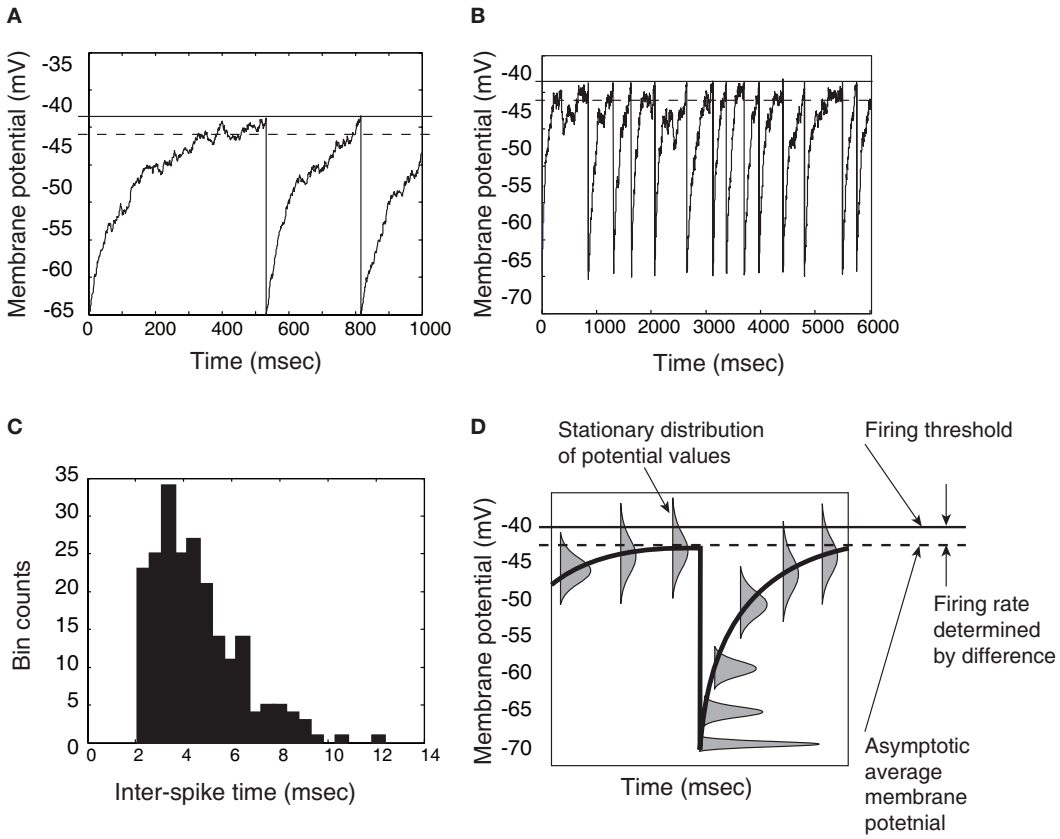


FIGURE 3 | Demonstration that a thresholded Ornstein-Uhlenbeck (OU) process, or leaky integrator, can give approximately exponential inter-spike times if the asymptotic level of the integrator is below threshold. (A) Close-up view of an OU process rising to a threshold, then resetting and rising again. **(B)** A longer time-course of the same process; inter-spike times are the times between resets (downward jumps). **(C)** Histogram of threshold-OU model's inter-spike intervals. The main deviation from true Poisson statistics is the lack of inter-spike intervals below a 2 ms limit – in other words, a refractory period. **(D)** Distributions of trajectories relative to spike time, showing that aside from the brief refractory period when the system has not reached asymptote (on average), the model is essentially a white noise process positioned just under the action potential threshold, leading to memoryless inter-spike times. The rate of those spikes depends on the distance between the threshold and the sub-threshold asymptote (Gardiner, 2004).

for all $t > 0$. This occurs because threshold-crossings will be due only to momentary noise. Formally, this is once again a problem of escape from an energy well. Note, here, however, that an idealized step-function threshold is employed for its simplicity. For the purpose of modeling population activity over the course of a perceptual decision as an emergent property of collective spiking, the assumption here is that the biophysical details of action potential generation make no difference.

The first-passage time distribution of this system would be approximately exponential (Gardiner, 2004), deviating from the exponential mainly by lacking a high probability of very short inter-spike times: the resetting of the membrane to its resting potential, followed by increase toward asymptote, must produce a refractory period that would prevent truly exponential inter-spike time distributions (see **Figure 3C**). The model's average inter-spike interval is an exponential function of the distance between the firing threshold and the asymptotic average potential (Gardiner, 2004): stronger inputs to a unit lead to a higher asymptote and higher firing rate. If inputs lift the asymptote above the firing threshold, the model breaks down, and inter-spike time distributions approach the Wald distribution (Gerstein and Mandelbrot, 1964).

Thus we can recreate the dynamics of the individual neuron at the population level by building a circuit consisting of a leaky integrator population feeding into a switch population (see the mapping from **Figures 2B,C**).

2.4. THE FORMAL MODEL OF POPULATION ACTIVITY

The model is based on the assumptions outlined above and in more detail in Simen and Cohen (2009); Simen et al. (2011b), most of which are common in neural network models of decision making. It is assumed that neural population activity (firing rate) can be represented approximately as a non-homogeneous Poisson spike train (equation 1). Its rate parameter λ is governed by an ordinary differential equation with a leak term and a sigmoidal activation function f. The input I to the activation function is a shot-noise process, equal to the weighted sum of input spikes received from

other populations, convolved with an exponential decay term (see **Figure 2A**):

$$\begin{aligned} \text{Poisson rate } \lambda(t) &= f\left(\int_{-\inf}^{t} e^{-s/\tau}\cdot I(s)\cdot ds\right) \\ \text{with } I_i(t) &= \sum_j w_{ij}\cdot\delta\left(t-s_j\right), \\ \text{and } f(y) &= \frac{1}{1+\exp\left(-\alpha y+\beta\right)} \end{aligned} \tag{1}$$

Here, w_{ij} = connection strength between input j and unit i, s_j = input spike time, δ = Dirac delta function.

The resulting population model approximates a leaky integrator, or stable Ornstein-Uhlenbeck (OU) process $x(t)$ (defined by equation (2)), as long as the input $I(t)$ is positive, and as long as inputs remain in the linear region of the sigmoidal activation function f:

$$\begin{aligned} \tau\cdot dx_i &= \left(-x_i+f(I_i)\right)\cdot dt + c\cdot dW_i, \\ I_i &= \sum_j w_{ij}x_j. \end{aligned} \tag{2}$$

Unlike a true OU process, when the population's net input I is negative, its output goes to 0, but not below it.

For readers not familiar with stochastic differential equations, the following discrete time approximation to equation (2) may be easier to understand:

$$\tau\cdot x_i\left(t+\Delta t\right) = x_i\left(t\right)+\left[-x_i\left(t\right)+f\left(I_i\left(t\right)\right)\right]\cdot\Delta t + c\sqrt{\Delta t}\cdot\mathcal{N}(0,1). \tag{3}$$

Here, a normal random variable is added to the deterministic system at each time step, with magnitude weighted by the square root of the time step size. Equation (3) also defines the "Euler method" by which simulations of the system in equation (2) are carried out (Gardiner, 2004).

When recurrent self-excitation, $k \equiv w_{ii}$, is included in the input term I, the population's intrinsic leak parameter can be reduced or, at a higher level of precisely tuned positive feedback, completely canceled (Seung, 1996). Exact cancelation gives rise to perfect (leakless) temporal integration, which is critical for implementing decision-making algorithms that are optimal in certain contexts (Gold and Shadlen, 2002). Such exact cancelation requires more precise tuning for larger leak terms (Simen et al., 2011a). At still higher levels of positive feedback, bistability results (see **Figure 4**). The likely consequence for an explicit, purely excitatory, spiking LIF model of populations in the case of stronger self-excitation would be synchronization of spiking, since stronger excitatory connectivity in a population would make any given spike more likely to produce another spike in another cell at nearly the same time. However, we assume further that a form of balanced inhibition can cancel spike-time correlations (Simen et al., 2011b). Under this assumption, any excitatory input of magnitude M to a population would be balanced by inhibitory input of magnitude γM, with γ defining the excitatory-inhibitory ratio. That is, if every excitatory spike received by a population produces an excitatory post-synaptic potential of magnitude 1 in one of its cells, then that population also receives an inhibitory post-synaptic potential of magnitude γ in one of its cells, on average.

These changes in self-excitation strength can be depicted in the form of an "effective" activation function, $f_{\gamma,k}$, specifying the output level of firing rate for a given level of net weighted input strength, parameterized by the level of positive, recurrent feedback, k, and the inhibitory balance, γ. As the model transitions from leaky (**Figure 4A**) to bistable (**Figure 4C**) behavior through the bifurcation point at which perfect integration occurs (**Figure 4B**), this effective activation folds back on itself (**Figure 4D**). This folding is known as a "cusp catastrophe" in the terminology of non-linear dynamical systems.

The bistable behavior demonstrated in **Figures 4C,E** now serves to establish a threshold level of input values: beginning with an output level on the bottom stable attracting arm of the folded sigmoid, inputs that exceed a critical level (the horizontal coordinate of the bottom fold) trigger a transition toward the upper stable attracting arm. Two key dynamical features now result. The first is a clearly defined quantization – in fact, binarization – of output levels. The upper stable arm of the folded sigmoid now represents an ON state (**Figure 4E**); the lower stable arm represents an OFF state; and potential confusion about which state the system is in is reduced by a large no-man's land of unstable activation levels. Hysteresis also results: a dip in the input below the threshold level does not now reduce the output to its OFF state; instead, it remains ON until input dips below the horizontal coordinate of the upper fold.

Figure 5 demonstrates an extremely simple circuit composed of an integrator sandwiched between two latches, along with their corresponding, predicted activation levels over time.

2.5. BEHAVIOR OF THE THRESHOLD MECHANISM, AND COMPARISON TO FIRING-RATE DATA

Since we are concerned with discriminating between accumulators and threshold mechanisms in perceptual decision making, a critical question is the following: what should we expect the population activity corresponding to each mechanism to look like? Previous work (e.g., Lo and Wang, 2006; Boucher et al., 2007) has suggested that, in the case of eye movements, a successive sharpening of bursting activity occurs as activity propagates from cortical circuits to the basal ganglia, and thereafter to the superior colliculus. That is, bursting activity begins more abruptly, with less gradual ramping, and also ends more abruptly, at later stages of processing.

The model shown in **Figure 6** illustrates similar progressive sharpening as signals progress from the accumulator to the threshold latch[2]. **Figure 6** also clearly shows, however, that the threshold mechanism's activation (in red) qualitatively matches the description of an accumulator, in that the level of activation rises throughout the stimulus presentation to a maximum that is time-locked to the response. Compared to the accumulator in this trace (green), the threshold is very distinct, since its activation accelerates during

[2]Matlab code for this model is available at: www.oberlin.edu/faculty/psimen/ThresholdModelCode.m

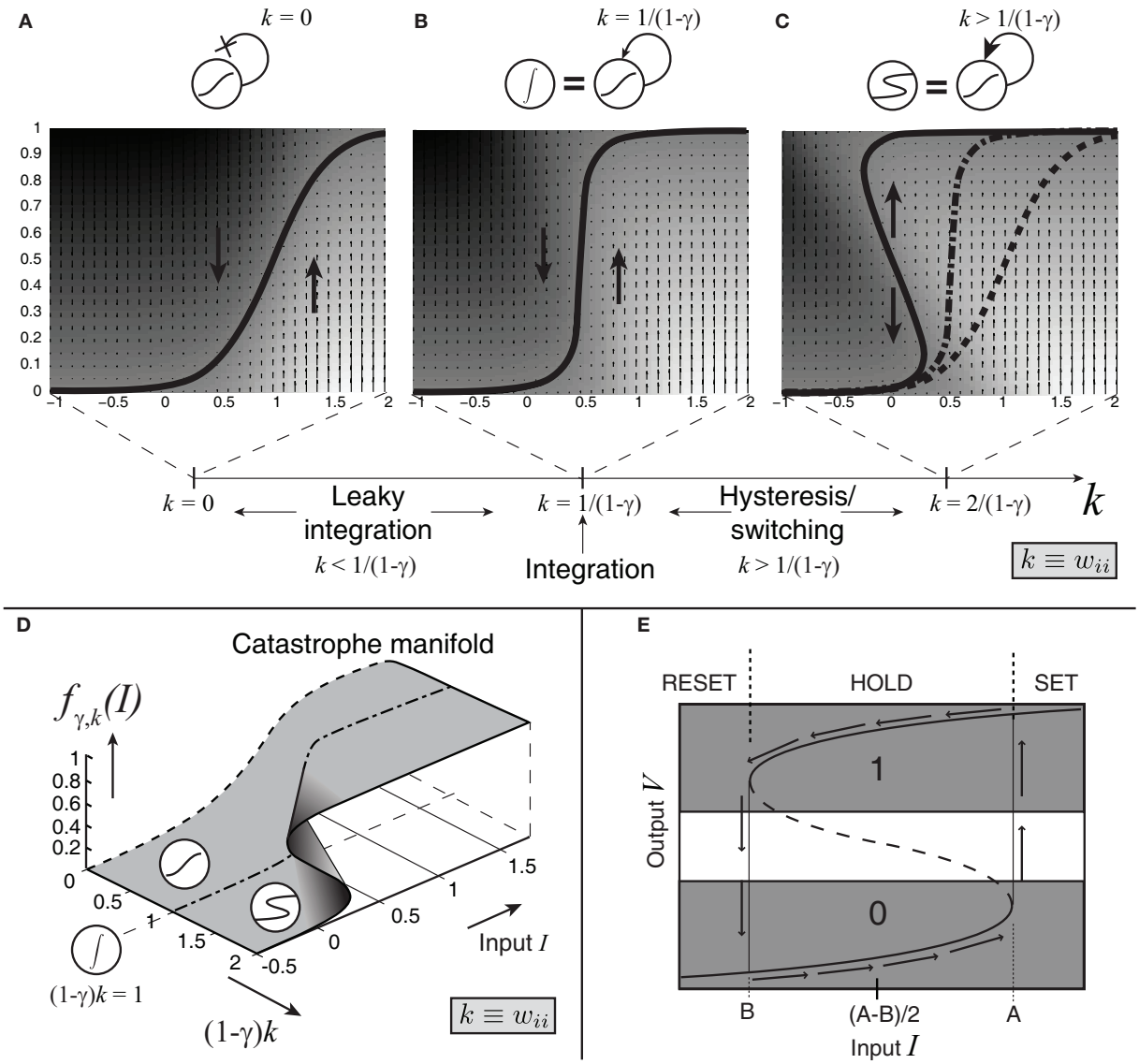


FIGURE 4 | Effect of self-excitation on the sigmoidal activation unit defined by equation 2. (A) The input-output equilibrium function $f_{\gamma,k}$ of the system (thick black curve) for the sigmoidal unit with self-excitation $k = 0$ and noise coefficient c set to 0; when $c > 0$, this function still describes the average behavior of the system. Upward and downward arrows and shading indicate system velocity at different input-output combinations (dark = negative; light = positive). **(B)** Self-excitation $k = 1/(1 - \gamma)$ cancels leak and produces a vertical effective activation function and leakless temporal integration [$k < 1/(1 - \gamma)$ produces leaky integration]. **(C)** Latching/switching behavior occurs for all $k > 1/(1 - \gamma)$; here $k = 2/(1 - \gamma)$. **(D)** A section of the complete catastrophe manifold defined over the space of input and self-excitation pairs. **(E)** Binarization diagram demonstrating separated ON (1) and OFF (0) areas of output activation and hysteresis. Dashed section of S-curve denotes unstable equilibria; solid sections denote stable equilibria.

the stimulus into a final ballistic component. However, without the green trace for comparison, the red threshold trace would be difficult to rule out as the signature of an evidence accumulator. Furthermore, as shown later, task variables that modulate the ramping activity of putative neural accumulators should also be expected to modulate threshold-unit activity, compounding the difficulty of discriminating between accumulator and threshold mechanisms.

One additional feature of the model in **Figure 6** – the "shutoff latch" component (black) – is worth mentioning at this point, since without it, connection weights in the circuit can easily be tuned to preserve a decision commitment over a delay period (note how long the threshold mechanism is active in **Figure 5**, for example). Without a shutoff signal, the circuit could make a decision commitment and maintain it indefinitely (the red threshold latch could latch into the ON state). Such behavior would be required in working memory tasks, and may occur even in tasks for which such behavior would seem to be suboptimal (e.g., Kiani et al., 2008). For tasks that require immediate responding and resetting for future decisions, however, adding a shutoff latch increases reset speed and can prevent undesired latching of the threshold-unit.

Examining traces of FEF activity during visual search tasks for neurons that have been interpreted as accumulators (reprinted in

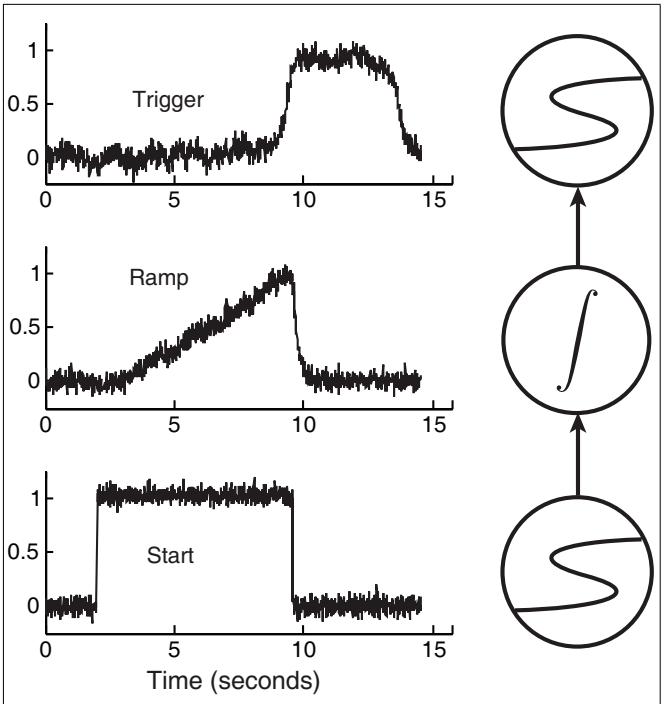


FIGURE 5 | A simple circuit composed of latches and an integrator. This particular circuit can be used for interval timing by tuning the slope of the ramp, which reduces to tuning how hard the Start unit drives the integrator (Simen et al., 2011b). Here the Start unit is clamped to an output of 1 at 2 s, then is clamped to 0 when the threshold-unit exceeds an output of 0.8.

Figure 7) shows how similar their average activation pattern is to the pattern produced by the threshold model in **Figure 6** (red trace). Much the same is true for FEF recordings collected during dot-motion discrimination tasks, as shown, for example, in Figure 2 of Ding and Gold (2012, reprinted here as **Figure 8**). In the model, furthermore, differences in evidence-accumulation rates lead to corresponding differences in the ramp-up rate of the threshold mechanism (see **Figure 9C**). Stochastic model simulations in **Figure 9C** show that changes in the threshold mechanism activation naturally mimic changes in the neural firing-rate data in **Figure 7** across fast/slow response-time conditions, as well as the firing-rate changes in **Figure 8** across motion coherence-level conditions.

The pattern of activity shown in **Figure 7** also motivates a theoretical alternative to the threshold-interpretation proposed in this paper. This alternative is known as the "gated accumulation" theory of FEF movement neuron activity (Purcell et al., 2010). According to this theory (at least as applied to monkey eye movements during visual search tasks), evidence accumulation is a process that may happen close to the time of the response, after a kind of initial quality check determines whether stimulus information should pass through a gate into the accumulator. Changes in the rate of this accumulation determine response time.

"Gate" is another word for threshold. Thus the gated accumulation theory incorporates a threshold mechanism between the retina and the accumulator that is not directly observed in the recordings of **Figure 7**. What seems more parsimonious and consistent with the predictions of the model presented here, however, is that gated accumulators are themselves the gates. The rise-time of any physical switching mechanism must be non-zero, so ramping, or accumulation, is probably a necessary feature of gate dynamics. **Figures 9B,C** show an example in which stronger, faster, negative feedback is applied to the threshold latch by the shutoff latch, as compared to the system in **Figure 6** (to reduce the height of the peak response). When noise is added to the processing, the stimulus-locked and response-locked averages of the threshold mechanism's activation in this case look remarkably like the gated-accumulator in **Figure 7**. In **Figure 9**, three different levels of input signal were applied to the model in **Figure 6** while keeping the noise level constant.

One objection to this characterization of the (Purcell et al., 2010) data might be based on the relatively long switching time of such FEF switches, which according to **Figure 7** could exceed 100 ms. Neural populations in the brain are clearly capable of transitioning much more quickly from low firing rates to high firing rates, after all. For example, Figure 3 of Boucher et al. (2007) shows a much more rapid transition from low to high firing rates, over the course of only a few msec, in bursting neurons in the brainstem. This extremely rapid switching behavior contrasts with hypothesized switch dynamics in FEF that take 100 ms to complete a transition from low to high firing rates. Such extended switch-on times might therefore reasonably be taken to support accumulator dynamics that do not involve bistability. However, it is important to note that any bistable positive feedback system can be made to linger at its inflection point of activation (the lower fold of the cusp catastrophe manifold) for arbitrarily long times, if the inputs to it are tuned precisely enough. The first gates/thresholds in the processing cascade, when faced with weak signals, may be expected therefore to ramp-up very gradually in some task conditions. Furthermore, time-locked averaging of abrupt activation-increases across trials can in any case smear out the abrupt onsets into a gradual ramp if the onset-times are not perfectly locked to the average-triggering event. **Figure 9** demonstrates similar smearing in panel b, in which Threshold Layer activation seems to rise to a different average level in the "Low" vs. the "High" condition, whereas response-locking as in panel **c** shows that activation rises to the same level in all conditions, and ramping is less gradual than in panel b.

Given these considerations, what phenomena would clearly distinguish threshold mechanisms from accumulator mechanisms? So far, the green accumulator component of the model in **Figure 6** lacks the red threshold trace's late, ballistic component, and the upward inflection point that initiates that component. This difference might seem to be a useful feature for distinguishing between accumulator and threshold activations. Unfortunately, there is good reason to suspect that the lack of feedback connections from the threshold mechanism to the accumulator in the model in **Figure 6** is unrealistic. There are known to be connections from FEF back to parietal cortex and extrastriate visual cortex that can conduct spatial attentional signals, for example (Moore and Fallah, 2001; Moore and Armstrong, 2003). When such feedback connections are included, as in **Figure 10**, the ballistic component produced in the threshold mechanism is transferred

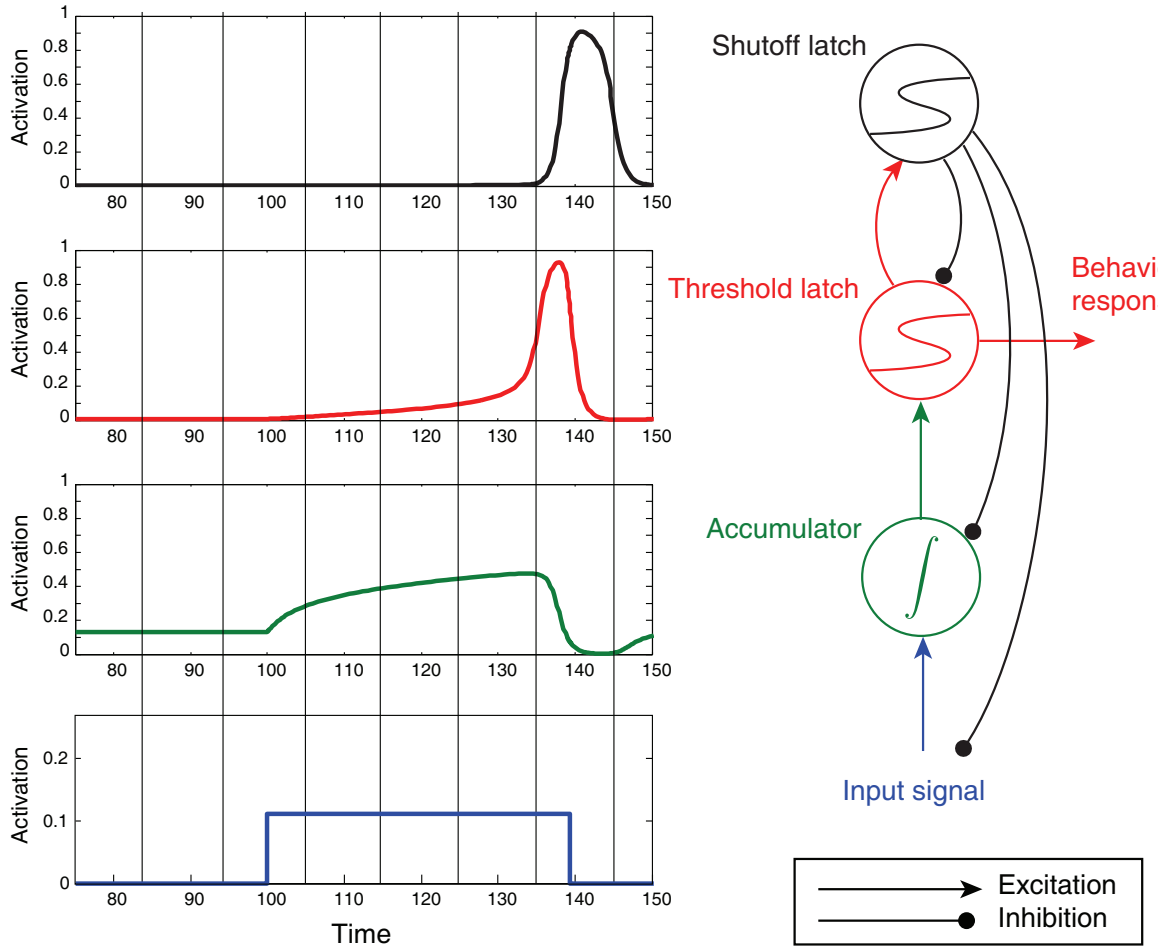


FIGURE 6 | Noiseless activation traces of a complete circuit comprising an accumulator/integrator, threshold latch, and shutoff latch, but lacking positive feedback from threshold-to-accumulator. Blue: input signal; green: accumulator; red: threshold latch; black: shutoff latch. Here the stimulus was arbitrarily terminated when the threshold latch exceeded 0.8 activation. Note the post-response dip below baseline in the accumulator.

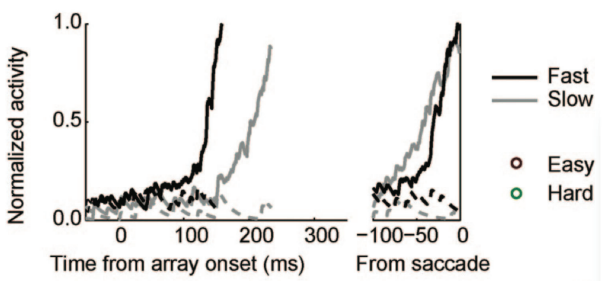


FIGURE 7 | Reprint of Figure 10A, from Purcell et al. (2010), showing movement neurons in monkey FEF whose activation profiles correlate with reaction-time in a visual search task.

backward to the accumulator as well. Under this plausible scenario, the activation profiles of evidence accumulators and threshold mechanisms become qualitatively indistinguishable.

According to the model in **Figure 10**, activation in both components should ramp up during the course of a decision, then finish with a ballistic jump near the time of the decision commitment. As **Figure 11** shows, activation in both components of this model should be modulated by the stimulus presented and the choice ultimately made – two criteria commonly used to associate firing-rate data with accumulators. Furthermore, activation should be modulated by the incentives offered, regardless of where in the circuit any incentive-induced control signals are actually applied. Applying a continuous, constant, additive biasing signal, either to the accumulator (**Figure 11A**) or to the threshold (**Figure 11B**), has qualitatively similar effects on the pre-ballistic components of ramping activation in both mechanisms.

One possibility for discriminating between two candidates for an accumulator/threshold pair would be to microstimulate one area while inactivating the other, e.g., with muscimol. A threshold mechanism should continue to display bistability in response to increasing current injections when accumulators are inactivated, whereas an accumulator should lack bistability and hysteresis – its response to increasing currents should be a monotonic function of that current when the feedback excitation from the corresponding threshold mechanism is disabled.

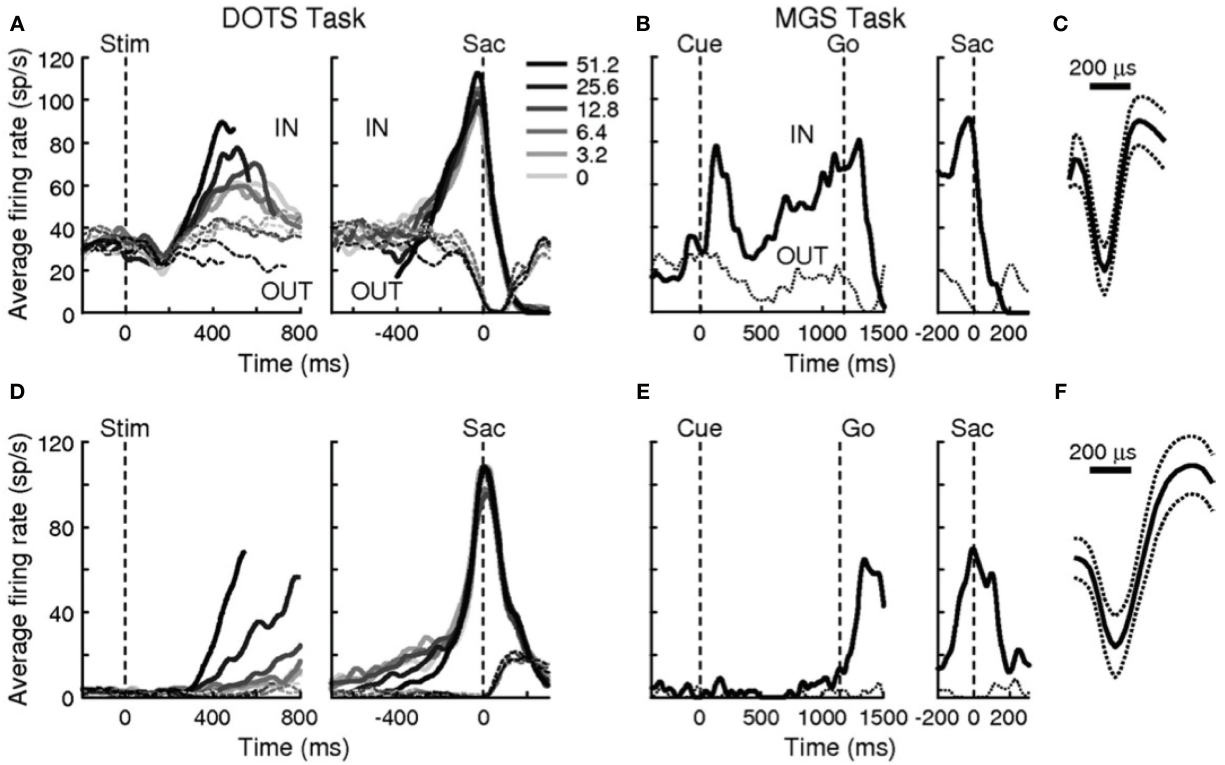


FIGURE 8 | Reprint of Figure 2, Ding and Gold (2012), showing accumulator-dynamics in FEF neurons likely to be inhibitory (top row) and excitatory (bottom row). Data were obtained during a reaction-time version of a dot-motion discrimination task **(A,D)** and a memory-guided saccade task **(B,E)**. **(C)** The narrow action potential profile of the presumably inhibitory interneuron in **(A,B)**. **(F)** The broader action potential profile of the presumably excitatory pyramidal cell in **(D,E)**.

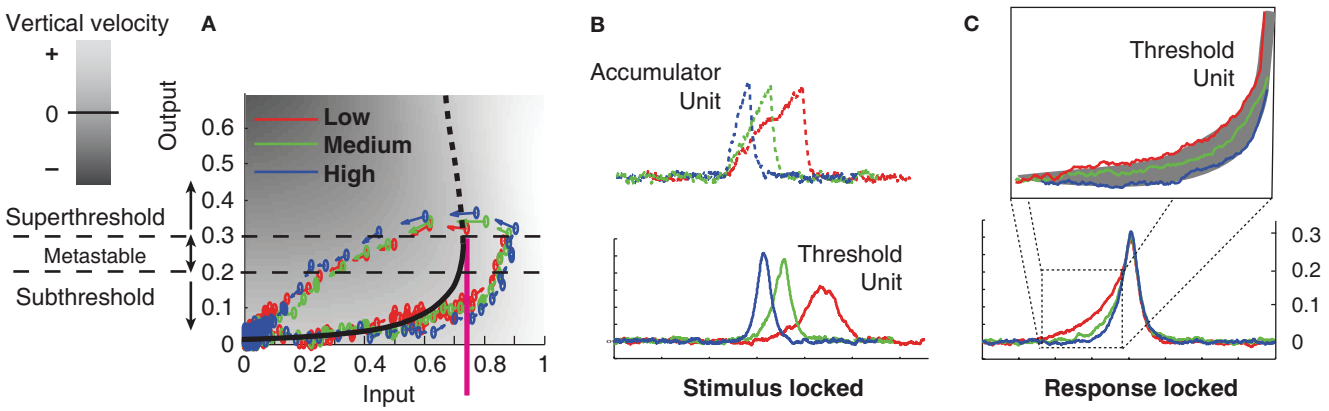


FIGURE 9 | Dynamics of a threshold latch in response to three different signal strengths, with non-zero noise. Blue: strong signal; green: medium signal; red: weak signal. **(A)** Input-output evolution of the threshold-unit's dynamics, starting at the origin and moving rightward, and upward along the bottom branch of the folded, effective activation function. Note the failure of the system's output value to exceed 0.4 before the shutoff inhibition drives the system down and to the left well before the system's output reaches its stable attracting value just below 1. **(B)** Stimulus-locked averages of activation (dashed: accumulator; solid: threshold-unit). **(C)** Response-locked averages of the threshold-unit activations, showing a common final level of activation, but differences in buildup rate for different signal qualities defining the gray area in the magnified inset (compare to **Figures 7** and **8**).

Of course, it is quite possible that accumulator and threshold mechanisms consist of networks of neurons distributed widely across the brain. In that case, it would be difficult to cleanly inactivate one component without affecting the other. If clean, independent inactivation could be achieved, however, then clear behavioral distinctions should

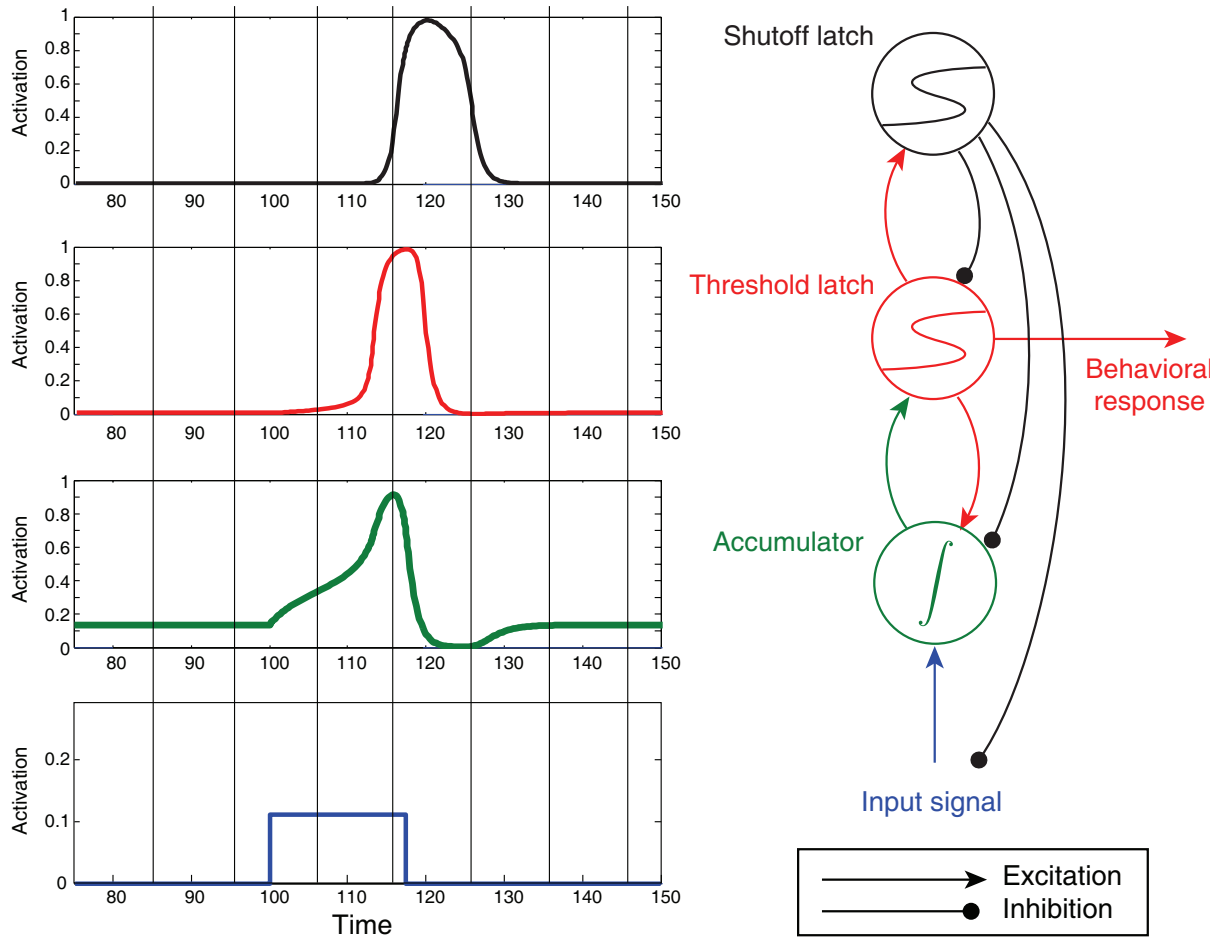


FIGURE 10 | Same circuit as in Figure 6, except that excitatory threshold-to-accumulator feedback is included. The green trace now has a late ballistic component, reflecting the ballistic activation of the threshold latch. Positive feedback also shortens the decision time considerably relative to the same model without threshold-to-accumulator feedback.

emerge: knocking out integrators should spare the ability to respond, while performance accuracy should approach chance; knocking out threshold mechanisms, in contrast, would abolish responding altogether. A more graded impairment could emerge from less potent inactivation: by inhibiting a threshold mechanism, more evidence would be required to accumulate to produce a decision, leading to improved accuracy, and increased response time; inhibiting an accumulator, in contrast, would presumably both increase response times and decrease accuracy.

Yet another possibility is that ramping activity in an area during decision making might be more or less epiphenomenal – relating neither to evidence accumulation nor threshold readout, *per se*, but instead to some sort of performance monitoring, or even simply to spreading correlations of activity that play no functional role. In this case, the inactivation tests proposed here would fail to produce the intended effects, but this outcome would at least suggest that a given area is not functionally relevant to the decision-making task at hand.

3. CONCLUDING REMARKS

Here I have assembled a list of reasons to consider a two-layer neural model, much like others discussed in neuroscience and psychology that are either explicitly composed of two layers, or that combine an evidence-accumulation process with an idealized decision threshold (e.g., Usher and McClelland, 2001; Corrado et al., 2005; Diederich and Busemeyer, 2006; Lo and Wang, 2006; Boucher et al., 2007; Ratcliff and McKoon, 2008; Gao et al., 2011). Like those models, the model proposed here splits different functions across different layers, rather than lumping them into a single layer (e.g., Wang, 2002). It thereby sacrifices parsimony for potentially better, more rewarding performance.

Unlike most multi-layer models, however, a model previously proposed by my colleagues and me (Simen et al., 2006) sends continuous, additive biasing signals to control the second layer (the threshold layer) rather than the first (the accumulator layer). Other models that we have proposed (Simen and Cohen, 2009; McMillen et al., 2011) tune multiplicative weights applied to the accumulated evidence before it is fed into the threshold mechanism. Increasing these weights amounts

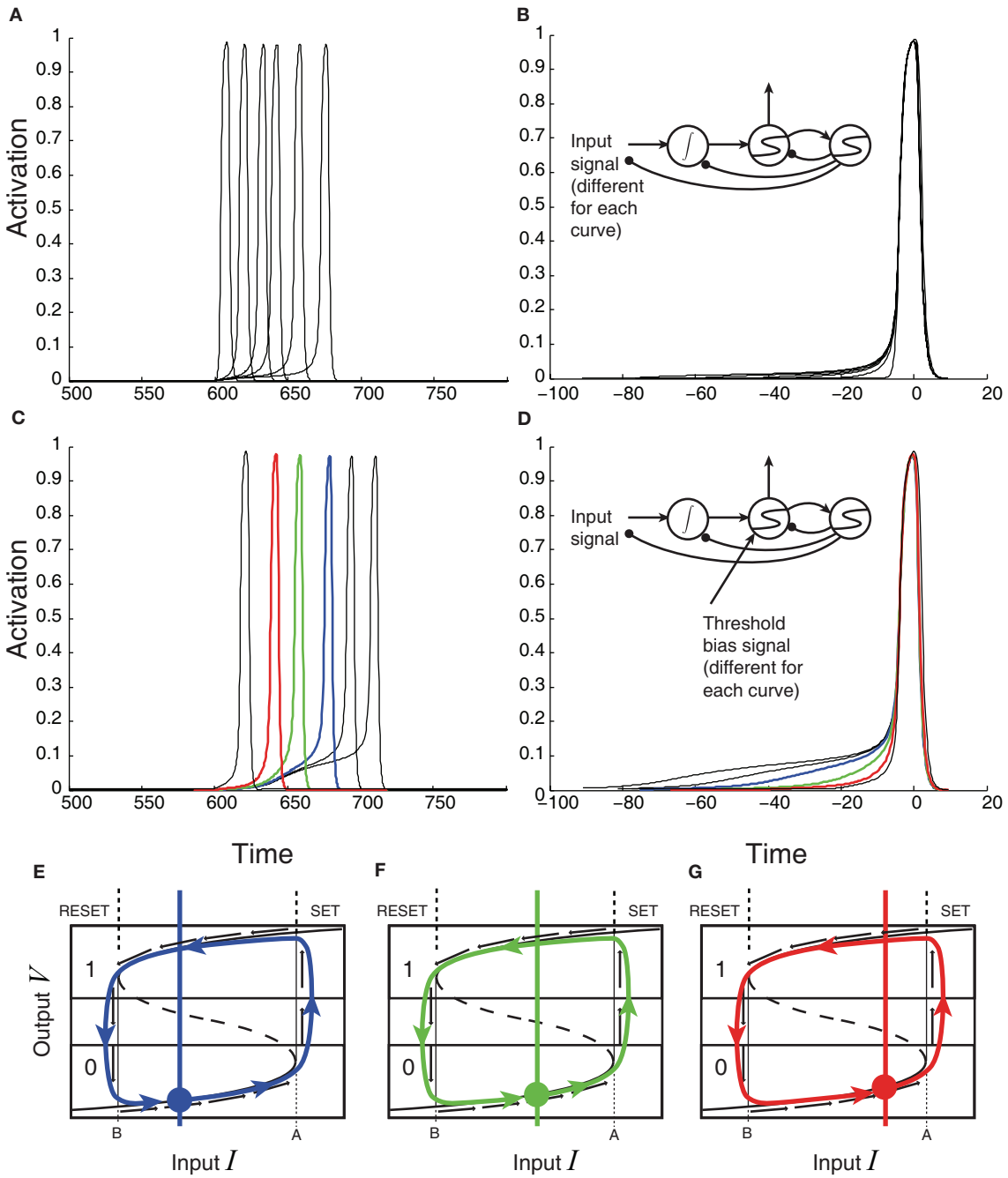


FIGURE 11 | Effects of biasing on the threshold-unit's activation. **(A,B)** Constant additive biasing of the *accumulator* can dramatically decrease response times for a fixed level of stimulus strength [offsets of stimulus-locked traces, **(A)**] and slightly modulates the initial rise period of the activation [slight changes in response-locked traces **(B)**]. **(C–G)** Constant additive biasing of the *threshold*-unit also affects stimulus-locked response times **(C)** and dramatically modulates the duration of the initial rise in response-locked averages **(D)**. **(E–G)** Three different levels of biasing (shifts of the initial input level denoted by the colored circle) with colors corresponding to the traces in **(C)** and **(D)** blue = low bias; green = medium bias; red = high bias.

to reducing thresholds divisively. In certain tasks (e.g., Bogacz et al., 2006), such approaches are both approximately optimal and mechanistically feasible.

Optimal biasing of accumulators rather than thresholds, in contrast, requires the biasing signals to be punctate rather than continuous, and I have raised doubts here about the physiological plausibility of punctate signals. These doubts are premised on the idea that any switch-like process in the brain must have a non-negligible rise-time and thus a non-zero, minimal duration – a duration that might plausibly take up a substantial proportion of

a typical response time in perceptual decision making. Human and non-human primate behavior is frequently suboptimal, however, so considerations of optimality do not rule out accumulator-biasing. Behavioral data, nonetheless often appear to support the notion that thresholds are strategically controlled, sometimes optimally (Simen et al., 2009; Bogacz et al., 2010a; Balci et al., 2011; Starns and Ratcliff, 2012). Furthermore, Ferrera et al. (2009) observed FEF activity that is consistent with a role for FEF as a tunable threshold mechanism, and (Ding and Gold, 2012) found a multitude of different functions expressed in FEF after previously finding much the same in the caudate nucleus (Ding and Gold, 2010). An inelegant but likely hypothesis supported by these and similar findings is that the separate functions of decision-making models are implemented by neural populations that are themselves distributed widely across the brain. What seems unquestionable, in any case, is that bistable switch mechanisms in either a one- or two-layer model would appear to play a necessary role in any full account of the data reviewed here.

It is also noteworthy that adaptive properties of the individual neuron's inter-spike time behavior have been explained by a form of threshold adaptation that is analogous to the population level threshold adaptation proposed here. As with neural population models, different models of the individual neuron can frequently mimic each other, whether they adapt action potential thresholds (Kobayashi et al., 2009) or the resting membrane potential after firing (Brette and Gerstner, 2005). In an international competition to model spike-time data, however, a relatively simple model with adjustable action potential thresholds (Kobayashi et al., 2009) defeated all other modeling approaches (Gerstner and Naud, 2009). Consistent with this evidence for adjustable action potential thresholds, recent findings suggest that sophisticated signal processing can occur at the axon's initial segment (Kole and Stuart, 2012). The same principles of accumulation, bistable readout, and threshold adaptation may therefore play out at multiple levels of neural organization.

I have argued that population threshold mechanisms are sufficiently non-ideal in their physical implementation that they should often be modeled explicitly rather than abstractly. What are the risks of getting this modeling choice wrong? Erring on the side of abstractness and simplicity risks:

1. misplacing the locus of evidence accumulation in the brain;
2. amplifying a disconnect between psychology and neuroscience in terms of which model-parameter or neural mechanism is modulated when speed-accuracy tradeoffs are adapted;
3. missing the possibility that cortical switch mechanisms might allow the cerebral cortex to implement a complex, sequential system (see Simen and Polk, 2010) without always requiring the involvement of structures such as the basal ganglia.

Conversely, insisting on an overly elaborate model of thresholds risks raising counterproductive doubts about neural data that is in fact tied to evidence accumulation. It also risks unnecessary inelegance and lack of parsimony. The optimal tradeoff between such risks is rarely obvious, and this article has not derived one. Its primary intent is to guard against the first kind of risk, since much of the decision neuroscience community currently seems safe from the second type.

A final, important point that should be kept in mind when neural evidence is brought to bear on psychological models (made elsewhere – e.g., Cohen et al., 2009 – but worth repeating) is that apparently slight changes in tasks may have dramatic consequences on the firing-rate patterns subserving performance of the tasks. Thus, although the conventional wisdom appears to be that there has been little electrophysiological evidence of threshold adaptation as a mechanism underlying behavioral performance adaptation, this lack probably depends heavily on the types of tasks that neuroscientists have examined. Fairly strong behavioral support (Simen et al., 2009; Bogacz et al., 2010a; Balci et al., 2011; Starns and Ratcliff, 2012) has been gathered for models of threshold adaptation in tasks for which such adaptation tends to maximize rewards (e.g., Bogacz et al., 2006). The tasks proposed in Bogacz et al. (2006) hold signal-to-noise ratios at a constant level across trials within any block of trials, whereas most physiological research with monkeys involves varying levels of signal-to-noise ratio from trial to trial. It therefore appears that valuable information could be gained about the neural mechanisms of economic influence on decision making if the exact task described in, for example, Bogacz et al. (2006), were tested directly in awake, behaving monkeys.

ACKNOWLEDGMENTS

Thanks to Michael Loose for comments on the manuscript.

REFERENCES

Balci, F., Simen, P., Niyogi, R., Saxe, A., Hughes, J., Holmes, P., and Cohen, J. D. (2011). Acquisition of decision making criteria: reward rate ultimately beats accuracy. *Atten. Percept. Psychophys.* 73, 640–657.

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., and Cohen, J. D. (2006). The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced choice tasks. *Psychol. Rev.* 113, 700–765.

Bogacz, R., Hu, P., Holmes, P., and Cohen, J. D. (2010a). Do humans select the speed-accuracy tradeoff that maximizes reward rate? *Q. J. Exp. Psychol.* 63, 863–891.

Bogacz, R., Wagenmakers, E.-J., Forstmann, B. U., and Nieuwenhuis, S. (2010b). The neural basis of the speed-accuracy tradeoff. *Trends Neurosci.* 33, 10–16.

Boucher, L., Palmeri, T. J., Logan, G. D., and Schall, J. D. (2007). Inhibitory control in mind and brain: an interactive race model of countermanding saccades. *Psychol. Rev.* 114, 376–397.

Brette, R., and Gerstner, W. (2005). Adaptive exponential integrate-and-fire model as an effective description of neuronal activity. *J. Neurophysiol.* 94, 3637–3642.

Brown, S., and Heathcote, A. (2005). A ballistic model of choice response time. *Psychol. Rev.* 112, 117–128.

Cisek, P., Puskas, G. A., and El-Murr, S. (2009). Decisions in changing conditions: the urgency-gating model. *J. Neurosci.* 29, 11560–11571.

Cohen, J. Y., Heitz, R. P., Woodman, G. F., and Schall, J. D. (2009). Letter to the editor: reply to Balan and Gottlieb. *J. Neurophysiol.* 102, 1342–1343.

Corrado, G. S., Sugrue, L. P., Seung, H. S., and Newsome, W. T. (2005). Linear-nonlinear-poisson models of primate choice dynamics. *J. Exp. Anal. Behav.* 84, 581–617.

Diederich, A., and Busemeyer, J. R. (2006). Modeling the effects of payoff on response bias in a perceptual discrimination task: bound-change, drift-rate-change, or two-stage-processing hypothesis. *Percept. Psychophys.* 68, 194–207.

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

Ding, L., and Gold, J. I. (2012). Neural correlates of perceptual decision making before, during, and after decision commitment in monkey frontal eye field. *Cereb. Cortex.* 22, 1052–1067.

Ferrera, V. P., Yanike, M., and Cassanello, C. (2009). Frontal eye field neurons signal changes in decision criteria. *Nat. Neurosci.* 12, 1454–1458.

Gao, J., Tortell, R., and McClelland, J. L. (2011). Dynamic integration of reward and stimulus information in perceptual decision making. *PLoS ONE* 6:e16479. doi:10.1371/journal.pone.0016749

Gardiner, C. W. (2004). *Handbook of Stochastic Methods*, 3rd Edn. New York, NY: Springer-Verlag.

Gerstein, G. L., and Mandelbrot, B. (1964). Random walk models for the spike activity of a single neuron. *Biophys. J.* 4, 41–68.

Gerstner, W., and Naud, R. (2009). How good are neuron models? *Science* 326, 379–380.

Gold, J. I., and Shadlen, M. N. (2002). Banburismus and the brain: decoding the relationship between sensory stimuli, decisions, and reward. *Neuron* 36, 299–308.

Green, D. M., and Swets, J. A. (1966). *Signal Detection Theory and Psychophysics.* New York: Wiley.

Hayes, J. (1993). *Introduction to Digital Logic Design.* New York, NJ: Addison-Wesley.

Kandel, E. R., Schwartz, J. H., and Jessell, T. M. (eds). (2000). *Principles of Neural Science.* Norwalk, CT: Appleton and Lange.

Kiani, R., Hanks, T., and Shadlen, M. N. (2008). Bounded integration in parietal cortex underlies decisions even when viewing duration is dictated by the environment. *J. Neurosci.* 28, 3017–3029.

Kobayashi, R., Tsubo, Y., and Shinomoto, S. (2009). Made-to-order spiking neuron model equipped with a multi-timescale adaptive threshold. *Front. Comput. Neurosci.* 3:9. doi:10.3389/neuro.10.009.2009

Kole, M. H. P., and Stuart, G. J. (2012). Signal processing in the axon initial segment. *Neuron* 73, 235–247.

Lo, C.-C., and Wang, X.-J. (2006). Cortico-basal ganglia circuit mechanism for a decision threshold in reaction time tasks. *Nat. Neurosci.* 9, 956–963.

Luce, R. D. (1986). *Response Times: Their Role in Inferring Elementary Mental Organization.* New York, NY: Oxford University Press.

McMillen, T., Simen, P., and Behseta, S. (2011). Hebbian learning in linear-nonlinear networks with tuning curves leads to near-optimal, multi-alternative decision making. *Neural Netw.* 24, 417–426.

Moore, T., and Armstrong, K. M. (2003). Selective gating of visual signals by microstimulation of frontal cortex. *Nature* 421, 370–373.

Moore, T., and Fallah, M. (2001). Control of eye movements and spatial attention. *Proc. Natl. Acad. Sci. U.S.A.* 98, 1273–1276.

Purcell, B. A., Heitz, R. P., Cohen, J. Y., Schall, J. D., Logan, G. D., and Palmeri, T. J. (2010). Neurally constrained modeling of perceptual decision making. *Psychol. Rev.* 117, 1113–1143.

Ratcliff, R., and McKoon, G. (2008). The diffusion decision model: theory and data for two-choice decision tasks. *Neural. Comput.* 20, 873–922.

Reddi, B. A. J., and Carpenter, R. H. S. (2000). The influence of urgency on decision time. *Nature* 3, 827–830.

Seung, H. S. (1996). How the brain keeps the eyes still. *Proc. Natl. Acad. Sci. U.S.A.* 93, 13339–13344.

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.

Simen, P., Balci, F., deSouza, L., Cohen, J. D., and Holmes, P. (2011a). Interval timing by long-range temporal integration. *Front. Integr. Neurosci.* 5:28. doi:10.3389/fnint.2011.00028.

Simen, P., Balci, F., deSouza, L., Cohen, J. D., and Holmes, P. (2011b). A model of interval timing by neural integration. *J. Neurosci.* 31, 9238–9253.

Simen, P., and Cohen, J. D. (2009). Explicit melioration by a neural diffusion model. *Brain Res.* 1299, 95–117.

Simen, P., Cohen, J. D., and Holmes, P. (2006). Rapid decision threshold modulation by reward rate in a neural network. *Neural Netw.* 19, 1013–1026.

Simen, P., Contreras, D., Buck, C., Hu, P., Holmes, P., and Cohen, J. D. (2009). Reward rate optimization in two-alternative decision making: empirical tests of theoretical predictions. *J. Exp. Psychol. Hum. Percept. Perform.* 35, 1865–1897.

Simen, P., and Polk, T. A. (2010). A symbolic/subsymbolic interface protocol for cognitive modeling. *Log. J. IGPL* 18, 705–761.

Smith, P. L. (2010). From Poisson shot noise to the integrated Ornstein-Uhlenbeck process: neurally principled models of information accumulation in decision-making and response time. *J. Math. Psychol.* 54, 266–283.

Starns, J., and Ratcliff, R. (2012). Age-related differences in diffusion model boundary optimality with both trial-limited and time-limited tasks. *Psychon. Bull. Rev.* 19, 139–145.

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

van Ravenzwaaij, D., van der Maas, H., and Wagenmakers, E.-J. (2011). Optimal decision making in neural inhibition models. *Psychol. Rev.* 119, 201–215.

Wang, X. J. (2002). Probabilistic decision making by slow reverberation in cortical circuits. *Neuron* 36, 955–968.

Conflict of Interest Statement: The author declares 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 February 2012; accepted: 21 May 2012; published online: 21 June 2012.

Citation: Simen P (2012) Evidence accumulator or decision threshold – which cortical mechanism are we observing? Front. Psychology ***3****:183. doi: 10.3389/fpsyg.2012.00183*

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

frontiers in
NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 17 July 2012
doi: 10.3389/fnins.2012.00106

EEG oscillations reveal neural correlates of evidence accumulation

M. K. van Vugt[1]*, P. Simen[2], L. E. Nystrom[3], P. Holmes[3,4] and J. D. Cohen[3]

[1] *Department of Artificial Intelligence, University of Groningen, Groningen, Netherlands*
[2] *Department of Neuroscience, Oberlin College, Oberlin, OH, USA*
[3] *Princeton Neuroscience Institute, Princeton University, Princeton, NJ, USA*
[4] *Department of Mechanical and Aerospace Engineering and Program in Applied and Computational Mathematics, Princeton University, Princeton, NJ, USA*

Edited by:
Marius Usher, Tel-Aviv University, Israel

Reviewed by:
Michael X. Cohen, University of Amsterdam, Netherlands
Tobias H. Donner, University of Amsterdam, Netherlands

****Correspondence:***
M. K. van Vugt, Department of Artificial Intelligence, University of Groningen, Nijenborgh 9, 9747 AG Groningen, Netherlands.
e-mail: m.k.van.vugt@rug.nl

Recent studies have begun to elucidate the neural correlates of evidence accumulation in perceptual decision making, but few of them have used a combined modeling-electrophysiological approach to studying evidence accumulation. We introduce a multivariate approach to EEG analysis with which we can perform a comprehensive search for the neural correlate of dynamics predicted by accumulator models. We show that the dynamics of evidence accumulation are most strongly correlated with ramping of oscillatory power in the 4–9 Hz theta band over the course of a trial, although it also correlates with oscillatory power in other frequency bands. The rate of power decrease in the theta band correlates with individual differences in the parameters of drift diffusion models fitted to individuals' behavioral data.

Keywords: EEG, drift diffusion model, decision making, oscillations

1. INTRODUCTION

Every day we make thousands of decisions, and modeling work has attempted to describe the nature of these decision processes (e.g., Ratcliff, 1978; Usher and McClelland, 2001). With the advent of cognitive neuroscience there has been a growing interest in its neural correlates. Here we introduce a novel approach to studying decision dynamics with human electrophysiology. By using model-predicted decision dynamics as regressors, we perform a comprehensive search for oscillatory features of electroencephalographic (EEG) activity that could reflect evidence accumulation.

There exist two main approaches to analyzing EEG data: looking at the raw potential, averaged time-locked to an event of interest ("event-related potentials") or looking at periodic activity or oscillations (not necessarily averaged). The presence of oscillations in EEG measurements indicates that neurons in a region have more synchronized synaptic and membrane activity (Wang, 2010). Through being synchronized, oscillations become strong enough in power to be detectable on the scalp. Synchronization is thought to allow groups of neurons to communicate with each other (Womelsdorf et al., 2007; Fries, 2009). Synchronized activity is crucial for plasticity and learning in the brain (STDP; Wang, 2010). The brain also appears to use oscillations in conjunction with spikes to encode specific information. Certain phases of oscillations often show an increased level of spiking relative to their baseline (Fries et al., 2007). A clear example of how the brain makes use of that oscillation-related change in excitability is phase coding, in which the phase of an oscillation at which a neuron fires encodes the spatial location of an animal (O'Keefe and Recce, 1993; Fries et al., 2007).

Without attempting to review the oscillations literature in full, we point out here a few relevant findings of this literature (see Buzsáki, 2006; Wang, 2010, for a more complete review). Probably the most-discussed oscillations are those in the 28–90 Hz gamma band, which have been studied extensively in the context of attention tasks. A prominent finding is that attention increases the amplitude of occipital 28–90 Hz gamma oscillations (e.g., Fries et al., 2007). Yet some studies have shown that oscillations of lower frequency are also important for attention and perception. For example Busch and VanRullen (2010) found that stimuli are better perceived at certain phases of the on-going 4–9 Hz theta oscillation than at other phases. This has led to the idea that so-called sustained attention is not uniformly sustained, but rather has an oscillating quality. Moreover, it suggests that in addition to hippocampal theta oscillations, which have primarily been associated with memory (e.g., Kahana et al., 2001) and spatial navigation (e.g., O'Keefe and Burgess, 1999), there exist cortical theta oscillations that are relevant to, among other things, perception.

In fact, it has also been suggested that cortical theta oscillations are crucial for the coordination of multiple sources of activity at decision points (Womelsdorf et al., 2010), and for combining various pieces of evidence (van Vugt et al., in press). Theta oscillations have also been found to covary with decision certainty (Jacobs et al., 2006) and prediction errors in decision making (Cavanagh et al., 2010). This suggests that theta oscillations could have a fundamental role in perceptual decision making and specifically in the accumulation of evidence. Nevertheless, other sources suggest that evidence accumulation is correlated with higher frequency oscillations in the beta and gamma bands (e.g., Donner et al., 2009).

The aim of this study is therefore to use a data-driven approach to find oscillatory correlates of evidence accumulation. To be able to do so, we need precise predictions for the dynamics of evidence accumulation provided by mathematical models. Probably

the most-discussed model for evidence accumulation is the Drift Diffusion Model (DDM; Ratcliff, 1978). This model posits that to make a decision, a person accumulates information until it reaches a threshold, at which time they make the response that corresponds to that threshold. Their response times (RTs) can be predicted by adding a fixed non-decision time to the time it takes to reach the threshold to account for sensory and motor latencies. The speed with which one accumulates evidence on average is referred to as the "drift rate" of the accumulation process. The height of the decision threshold reflects response caution. This model, and variants of it, is capable of explaining complete RT distributions, not just average RTs like most other models of cognition (Ratcliff and Smith, 2004).

In this study we examine what frequency bands of brain oscillations best reflect evidence accumulation as predicted by accumulator models. We also test whether the dynamics of the thus-selected oscillations covary with individual differences in DDM parameters estimated on the basis of participants' behavioral data. This work not only furthers our understanding of human decision making, but may eventually allow us to distinguish different implementations of the DDM that cannot be disentangled based on behavioral data alone (Ditterich, 2010).

2. MATERIALS AND METHODS

2.1. TASK

Participants performed a perceptual decision making task in which they judged the direction of motion (left or right) of a display of randomly moving dots of which a percentage moved to the left or the right. These random dot kinematograms were similar to those used in a series of psychophysical and decision making experiments with monkeys as participants (e.g., Britten et al., 1992; Shadlen and Newsome, 2001; Gold and Shadlen, 2003). Stimuli consisted of an aperture of ∼7.6 cm diameter viewed from ∼100 cm (∼4° visual angle) in which white dots (2 × 2 pixels) moved on a black background. A subset of dots moved coherently either to the left or to the right on each trial, whereas the remainder of dots jumped randomly from frame to frame. Motion coherence was defined as the percentage of coherently moving dots. Dot density was 17 dots/square degree, selected such that individual dots could not easily be tracked. Tracking was further discouraged by using three interleaved sets of dots of equal size, each of which was used in every three successive video frames. Therefore each set of dots returned only after three frames, with a random displacement. The speed of the dots was ∼7°/s. Participants indicated their responses by pressing the "Z" key with their left index finger for left-ward motion or the "M" key with their right index finger for right-ward motion.

Participants also performed a control task in which they did not need to integrate motion evidence (non-integration condition). In this condition, each trial started with entirely random (0% coherence) dot-motion, followed by an arrow indicating the direction to which a participant should respond. The arrow onset time was calibrated (based on RTs in previous blocks of the non-integration condition) such that the dot-motion-viewing times in these trials mirrored the response time distribution of the dots trials. This was done by taking the RT distributions from previous blocks, and subtracting from that the average time required for pressing a button in response to a stimulus ("signal detection trials").

The experiment presentation code was written in PsychToolbox (Brainard, 1997). Dots were presented with PsychToolbox extensions written by J. I. Gold[1].

2.2. PARTICIPANTS

Twenty-three participants (12 female; 21 right-handed; mean age 25; range 18–38) participated in our experiment in exchange for payment. The experiment was approved by the Institutional Review Board of Princeton University. Participants engaged in three separate hour-long training sessions in which they became familiar with the task. At the beginning of these training sessions, performance on a psychometric block (with fixed viewing times of 1000 ms and five different coherence levels) was used to determine the coherences at which they performed at ∼70 and 90% correct. These coherence levels were used for the remainder of the session, and the coherences from the last training session were used for the two EEG sessions.

2.3. RECORDING METHODS

We recorded EEG data from 128 channels using Neuroscan EEG caps with a Sensorium EPA-6 amplifier. Data were digitized at 1000 Hz and band-pass filtered from 0.02 to 300 Hz; all impedances $<30\,k\Omega$. Acquisition was controlled by Cogniscan software. All data were referenced to the left mastoid and off-line rereferenced to an average reference after automatic bad-channel removal (Friederici et al., 2000; Hestvik et al., 2007). We wavelet-transformed the data in six standard frequency bands (delta (2–4 Hz), theta (4–9 Hz), alpha (9–14 Hz), beta (14–28 Hz), low gamma (28–48 Hz), and high gamma (48–90 Hz); van Vugt et al., 2010) using six-cycle Morlet wavelets. Morlet wavelets have an optimal trade-off between temporal- and frequency resolution for EEG data (van Vugt et al., 2007) and six cycles are often used for the analysis of EEG data (e.g., Caplan et al., 2001).

2.4. GENERAL LINEAR MODEL FOR EEG

To find neural correlates of the dynamics of decision making in EEG data, we developed a General Linear Model (GLM) method, in which we correlated the predicted accumulator dynamics with the EEG time series. GLMs are a type of regression that is generally used with functional magnetic resonance imaging (fMRI) to find voxels that display a hypothesized pattern of activation, such as "high" when an object is presented and "low" when a scrambled object is presented. Here we used a similar technique with EEG data, to detect electrodes that display a pattern of activation that is predicted by the DDM.

For every trial, we modeled a ramp of activity starting at stimulus onset and ending at the response. This "upramp" had a height of one, and its slope was constrained by the response time for that trial (see **Figure 1** for examples). We compared the correlations of the ramp regressors to those of regressors reflecting the alternative hypotheses of neural activity that is "on" during the trial ("boxcar") and of neural activity that reflects a transient initial

[1]http://code.google.com/p/dotsx/

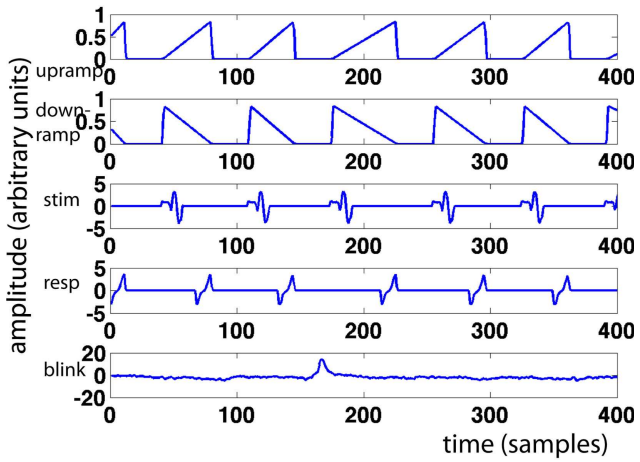


FIGURE 1 | Sample regressors. Pictured are, from top to bottom: upramp, downramp, stimulus regressor, response regressor, and eyeblink regressor. *X*-axis represents time in samples, *y*-axis regressor amplitude in arbitrary units.

response slowly decreasing over the trial duration ("downramp"). Evidence accumulation activity extracted with the upramp regressor should look clearly different from these alternative hypotheses. We did not employ separate regressors for left- and right-ward dot-motion. In other work (van Vugt et al., in revision), we have similarly detected lateralized readiness potentials, but these depart from baseline much later than the theta band activity discussed below, and we believe they are primarily motor (response) related.

We created a set of parallel upramps, downramps, and boxcars for the arrow control task, which had the same height as the dots ramps (i.e., a unit height), and also started at dots onset and ended at dots offset. Note, however, that in that case the arrows appeared somewhere in the middle of the interval between dots onset and offset, and the response followed fairly promptly thereafter.

We compared the fits of these regressors to an alternative model that did not take trial-by-trial variation in response time (RT) into account. In that model, we created regressors with a slope and height modulated by the individual's drift rate and threshold, respectively. In addition, the slope did not depart from baseline until T_0 (non-decision time) milliseconds after stimulus onset. We created one such regressor for the low-coherence condition, and a second for the high-coherence condition. We inserted in each trial the respective regressor shape. Importantly, these regressors did not covary with individual trial RTs (see **Figure 2**). As a result, the DDM model only has five parameters (low- and high-coherence drift rates, low- and high-coherence non-decision times, and decision thresholds), whereas the RT model has as many parameters as there are trials (namely, the RT for every trial). The DDM parameters were obtained by fitting the pure DDM (i.e., a DDM without variability in starting point, non-decision time, and drift rates) to each participant's behavioral data with the DMA toolbox (VandeKerckhove and Tuerlinckx, 2007, 2008).

In addition to these regressors of interest, we created a set of nuisance regressors that modeled transient neural responses to stimulus onset and button press, as well as eye activity. These nuisance regressors are used to remove those sources of variance from the EEG signal, such that only the signal of interest remained. To determine the transient response to a stimulus, we first looked at the grand average of stimulus-related EEG activity (**Figure 3A**), from which we chose a time window to plot a topographical distribution (**Figure 3B**). Although this topography does not exhibit a single clear maximum, we chose to use electrode Cz to compute for every participant individually the stimulus-related average. We then inserted this average waveform (from 0 to 300 ms post-stimulus) in the regressor at any time point where a stimulus was presented ("stimulus regressor").

Similarly, we examined the grand average response-locked ERP (**Figure 3C**) to define a time window for which to compute a topographical plot (**Figure 3D**) which was then used to define a response-related electrode. We chose CPz (Cz is slightly more anterior than CPz). We then used the average response-locked waveform from −200 to 0 ms in CPz to model the transient neural response to a button press ("response regressor"). See **Figure 3** for an illustration of the locations of Cz and CPz. Response-related ERP peaks typically exhibit their maximum more posteriorly than stimulus-related ERP peaks. CPz and Cz are two central electrodes that show peak responses to stimulus presentation and button presses, respectively. The eye blink regressor was created from the activity of the eye channel. We focused exclusively on eye blinks and not on horizontal eye movements because we only collected eye movements from a single eye electrode placed underneath the left eye. We set the eye blink regressor to zero outside the eye blink episodes detected with an amplitude threshold to ensure that no random fluctuations in the eye channel could distort our results.

A major problem in GLM analyses of EEG data is the poor signal-to-noise ratio (SNR). To improve the SNR we created features (independent variables in the regression) that only consisted of the trials themselves. This thus excluded the inter-trial time in which participants may have moved or been engaged in unspecified cognitive processes such as contemplating their lunch). We padded the trials with 300 ms before the stimulus and after the response. The reason for including this extra-trial time is that the neural activity of interest should display the hypothesized ramping behavior during the trial, but should also be relatively quiet outside this period, since at that time no evidence is being accumulated, and participants' attention will not yet have wandered very far away just after the trials. Moreover, not including this extra-trial period will make the upramp and downramp regressors identical up to an inversion, and this causes problems for the analysis. We can exploit the additional variance provided by this baseline to find our signal of interest in the EEG data. We then appended all these padded trials into a feature vector. The features were created both from the raw EEG time series, and wavelet-convolved signals in the delta (2–4 Hz), theta (4–9 Hz), alpha (9–14 Hz), beta (14–28 Hz), low gamma (28–48 Hz), and high gamma (48–90 Hz) ranges. After construction, we downsampled these features to 50 Hz, and *z*-transformed them to put them on the same scale across participants (van Vugt et al., 2010). Downsampling was done to reduce memory load for the computations.

We ran the GLM in two steps. In the first step we modeled all the nuisance regressors. The regressors of interest (ramps) were then modeled on the residuals of this first regression, which ensured that

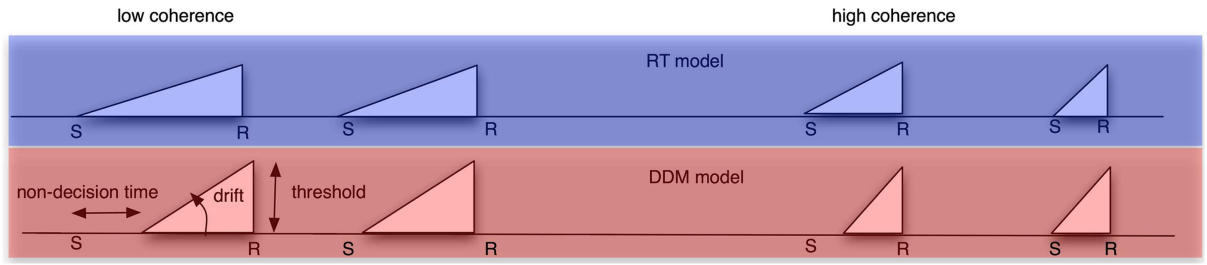


FIGURE 2 | Schematic of the different models we used to create regressors from. Top row: RT model. Bottom row: DDM model. Left column: low-coherence condition. Right column: high-coherence condition. In the RT model, the ramp always starts at stimulus onset and ends always at the response, and it always has a height of one. It therefore has a different length for the slower and faster trials within a coherence condition. Conversely, the DDM model has a fixed shape for all trials within the low-coherence condition, and another shape for all trials within the high-coherence condition. This shape is determined by three DDM parameters: non-decision time (which determines ramp onset), decision threshold (which determines ramp height) and drift rate (which determines the slope of the ramp).

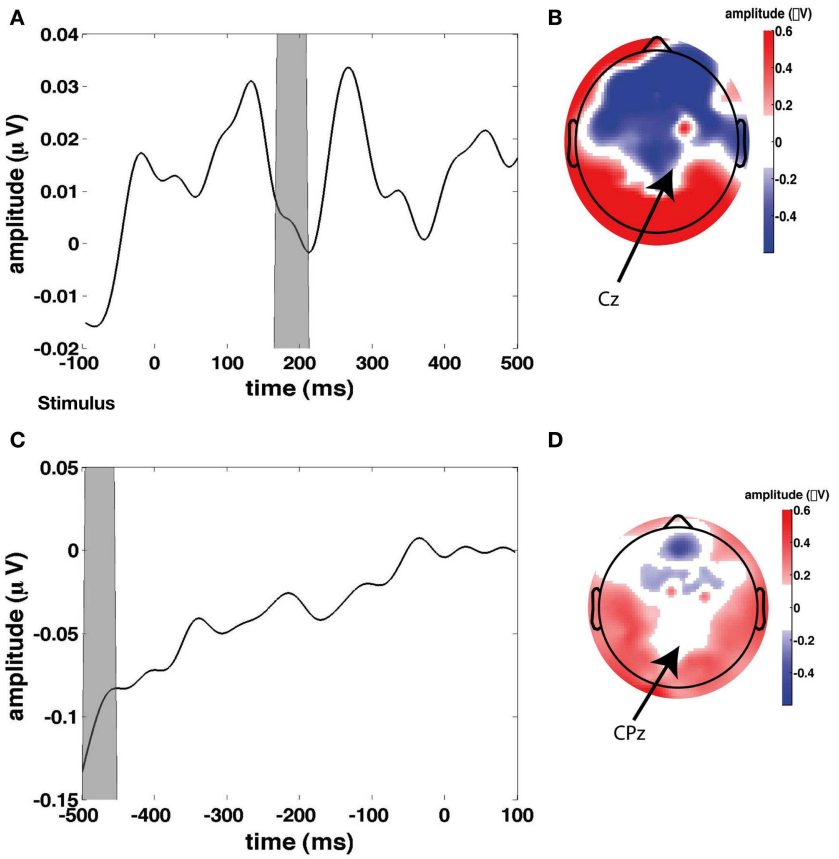


FIGURE 3 | Stimulus and response-locked event-related potentials (ERPs) used for creating the stimulus and response regressors. (A) Grand average stimulus-locked ERP (i.e., average across all channels). **(B)** Topographical plot of ERP amplitude in the gray time window in the time course in **(A)** which represents the first stimulus-evoked peak. **(C)** Grand average response-locked ERP. **(D)** Topographical plot of ERP amplitude in the gray time window in the time course in **(C)** which represents the maximum response-related peak.

the nuisance regressors could not influence the fits for the regressors of interest. In addition to computing the regression coefficients for each feature, we also computed the (square root of the) variance explained by correlating the feature with the fitted regressors as a measure of goodness of fit (Tabachnick and Fidell, 2005).

2.5. MULTIVARIATE GROUP ANALYSIS

To combine across participants, we included all participants' data into a single canonical correlation analysis (CCA; Calhoun et al., 2009). In general, CCA is a multivariate technique to find correlated components between two datasets. When given two matrices

(e.g., a set of regressors and a set of time courses from electrodes), it finds a set of weights on these two matrices such that they are maximally correlated. The CCA method we used (see **Figure 4**) was designed to do a group analysis across all of the participants.

There is no clear agreement in the literature on how to combine data across participants when using decomposition methods such as ICA and CCA (Calhoun et al., 2001). One approach is to perform the decomposition for each participant separately and then sort the resulting components. A problem with that method is that it is unclear how this sorting should be done reliably because the decomposition may have resulted in somewhat different components for every participant. Another approach is to concatenate all participants' data and to perform the decomposition on the resulting group data. The advantage of this method is that there is

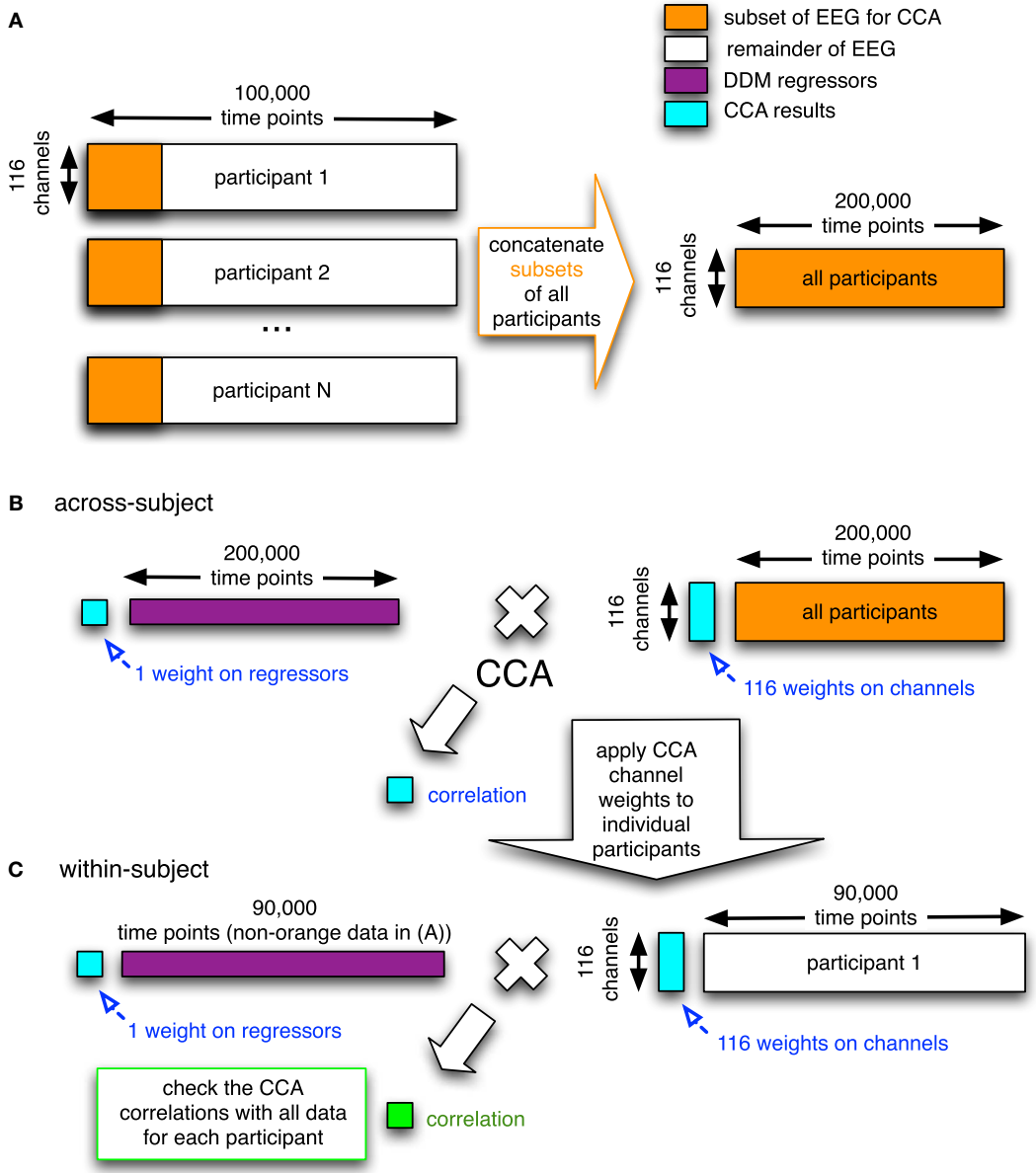


FIGURE 4 | Graphical overview of the CCA/GLM method we developed. In the first step **(A)**, the electrodes-by-time matrices are concatenated in the time dimension for all participants, where only a subset of the data of each participant is used (orange rectangles). This concatenated matrix with EEG data is then used together with the corresponding concatenation of regressors (purple; the DDM-inspired model time series) in the second step. **(B)** In this second step, the electrode-by-time matrix that contains data from all participants is correlated with the corresponding time course of the regressor (e.g., upramp) using CCA. This yields a correlation value, a weight on the regressor, and a set of weights on the electrodes (all in cyan). In the third step **(C)**, the weight map on the electrodes is applied to the remainder of every participants' EEG data (white) and the correlation (green) of this weighted EEG data to the regressor (purple) is compared to the correlation based on the group data in the random effects analysis (cyan). This whole procedure is done for the raw EEG data and separately for each frequency band (cf. **Figure 9** below). The number of time points indicated above the matrices are just an indication.

only a single set of weights on electrodes across all participants (i.e., a common spatial structure), while each participant has their own time course. It therefore becomes very easy to compare results across participants. A disadvantage of this method is that it forces components to have a similar spatial distribution across participants (Calhoun and Eichele, 2010). However, the added benefit of increased reliability derived from not having to aggregate potentially disparate components outweighed this limitation.

In this study we therefore decided to use temporal concatenation, such that we still had 116 features (corresponding to one feature for every EEG channel) but had hundreds of thousands of time points (see **Figure 4A**). The CCA then resulted in a single set of canonical correlation weights on electrodes across participants. We had sufficiently many time points to run into memory problems. To overcome these, we took a representative subset of each participants' data (i.e., six 4-min blocks from each of the different conditions). We checked that including only a subset of each participants' data in the group decomposition is a reasonable approach in **Figure 10** (see Results for more details). We computed a Bayes Factor to check whether indeed there was no strong evidence for a difference between the subset of the data and the full data by using the Bayes Factor calculator[2] (Rouder et al., 2009).

The CCA produced a single set of electrode weights across all participants for each regressor (**Figure 4B**). These electrode weights can take either sign, a negative correlation with an upramp implying that power decreases as time progresses.

2.6. RANDOM EFFECTS ANALYSIS

We developed a random effects analysis as a complementary means to assess the significance of the various correlations between regressors and EEG data. To this, we made use of the fact that we had only used a subset of each participants' data to run the CCA. We used the remaining data to compute for each participant separately a correlation between the EEG data to which the CCA-derived electrode weights were applied, and the corresponding regressors. We then Fisher-transformed these correlations and compared them with t-tests. As such, we could for example assess whether the correlation between EEG and the dots time course was larger than the correlation between EEG and the arrows time course. Because this involved many t-tests, we applied a False Discovery Rate procedure (Benjamini and Hochberg, 1995). A False Discovery Rate of 0.001, which is the level we used, indicates that on average 1 in 1000 of the significant effects we find is a false positive.

2.7. PERMUTATION ANALYSIS

Since it is possible that the results we obtained were due to random correlations between the EEG data and the regressors, we performed an additional permutation analysis to assess what the levels of canonical correlation would be for random data. We created random data by shuffling the ramps in time for random amounts and repeated the canonical correlation analysis with these regressors. We did this for 1000 iterations. We then compared the correlations obtained from the empirical data to those obtained from the random data, and computed the probability that our empirical data were derived from these random distributions.

3. RESULTS

3.1. BEHAVIORAL DATA

Before turning to the electrophysiological results, we discuss our behavioral data. Participants were engaged in a random dot-motion discrimination paradigm, where the level of motion coherence was set such that they performed at ∼70 and 90% correct (**Figure 5C**). Accuracy was significantly higher [$t(22) = 21.6$, $p < 0.001$, **Figure 5A**] and RT was significantly faster [$t(22) = 5.7$, $p < 0.001$, **Figure 5B**] in the 90% correct condition. The two coherence levels used to create the 70 and 90% correct conditions were also significantly different from each other [$t(22) = 17.3$, $p < 0.001$].

These results are consistent with a DDM parametrization in which thresholds are approximately constant, starting points are approximately midway between thresholds, and the drift rate is high for the high-coherence condition, and low for the low-coherence condition. This was confirmed by fitting a DDM to the behavioral data (**Table 1**). The mean (sem) Maximum Likelihood of these fits across subjects was 3466 (167) and the mean BIC (Bayesian Information Criterion) was 3508 (167). The fits showed that indeed the drift rate was higher for the high-coherence compared to the low-coherence condition. The drift rate was even higher for the arrows non-integration control condition, where evidence was so abundantly clear that participants barely needed to integrate information.

3.2. ELECTROPHYSIOLOGICAL DATA

Before turning to our novel model-based EEG analyses, we first examine the basic EEG data. We looked at all electrodes and picked a few representative samples of standard electrodes that are typically shown in EEG studies. **Figure 6** shows the basic characteristics of our EEG data. The spectrograms of oscillatory power for central electrode Cz show task modulation and a decrease in 4–9 Hz theta power over the course of the trial (**Figure 6A**). **Figure 6B** shows the effect of motion coherence (i.e., task difficulty), where trials in the easy high-coherence condition tend to show higher theta power near the response, compared to trials in the more difficult low-coherence condition. There also seems to be a difference in 14–28 Hz beta power occurring after the mean response time. This may reflect motor activity, which occurs later for the difficult compared to the easy trials, because the difficult trials have longer response times. **Figure 6C** shows the difference between the integration (dot-motion) and non-integration (arrows) conditions. Across the whole task, the non-integration condition is associated with higher 9–14 Hz alpha power than the integration condition. Furthermore, a decrease in theta is visible between stimulus and response, where the response is associated with lower theta.

Another notable effect that can be observed in the raw data is increased gamma oscillatory power for the integration condition, compared to the non-integration condition (shown for the frontal electrode FPz in **Figure 7**). This difference appears to be fairly constant across the whole task period. On the basis of these spectrograms, we expect evidence accumulation dynamics in the form

[2]http://pcl.missouri.edu/bf-one-sample

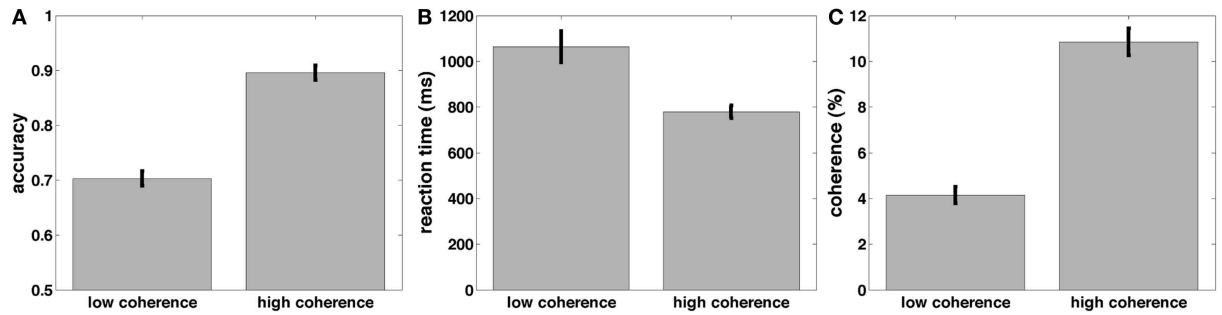


FIGURE 5 | Mean accuracy (A), response time (B), and coherences (C) across subjects for the low and high-coherence (difficult and easy) task conditions.

Table 1 | Mean (sem) DDM parameters for best fitting model to data from Experiment 1, separately for low and high-coherence trials (integration conditions), and arrows trials (non-integration condition).

Condition	Drift	Decision threshold	Non-decision time	Starting point
Low-coherence	0.057 (0.004)	0.151 (0.007)	0.439 (0.011)	0.077 (0.004)
High-coherence	0.167 (0.008)	0.151 (0.007)	0.408 (0.014)	0.078 (0.005)
Arrows control	0.784 (0.069)	0.210 (0.036)	0.219 (0.009)	0.094 (0.027)

The threshold was held constant between the low and high-coherence conditions. The fits adhere to the Ratcliff convention where the non-biased starting point is half the decision threshold value.

of ramps (**Figure 1**) to be primarily associated with low-frequency oscillations as shown in **Figure 6A**.

3.3. GLM METHOD CHECKS

Before running the CCA on the ramp regressors, we verified our method by plotting the weights for the eye blink, stimulus, and response regressors on a topographical plot. If the GLM method works correctly, we expect the highest regression coefficients on the electrodes that were used to generate the relevant regressor. That is, for the eyeblink regressor, we expect the largest weights on the frontal electrodes. For the stimulus regressor that was generated based on Cz, the maximum correlation should occur with this electrode, and the response regressor should maximally correlate with CPz. **Figure 8** shows that this was indeed the case.

Having established that the GLM, which predicts the time series of a single electrode by a linear combination of regressors, is a viable method to analyze EEG data, we used the multivariate CCA method to search for the hypothesized ramp dynamics in our data. The advantage of CCA is that rather than focusing on a single electrode at a time, it allows linear combinations of channels to predict linear combinations of regressors. It is therefore much more flexible. Yet, because it is more difficult to interpret linear combinations of regressors than single regressors, we restricted our attention to single regressors. Nevertheless, preliminary observations indicate that ∼40% larger canonical correlations can be obtained by allowing linear combinations of regressors. Linear combinations of regressors could for example create boxcars from a roughly equal weighting of upramp and downramp regressors, in a completely data-driven manner. Future work should further explore this option.

We performed a CCA between the regressors and the EEG time series for every channel. We did this analysis separately for every frequency band (2–4 Hz delta; 4–9 Hz theta; 9–14 Hz alpha; 14–28 Hz beta; 28–48 Hz low gamma; 48–90 Hz high gamma) because we sought to make inferences about which band shows most evidence of ramping activity.

Figure 9A shows the canonical correlation of the upramp with EEG data in every frequency band, as well as for non-wavelet-transformed (plain) EEG. The correlation between hypothesized ramping dynamics and EEG activity is largest in the 2–4 Hz delta and 4–9 Hz theta bands. Note that all correlations given by CCA are constrained to be positive, and that any negative relations between regressors and EEG will be reflected in a negative sign of the weights on the electrodes. We did a random effects analysis as a follow-up, in which we applied the CCA-derived electrode weights to each participants' left-out EEG data (only a subset of each participants' data was used for CCA due to memory constraints) and computed the correlation between these data and the regressors. We then assessed whether the Fisher-transformed individual-subject correlation values were significantly different from zero. We found that this was the case for all frequencies, except for the high gamma band [all t-values up to low gamma >6, which reflects p-values smaller than 0.0005, which is the p-value level corresponding to a False Discovery Rate of 0.001].

We also assessed whether the different frequency bands reflected evidence accumulation or rather a more general dot-motion-induced ramping process by repeating the same analysis for the non-integration condition. **Figure 9B** shows that the canonical correlate in the theta band is specific to the integration (dots) condition, whereas the canonical correlate in the delta

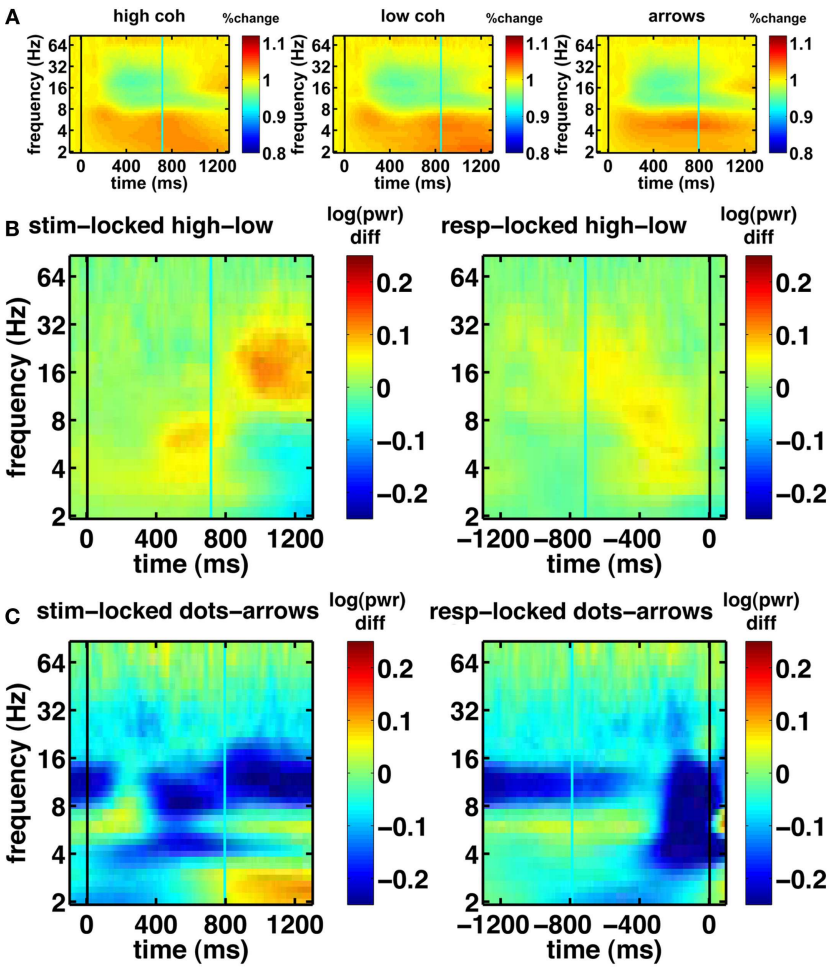


FIGURE 6 | Basic spectrograms for electrode Cz (a representative electrode). Black line indicates the event to which the data are aligned (onset of dot-motion for the stimulus-locked graphs and the response for the response-locked graphs). Cyan line indicates the average response time and stimulus onset time, respectively. **(A)** stimulus-locked spectrograms for the high-coherence, low-coherence, and non-integration condition. There is a gradual decrease in oscillatory power over the course of the trial. **(B)** Difference spectrograms comparing low- and high-coherence conditions. Left column: stimulus-locked. Right column: response-locked. **(C)** Difference spectrograms indicating the contrast between integration and non-integration conditions. Left column: stimulus-locked. Right column: response-locked.

band is also fairly high for the non-integration (arrows) condition. Indeed, comparing the distributions of individual-participant correlation values between the dots and arrows conditions for the delta band indicates that those are not significantly different when using a False Discovery Rate of 0.001 [$t(22) = 2.98$, $p = 0.0069$]. This means that in the non-integration condition, delta oscillatory power increases (or decreases-depending on the sign of the canonical correlation weights) over the course of a trial from dots onset until response, while according to our theory the participant only starts to accumulate evidence (and does so quite rapidly) once the arrow stimulus appears on the screen at a later point.

Theta oscillations show a dramatic drop in (canonical) correlation value in the non-integration condition, as would be expected from a neural correlate of evidence accumulation, because in the non-integration condition virtually no evidence has to be accumulated, and evidence accumulation only starts when the arrows appear on the screen. Indeed, there is a significant difference between the correlations of theta power with the upramp for dots versus arrows conditions [$t(22) = 11.6$, $p < 0.001$]. This suggests that theta oscillations are a more likely candidate for a neural correlate of evidence accumulation than delta oscillations because only the theta oscillations are specific to the integration condition.

We further tested whether the correlations we obtained could be due to a better match between temporal structure of the EEG data in the delta and theta frequency bands and the structure of the regressor than with EEG data in other frequency bands. In other words, we tested the alternative hypothesis that any random sequence of ramps would produce the correlations we observed, with the highest correlations in the delta and theta bands. To that end, we performed a permutation test. We created a set of 1000 regressors in which we randomly moved the ramps around across time, and redid the CCA. The white dots in **Figure 9** indicate the 97.5th percentile of the distribution of correlations that would

be expected based on random regressors. These random correlations are clearly below the observed correlations. Moreover, all of the observed canonical correlations based on the empirical data are larger than the canonical correlations based on random data, so the probability that the empirical canonical correlations are obtained from random data is <0.001.

We further asked whether our CCA, which was based on only a subset of each participant's data (see Methods), is a good representation of the participant's data. To this end we compared the correlation value of the canonical correlate of interest with the correlations between the regressor and the weighted set of electrodes for the entire time course within each subject (i.e., using the complete data for each subject). If the CCA decomposition were the same for every participant, then the canonical correlation of the across-subject data would be identical to the within-subject correlation based on the same weights. **Figure 10** shows that the within-subject correlations based on a subject's complete data, weighted by the coefficients obtained from the CCA, have generally similar values to the across-subject CCA based on a subset of a subject's data [one-sample t-test comparing mean proportion between the within and across-subject correlations to one: $t(22) = 2.0$, $p < 0.1$, Bayes factor for the alternative hypothesis of a mean correlation different from one versus the null hypothesis of a mean correlation equal to one: 1.02, indicating there is also little evidence for the null hypothesis].

We then wondered to what extent EEG activity would show correlation with alternative patterns of activity. Rather than upramps, there could also be neural correlates of downward ramps, which start at a high level and then return to baseline by the time of the response. Note that such a downramp is clearly different from a negative correlation of EEG with an upramp: the downramp starts at a high level and then drops down to baseline by the end of the trial, while the negative upramp starts at baseline and then goes down to a level of −1 at the end of the trial. The inclusion of between-trial baseline data allows us to distinguish between negative weights on upramps and (positive weights on) downramps. Finally, there could be patterns of neural activity that turn on at the start of the trial, and turn off at the end that reflects task engagement. We modeled this with a boxcar between stimulus onset and response.

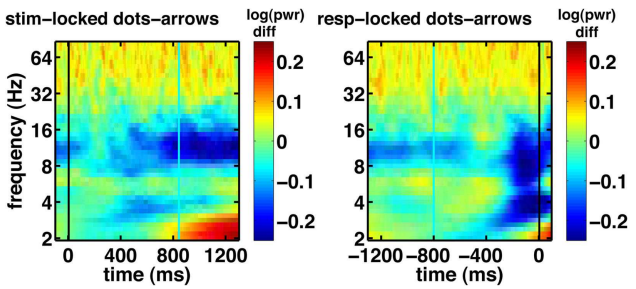


FIGURE 7 | Basic electrophysiological data for electrode FPz. Black line indicates the event to which the data are aligned (onset of dot-motion for the stimulus-locked graphs and the response for the response-locked graphs). Cyan line indicates the average response time and stimulus onset time, respectively. These are difference spectrograms contrasting the integration and non-integration conditions. Left column: stimulus-locked. Right column: response-locked.

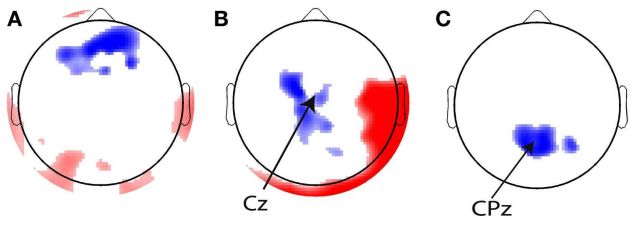


FIGURE 8 | Topographical overviews of the 10% most significant electrodes for the (A) eye blink, (B) stimulus, and (C) response regressors. Red and blue reflect positive and negative regression weights, respectively, and the intensity of the shading indicates the magnitude of the regression weights. The most significant electrodes are in the locations from which the regressors were generated [indicated with arrows in **(B,C)**].

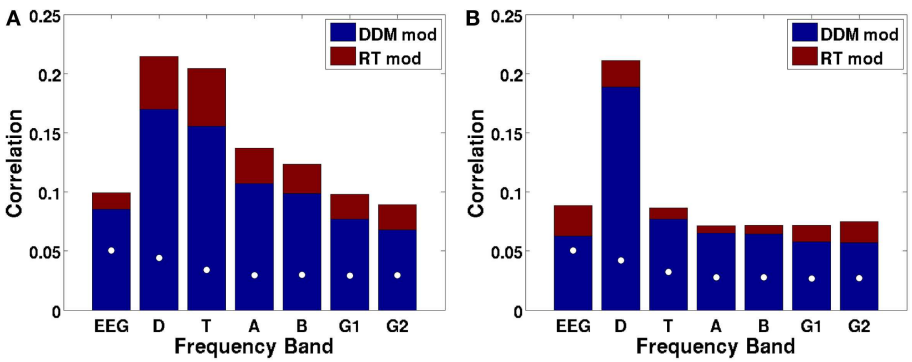


FIGURE 9 | Canonical correlations as a function of frequency, with the DDM-modulated model in blue and additional correlation achieved by the RT-modulated model in red, shown separately for the ramp regressor of (A) dots (integration condition) and (B) arrows (non-integration condition). White dots indicate the 97.5th percentile of the distribution of canonical correlations expected based on random regressors. Letters indicate frequency bands: EEG = raw EEG, D = 2–4 Hz delta, T = 4–9 Hz theta, A = 9–14 Hz alpha, B = 14–28 Hz beta, G1 = 28–48 Hz low gamma, and G2 = 48–90 Hz high gamma.

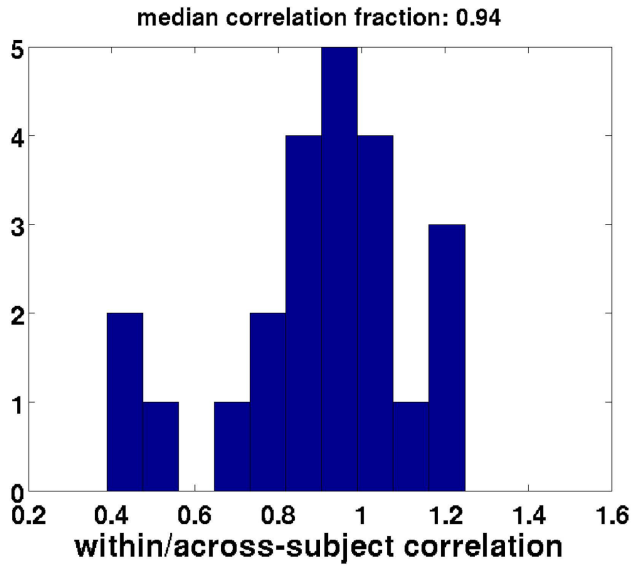


FIGURE 10 | Validation of the subset method: within-subject correlation of weighted regressors with EEG data divided by the across-subject canonical correlates. Perfect validity of the subset method would yield a fraction of one (within-subject correlations equal to across-subject canonical correlation). Each datapoint used to create this histogram reflects a single participant. This distribution has a mean that is not different from one.

Our results show that potential accumulators (upramps) are more consistent with the EEG data than these alternative hypothesis. Both downramp and boxcar regressors showed much lower correlations [in the theta band the canonical correlations are 0.11 for the boxcar, 0.076 for downramp dots, and 0.15 for downramp arrows]. This was corroborated by a random effects analysis, which showed that for all but the high gamma band, the upramp had a significantly higher correlation than the downramp [all $ts > 6.04$, $p < 0.001$].

We then asked to what extent the DDM, free from trial-to-trial variations in RT, could predict our EEG data. To do this, we compared the canonical correlations for a regressor that was ramping up or down exactly in accord with RT to that of a regressor that was more stereotyped, having a fixed length (time-locked to the response) but modulated by an individual's DDM parameters as obtained from fitting the DDM to the participant's behavioral data. Regressor height was modulated by the threshold parameter; its slope by the drift parameter and ramp onset was delayed by the participants' non-decision time (see **Figure 2**). Because the DDM-modulated regressor is not yoked to RT, it fails to capture the stochastic noise in RT. Although, as would be expected, the canonical correlations are uniformly higher for the RT-based regressor than for the DDM-modulated regressor, it is remarkable that the DDM still explains a large fraction (0.58–0.73) of the variance that the RT-yoked regressor can (**Figure 9**, red boxes labeled RT mod). In other words, the model is able to account for a large portion of the neural variance in ramp-like behavior.

Figures 11A,B show the time courses of the canonical correlate in the theta band: the frequency band that shows the greatest difference in upramp weights between dots and arrows (**Figure 9**). The time course of the upramp regressor is much more peaked for the integration condition (green) than for the non-integration condition (arrow trials; magenta). In the stimulus-locked average, the dots upramp time course departs significantly from baseline around 240 ms [one-tailed t-test with a $p < 0.01$ significance level], whereas the arrows upramp time course does not depart significantly from baseline until 400 ms post-stimulus. From 260 ms post-stimulus, the dots upramp and arrows upramp are significantly different. Similarly, in the response-locked time courses, the dots upramp differs from baseline from −580 ms, but the arrows upramp not until −380 ms. The two time courses start to differ significantly at −520 ms. The smaller amplitude of the arrows upramp is what we expected based on the lower correlations of theta activity with the upramp regressor in the non-integration compared to the integration condition. **Figure 11C** shows the topographical distribution of the weights on the electrodes that define the canonical correlate in the theta band. They have a posterior parietal distribution and are negative. This means that theta power in these parietal channels starts near baseline, and then as the trial progresses, theta power decreases away from baseline.

We next asked whether the dynamics of this theta upramp (the most promising candidate for evidence accumulation) also covaried with individual differences in DDM parameters. We found a significant correlation [$r(45) = 0.44$, $p = 0.0024$] between an individual's drift rate and the slope of the average theta band time course between 500 and 100 ms pre-response for that same person. There was no significant correlation between the level the time course reached at the end of the response interval and the individual's behaviorally fitted decision threshold [$r(45) = 0.08$, $p = 0.48$].

4. DISCUSSION

We have shown that EEG oscillations exhibit dynamics consistent with evidence accumulation in a perceptual decision making task. In addition, the magnitude of the slopes of these potential "neural accumulators" in the 4–9 Hz theta band covaried with individual differences in the drift rates obtained from the behavioral data.

While correlations with accumulator dynamics occur in different frequency bands, previous studies have implicated theta oscillations in various aspects of decision making (e.g., Jacobs et al., 2006; Cavanagh et al., 2010). For example Jacobs et al. (2006) showed an increase in parietal theta oscillations with decision confidence during a recognition memory task. Cavanagh et al. (2010) observed a correlation between post-decision error monitoring and theta oscillations in frontal regions. Our findings add to this body of evidence. One may wonder what is particular about the theta frequency that would make it suitable for a function in decision making. A modeling study by Smerieri et al. (2010) suggests an answer. They showed that in simulated spiking neural networks of two populations of mutually inhibiting neurons, RTs decreased and drift rates increased with increasing theta power. This effect was specific to the theta range because higher frequencies are too fast to modulate the cell's membrane potential, which acts as a low-pass filter.

Nevertheless, the correlations of accumulator dynamics with other frequency bands are also not too surprising. For example, a previous study has associated 14–28 Hz beta, rather than theta,

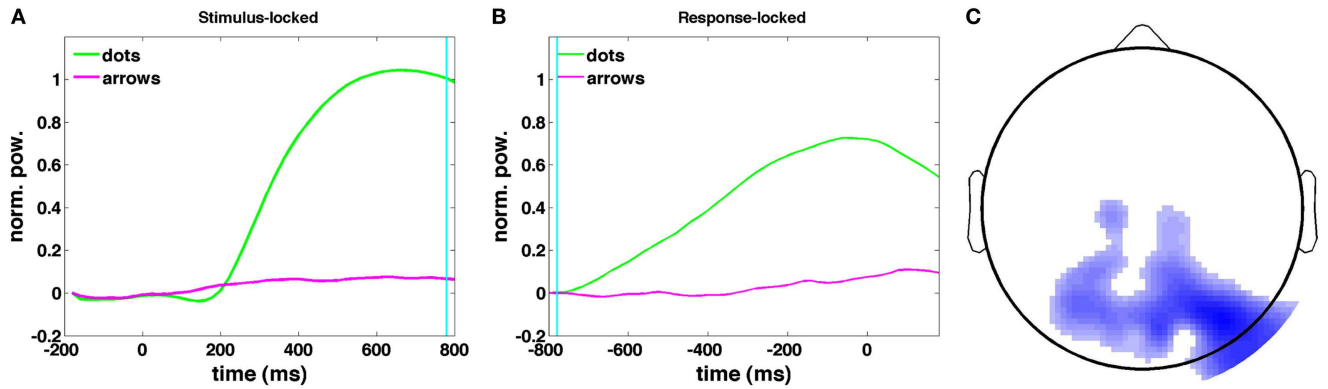


FIGURE 11 | Stimulus-locked (A) and response-locked (B) time courses of the canonical correlate in the 4–9 Hz theta band, correlated with the dots upramp, cyan line indicates RT in (A) and dots onset in (B). (C) Dots upramp topography (red indicates a positive correlation between oscillatory power and the regressor, blue a negative correlation). The shade of the color indicates the magnitude of the correlation. Note that there are no positive correlations for the dots upramp.

oscillations with evidence accumulation (Donner et al., 2009). One reason for that difference with our findings may be that that study was conducted with magnetoencephalography (MEG) rather than EEG recordings. MEG may be better able to detect higher frequency oscillatory activity because of its higher spatial sensitivity (e.g., Dominguez et al., 2005). Another difference with Donner's study is that the beta oscillations they observed were lateralized, that is, they increased in one hemisphere for one choice and in the other hemisphere for the other choice. While we have examined such lateralized oscillations as well with our model, this did not yield very strong correlations, because the lateralized beta and gamma oscillations only started to rise just before the response.

To further explore how analysis methods affect our results, we redid the same analysis on a Laplace-transformed version of our data, which is a method that improves spatial resolution for some sources, but decreases resolution for other sources (Hauk et al., 2002). This version of the analysis (available from the first author on request) still yielded correlations of EEG with evidence accumulation dynamics, but now the dots-selective accumulator dynamics shifted to more lateral locations and to the alpha and beta bands, quite similar to those observed by Donner et al. (2009). We should also note that our use of wavelets biases us to finding results in the lower frequencies, while multitapers would instead cause a bias toward finding results in higher frequencies (van Vugt et al., 2007). Together, this suggests that evidence accumulation involves oscillations at different frequencies, and the type of oscillations one observes most prominently depends on the recording and analysis methods.

The ramp regressor not only correlates with theta activity but also loads fairly strongly on EEG activity in the 2–4 Hz delta band. Nevertheless, unlike the theta band, the delta band shows a significant correlation with the ramp regressor in *both* the integration and the non-integration control task (**Figures 9A,B**). This suggests that while theta may be more specific to evidence accumulation, delta may reflect a more generic "on-task" process that is triggered by the dot stimuli on the screen. The relatively large loading on the alpha regressor may reflect bleeding-in of theta activity because there are individual differences in the ranges of alpha and theta oscillations (Klimesch et al., 1993). It may thus be the case that the 9–14 Hz alpha band contains theta accumulator activity from individuals with a high-frequency theta band.

It is also important to consider alternative explanations for the observed correlation with upramp dynamics. For example, the upramp could alternatively reflect time-on-task, which also increases with RT. Arguing against this interpretation is the fact that the correlate of the upramp in the theta band is much lower for the non-integration control task, in which stimulus-viewing-duration has not changed, but the need for evidence accumulation has mostly disappeared. Alternatively, upramp dynamics could reflect response preparation. While some accounts argue that motor preparation already starts at the time of stimulus presentation (Miller and Hackley, 1992), it mostly occurs a few hundred milliseconds before the response (see also **Figure 3**). A third alternative explanation is that pre-response increases in medial frontal theta increases have previously been associated with response conflict (Cohen and Cavanagh, 2011). The dots and arrows condition do not only differ in accumulation dynamics, but also in the amount of response conflict, and while the effect we observe has different sign and topography from the effects reported by Cohen and Cavanagh (2011), we cannot exclude the possibility that our results represent response conflict.

Another issue to consider is that while the correlation with the ramp regressor is stronger in the theta band than in other frequency bands, evidence accumulation is also a broad-band phenomenon: it is significantly different from zero in almost all frequency bands according to our random effects analysis. Several recent studies have suggested that broad-band increases in oscillatory power reflect increased neuronal spiking more than increases in power in specific frequency bands (Manning et al., 2009). Furthermore, changes in broad-band power have been associated with cognitive processes, such as verbal and spatial memory (Ekstrom et al., 2007; Sederberg et al., 2007).

Although the correlations we obtained between the regressors and the EEG data are on the order of magnitude of correlations obtained from GLMs applied to fMRI data, there is room for methodological improvement. Correlations of fMRI and EEG with

task conditions or cognitive models tend to be fairly low due to the large amount of noise in neural data. Nevertheless, we specifically showed that the correlations we obtained are larger than correlations found based on random regressors with a similar temporal structure (**Figure 9**). Future studies should investigate whether correlations could be improved by applying e.g., Independent Component Analysis (Delorme et al., 2001). In addition, the use of regularization, which zooms in on the informative features in the data, could potentially help to increase the correlation between model dynamics and EEG data.

What may seem surprising about the neural correlate of evidence accumulation in the theta band is that instead of increasing, oscillatory power decreases over the course of the decision interval (**Figure 11C**). Nevertheless, these decreases in oscillatory power may in fact reflect increases in functional brain activity. This is consistent with (Lorist et al., 2009), who found that oscillatory power increases with fatigue, thereby implying it should decrease with productive task performance. It may also be the case that over the course of evidence accumulation, one moves from a more global mode of processing, in which information is combined from a large number of neurons, to combining information from a much smaller set of neurons associated with less synchronization and lower oscillatory power (von Stein and Sarnthein, 2000). Both of these hypotheses could be tested with more localized neural recordings obtained from e.g., intracranial EEG. A third possibility is that theta may reflect the amount of uncertainty or an urgency to respond (Cisek et al., 2009), rather than the evidence accumulation process *per se*, and that other oscillations (e.g., beta which is more prominently observed in Laplace-transformed EEG and MEG data) may reflect evidence accumulation itself.

Our findings have several implications for future research. First, the correlates of the DDM that are observable in EEG can be used to assess the effect of task manipulations (such as speed-accuracy trade-off or reward rate) on accumulation dynamics. Second, there are large individual differences in decision making (e.g., Forstmann et al., 2010). EEG signatures of neural accumulators may allow us to distinguish different types of participants or strategies, given that individual differences in DDM parameters covaried with the slope of the neural accumulation signal. The "neural accumulators" could thereby soak up some portion of the noise in the model. These "neural accumulators" may also capture individual trial noise, such as attentional fluctuations, although that remains to be proven. Third, we could use the same multivariate methods to clarify the topographical location of possible neural accumulators with fMRI data, which has poorer temporal but better spatial resolution than EEG. Using identical methods for the analysis of EEG and fMRI data in the same task could thus provide new perspectives on data fusion.

Finally, it is important to consider what implications our results have for models of decision making. For example, the nonlinearity of the accumulator time courses suggests that evidence accumulation may better be described by a competitive evidence integration than by a linear ballistic accumulator (Brown and Heathcote, 2008). Yet, it is difficult to distinguish between the remaining accumulator models based solely on their dynamics in two-alternative forced choice tasks (Ditterich, 2010). In fact, we have tried different versions of our evidence accumulation model, such as a version where only onset of evidence accumulation changed rather than both onset of evidence accumulation and the slope. We found no appreciable change in our results. However, it is possible to distinguish between some models by employing brief pulses of strong evidence, as in Wong and Huk (2008) and Zhou et al. (2009).

In short, we have developed a novel method for detecting and examining the electrophysiological correlates of model dynamics. Using this method, we have provided evidence for a neural correlate of the dynamics of evidence accumulation in decision making measured in human EEG. Accumulation dynamics were captured best by 4–9 Hz theta oscillations in a set of superior parietal channels, and they covaried with individual differences in DDM parameters fitted to behavioral data.

ACKNOWLEDGMENTS

The authors gratefully acknowledge funding from the AFOSR Multi-University Research Initiative FA9550-07-1-0537 and Conte Center MH062196 grants. They also thank Valerie Karuzis, Jillian Brinberg, Peter Foster, Laura deSouza, and Brian Schilder for help with testing participants. We thank the reviewers for suggesting additional analyses and for comments that allowed us to improve this paper. This research was conducted in the absence of any financial or commercial relationships that could be construed as a potential conflict of interest.

REFERENCES

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Series B Stat. Methodol.* 57, 289–300.

Brainard, D. H. (1997). The psychophysics toolbox. *Spat. Vis.* 10, 443–446.

Britten, K. H., Shadlen, M. N., Newsome, W. T., and Movshon, J. A. (1992). The analysis of visual motion: a comparison of neuronal and psychophysical performance. *J. Neurosci.* 12, 4745–4765.

Brown, S. D., and Heathcote, A. (2008). The simplest complete model of choice reaction time: linear ballistic accumulation. *Cogn. Psychol.* 57, 153–178.

Busch, N. A., and VanRullen, R. (2010). Spontaneous EEG oscillations reveal periodic sampling of visual attention. *Proc. Natl. Acad. Sci. U.S.A.* 107, 16048–16053.

Buzsáki, G. (2006). *Rhythms of the Brain.* New York: Oxford University Press.

Calhoun, V. D., Adali, T., Pearlson, G. D., and Pekar, J. J. (2001). A method for making group inferences from functional MRI data using independent component analysis. *Hum. Brain Mapp.* 14, 140–151.

Calhoun, V. D., and Eichele, T. (2010). "Fusion of EEG and fMRI by Parallel Group ICA," in *Simultaneous EEG and fMRI*, eds M. Ullsperger and S. Debener (Oxford: Oxford University Press), 161–174.

Calhoun, V. D., Liu, J., and Adah, T. (2009). A review of group ICA for fMRI data and ICA for joint inference of imaging, genetic, and ERP data. *Neuroimage* 45, S163–S172.

Caplan, J. B., Madsen, J. R., Raghavachari, S., and Kahana, M. J. (2001). Distinct patterns of brain oscillations underlie two basic parameters of human maze learning. *J. Neurophysiol.* 86, 368–380.

Cavanagh, J. F., Frank, M. J., Klein, T. J., and Allen, J. J. B. (2010). Frontal theta links prediction errors to behavioral adaptation in reinforcement learning. *Neuroimage* 49, 3198–3209.

Cisek, P., Puskas, G. A., and El-Murr, S. (2009). Decisions in changing conditions: the urgency-gating model. *J. Neurosci.* 29, 11560–11571.

Cohen, M. X., and Cavanagh, J. F. (2011). Single-trial regression elucidates the role of prefrontal theta oscillations in response conflict. *Front. Psychol.* 2:30. doi:10.3389/fpsyg.2011.00030

Delorme, A., Makeig, S., and Sejnowski, T. (2001). "Automatic artifact rejection for EEG data using high-order statistics and independent component analysis," in *Proceedings of the Third International ICA Conference*, San Diego.

Ditterich, J. (2010). A comparison between mechanisms of multi-alternative perceptual decision making: ability to explain human behavior, predictions for neurophysiology, and relationship with decision theory. *Front. Neurosci.* 4:184. doi:10.3389/fnins.2010.00184

Dominguez, L. G., Wennberg, R. A., Gaetz, W., Cheyne, D., Snead, O. C., and Velazquez, J. L. P. (2005). Enhanced synchrony in epileptiform activity? Local versus distant phase synchronization in generalized seizures. *J. Neurosci.* 25, 8077–8084.

Donner, T., Siegel, M., Fries, P., and Engel, A. K. (2009). Buildup of choice-predictive activity in human motor cortex during perceptual decision making. *Curr. Biol.* 19, 1581–1585.

Ekstrom, A. D., Viskontas, I., Kahana, M. J., Jacobs, J., Upchurch, K., Bookheimer, S., and Fried, I. (2007). Contrasting roles of neural firing rate and local field potentials in human memory. *Hippocampus* 17, 606–617.

Forstmann, B. U., Anwander, A., Schäfer, 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.

Friederici, A. D., Wang, Y., Herrmann, C. S., Maess, B., and Oertel, U. (2000). Localization of early syntactic processes in frontal and temporal cortical areas: a magnetoencephalographic study. *Hum. Brain Mapp.* 11, 1–11.

Fries, P. (2009). Neuronal gamma-band synchronization as a fundamental process in cortical computation. *Annu. Rev. Neurosci.* 32, 209–224.

Fries, P., Nikolić, D., and Singer, W. (2007). The gamma cycle. *Trends Neurosci.* 30, 309–316.

Gold, J. I., and Shadlen, M. N. (2003). The influence of behavioral context on the representation of a perceptual decision in developing oculomotor commands. *J. Neurosci.* 23, 632–651.

Hauk, O., Keil, A., Elbert, T., and Müller, M. M. (2002). Comparison of data transformation procedures to enhance topographical accuracy in time-series analysis of the human EEG. *J. Neurosci. Methods* 113, 111–122.

Hestvik, A., Maxfield, N., Schwartz, R. G., and Shafer, V. (2007). Brain responses to filled gaps. *Brain Lang.* 100, 301–316.

Jacobs, J., Hwang, G., Curran, T., and Kahana, M. J. (2006). EEG oscillations and recognition memory: theta correlates of memory retrieval and decision making. *Neuroimage* 15, 978–987.

Kahana, M. J., Seelig, D., and Madsen, J. R. (2001). Theta returns. *Curr. Opin. Neurobiol.* 11, 739–744.

Klimesch, W., Schimke, H., and Pfurtscheller, G. (1993). Alpha frequency, cognitive load and memory performance. *Brain Topogr.* 5, 241–251.

Lorist, M. M., Bezdan, E., ten Caat, M., Span, M. M., Roerdink, J. B. T. M., and Maurits, N. M. (2009). The influence of mental fatigue and motivation on neural network dynamics; an EEG coherence study. *Brain Res.* 1270, 95–106.

Manning, J. R., Jacobs, J., Fried, I., and Kahana, M. (2009). Broadband shifts in LFP power spectra are correlated with single-neuron spiking in humans. *J. Neurosci.* 29, 13613–13620.

Miller, J., and Hackley, S. A. (1992). Electrophysiological evidence for temporal overlap among contingent mental processes. *J. Exp. Psychol. Gen.* 121, 195–209.

O'Keefe, J., and Burgess, N. (1999). Theta activity, virtual navigation and the human hippocampus. *Trends Cogn. Sci. (Regul. Ed.)* 3, 403–406.

O'Keefe, J., and Recce, M. L. (1993). Phase relationship between hippocampal place units and the EEG theta rhythm. *Hippocampus* 3, 317–330.

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

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

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.

Sederberg, P. B., Schulze-Bonhage, A., Madsen, J. R., Bromfield, E. B., Litt, B., Brandt, A., and Kahana, M. J. (2007). Gamma oscillations distinguish true from false memories. *Psychol. Sci.* 18, 927–932.

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.

Smerieri, A., Rolls, E. T., and Feng, J. (2010). Decision time, slow inhibition, and theta rhythm. *J. Neurosci.* 30, 14173–14181.

Tabachnick, B. G., and Fidell, L. S. (2005). *Using Multivariate Statistics.* Boston: Pearson Education.

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

van Vugt, M. K., Schulze-Bonhage, A., Litt, B., Brandt, A., and Kahana, M. J. (2010). Hippocampal gamma oscillations increase with working memory load. *J. Neurosci.* 30, 2694–2699.

van Vugt, M. K., Sederberg, P. B., and Kahana, M. J. (2007). Comparison of spectral analysis methods for characterizing brain oscillations. *J. Neurosci. Methods* 162, 49–63.

van Vugt, M. K., Sekuler, R., Wilson, H. R., and Kahana, M. J. (in press). Electrophysiological correlates of similarity-based interference in visual working memory. *J. Exp. Psychol. Gen.*

VandeKerckhove, J. A., and Tuerlinckx, F. (2007). Fitting the Ratcliff diffusion model to experimental data. *Psychon. Bull. Rev.* 14, 1011–1026.

VandeKerckhove, J. A., and Tuerlinckx, F. (2008). Diffusion model analysis with MATLAB:A DMAT primer. *Behav. Res. Methods* 40, 61–72.

von Stein, A., and Sarnthein, J. (2000). Different frequencies for different scales of cortical integration: from local gamma to long range alpha/theta synchronization. *Int. J. Psychophysiol.* 38, 301–313.

Wang, X. J. (2010). Neurophysiological and computational principles of cortical rhythms in cognition. *Physiol. Rev.* 90, 1195–1268.

Womelsdorf, T., Schoffelen, J., Oostenveld, R., Singer, W., Desimone, R., Engel, A., and Fries, P. (2007). Modulation of neuronal interactions through neuronal synchronization. *Science* 316, 1609.

Womelsdorf, T., Vinck, M., Leung, L. S., and Everling, S. (2010). Selective theta-synchronization of choice-relevant information subserves goal-directed behavior. *Front. Hum. Neurosci.* 4:210. doi:10.3389/fnhum.2010.00210

Wong, K.-F., and Huk, A. C. (2008). Temporal dynamics underlying perceptual decision making: insights from the interplay between an attractor model and parietal neurophysiology. *Front. Neurosci.* 2:2. doi:10.3389/neuro.01.028.2008

Zhou, X., Wong-Lin, K.-F., and Holmes, P. (2009). Time-varying perturbations can distinguish among integrate-to-threshold models for perceptual decision making in reaction time tasks. *Neural. Comput.* 21, 2336–2362.

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 February 2012; accepted: 23 June 2012; published online: 17 July 2012. Citation: van Vugt MK, Simen P, Nystrom LE, Holmes P and Cohen JD (2012) EEG oscillations reveal neural correlates of evidence accumulation. Front. Neurosci. ***6****:106. doi: 10.3389/fnins.2012.00106*

This article was submitted to Frontiers in Decision Neuroscience, a specialty of Frontiers in Neuroscience.

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 31 January 2012
doi: 10.3389/fpsyg.2011.00398

The nature of belief-directed exploratory choice in human decision-making

W. Bradley Knox[1†], A. Ross Otto[2†], Peter Stone[1] and Bradley C. Love[3]*

[1] *Department of Computer Science, University of Texas at Austin, Austin, TX, USA*
[2] *Department of Psychology, University of Texas at Austin, Austin, TX, USA*
[3] *Department of Cognitive, Perceptual and Brain Sciences, University College London, London, UK*

Edited by:
Erica Yu, University of Maryland, USA

Reviewed by:
Timothy Pleskac, Michigan State University, USA
Erica Yu, University of Maryland, USA

****Correspondence:***
Bradley C. Love, Department of Cognitive, Perceptual and Brain Sciences, University College London, 26 Bedford Way, Room 235, London WC1H 0AP, UK.
e-mail: bradley.c.love@gmail.com

[†]W. Bradley Knox and A. Ross Otto have contributed equally to this work.

In non-stationary environments, there is a conflict between exploiting currently favored options and gaining information by exploring lesser-known options that in the past have proven less rewarding. Optimal decision-making in such tasks requires considering future states of the environment (i.e., planning) and properly updating beliefs about the state of the environment after observing outcomes associated with choices. Optimal belief-updating is *reflective* in that beliefs can change without directly observing environmental change. For example, after 10 s elapse, one might correctly believe that a traffic light last observed to be red is now more likely to be green. To understand human decision-making when rewards associated with choice options change over time, we develop a variant of the classic "bandit" task that is both rich enough to encompass relevant phenomena and sufficiently tractable to allow for ideal actor analysis of sequential choice behavior. We evaluate whether people update beliefs about the state of environment in a *reflexive* (i.e., only in response to observed changes in reward structure) or reflective manner. In contrast to purely "random" accounts of exploratory behavior, model-based analyses of the subjects' choices and latencies indicate that people are reflective belief updaters. However, unlike the Ideal Actor model, our analyses indicate that people's choice behavior does not reflect consideration of future environmental states. Thus, although people update beliefs in a reflective manner consistent with the Ideal Actor, they do not engage in optimal long-term planning, but instead myopically choose the option on every trial that is believed to have the highest immediate payoff.

Keywords: decision making, reinforcement learning, Ideal Actor, Ideal Observer, POMDP, exploration, exploitation, planning

INTRODUCTION

Effective decision-making often requires a delicate balance of exploratory and exploitative behavior. For example, consider the problem of choosing where to dine out from a set of competing options. The quality of restaurants changes over time such that one cannot be certain which restaurant is currently best. In this non-stationary environment, one either chooses the best-experienced restaurant so far (i.e., exploit) or visits a restaurant that was inferior in the past but now may be superior (i.e., explore). The actions a diner should take in a series of choices is a non-trivial problem as optimal decision-making requires factoring in the uncertainty of the environment and the impact of the current action on one's future understanding of restaurant quality.

How humans and artificial agents balance and structure exploratory and exploitative actions is an important topic in reinforcement learning (RL) research (Sutton and Barto, 1998; Cohen et al., 2007; Lee et al., 2011). Exploring when one should exploit and, conversely, exploiting when one should explore both incur costs. For example, an actor who excessively exploits will fail to notice when another action becomes superior. Conversely, an actor who excessively explores incurs an opportunity cost by frequently forgoing the high payoff option.

In deciding whether to explore or exploit, an agent should consider its uncertainty about the environmental state. In the dining example above, the agent's decision to explore or exploit should depend on the volatility of the environment (e.g., the rate at which restaurant quality changes over time) and how recently the agent has explored options observed to be inferior in the past. For example, an agent should exploit when it has recently confirmed that alternative restaurants remain inferior and the environment is fairly stable (i.e., restaurant quality does not rapidly change). On the other hand, an agent should explore when alternatives have not been recently sampled and the environment is volatile. Between these two extremes lie a host of intermediate cases.

In this contribution, we examine how people update their belief states about the relative superiority of actions. In one view, *reflective* belief updates incorporate predictions of unobserved changes in the environment. For example, a reflective belief-updater would be more likely to believe that an inferior restaurant has improved as time passes since its last visit to the restaurant. In contrast, a *reflexive* model of choice is only informed by direct observations of rewards and, therefore, does not fully utilize environmental structure to update beliefs and guide actions. This distinction closely echoes contemporary dual-system frameworks of RL in

which a reflexive, computationally parsimonious model-free controller putatively competes for control of behavior with a reflective and model-based controller (Daw et al., 2005).

For reflexive models, exploratory choices are the result of a *purely* stochastic decision process. Thus, reflexive accounts do not predict sequential structure in humans' patterns of exploratory choice (cf. Otto et al., 2010). Perhaps because of their simplicity and unexamined intuitions about the "randomness" of exploratory behavior, reflexive approaches are commonly adopted to model human behavior (Daw et al., 2006; Worthy et al., 2007; Gureckis and Love, 2009; Pearson et al., 2009; Jepma and Nieuwenhuis, 2011). Reflexive approaches are also prominent in the design of artificial agents (Sutton and Barto, 1998).

THE LEAPFROG VARIANT OF THE CLASSIC BANDIT TASK

To understand how people balance exploration and exploitation given uncertainty about the state of the environment, we developed a variant of the commonly used *n*-armed bandit task, as the restaurant example given above can be formally described. In the classic *n*-armed bandit task, there are multiple actions (i.e., bandits) with unknown payoffs associated with them. Crucially, the payoffs at each time point are not explicitly revealed to decision-makers but instead must be determined by repeated sampling of actions. In restless bandit tasks, the actions' payoffs change over time, necessitating the aforementioned balancing of exploratory and exploitive actions.

Previous studies of exploratory choice have utilized *n*-armed bandit tasks in which the payoff distributions associated with the actions noisily drift over the course of decision-making (Daw et al., 2006; Pearson et al., 2009; Jepma and Nieuwenhuis, 2011). Although these tasks assess how people behave in changing environments, one major drawback of existing tasks is that there is no accompanying formal analysis of what decisions people should ideally make. In particular, existing work does not specify a statistically optimal process for updating estimates of action payoffs, and the action selection methods posited ignore the informational value of exploring. The failure of existing tasks to prescribe optimal choice behavior makes it difficult to assess how people differ from an optimal agent. In part, these shortcomings reflect that formulating ideal agents for existing tasks is an intractable problem (e.g., Daw et al., 2006).

In this contribution, we develop and use a novel laboratory task that is sufficiently constrained to allow for formal specification of the ideal agent. This formulation will be used to assess human behavior (e.g., Are people reflective or reflexive belief updaters? Do they act optimally given their beliefs?). In our Leapfrog task, the rewards for two possible actions continually alternate in their superiority, "leapfrogging" over each other. The underlying state of the environment – that is, which option currently has the higher payoff – is only partially observable to the decision-maker. This class of problems is referred to as a partially observable Markov decision process (POMDP) in the Artificial Intelligence literature (Kaelbling et al., 1998). Choosing the best-observed action ("exploiting") generally provides little information about the underlying state while exploration can resolve the underlying state but potentially incurs opportunity costs.

An example instantiation of the Leapfrog task is depicted in **Figure 1A**. In the Leapfrog task there are two actions, A and B, with different payoffs. The participants' task is to try to choose the higher payoff option as often as possible (this proportion, not total points, is the key metric). Option B's payoff is initially higher (at 20 points) than option A's payoff (10 points). On each trial, there is a fixed probability, which we refer to as *volatility*, that the inferior action increases its payoff by 20 points, "leapfrogging" the other option to give higher payoff. In summary, jumps are subject to three constraints: they occur at a fixed volatility unknown to the decision-maker, the two actions alternate in making jumps, and a jump always increases an action's point payoff by 20. Critically, instructions make clear to participants that they will be rewarded at the end of the experiment based on the proportion of "correct" choices (i.e., choices for which the option with the higher true payoff was chosen) as opposed total points earned. Jumps are not explicitly made known to the decision-maker, but rather must be inferred indirectly from observing choice payoffs.

Consistent with previous frameworks (Daw et al., 2006), we define exploitation as choosing the action with the highest observed payoff. A decision-maker must explore to detect when the alternative action has leapfrogged over the action presently being exploited. The Leapfrog task has only two actions, and all possible underlying environment states can be mapped to the number of unobserved jumps. For example, when there are no unobserved jumps, the exploitative option still yields the higher payoff. On the other hand, when there is one unobserved jump, the exploratory option has the higher payoff. Consequently, unlike previously studied bandit tasks, the prescription for optimal choice behavior in the Leapfrog task is tractable, though non-trivial. Our task also affords a straightforward manipulation of the rate at which decision-makers should explore. Namely, across conditions, we vary the volatility of the environment to examine whether subjects in low and high volatility conditions differ in their balance of exploratory and exploitative choice. Intuitively and in accord with the Ideal Actor, we expect that subjects should explore more frequently in more volatile environments.

MODELS EVALUATED

A number of model variants are evaluated to shed light on human choice behavior in the Leapfrog task. In addition to examining whether choice is better described by reflexive versus reflective strategies, the second main question we ask is whether people *plan ahead optimally*, taking the value of the information gained through exploration into account when acting, or *myopically* choose the action expected to receive the larger reward, regardless of the action's impact on later reward. We compare human data from the Leapfrog task to three models: a reflexive and myopic model we term the "Naïve RL" model, which expects payoffs (or rewards) to be as they were last experienced; a reflective and myopic model we call the "Belief model," which directly acts on the basis of its beliefs about current payoffs; and a model that plans optimally from reflective beliefs, the "Ideal Actor." For full algorithmic descriptions of the models we refer the reader to the Appendix.

Both reflective models employ an "Ideal Observer," which optimally updates beliefs based on past actions and observed payoffs.

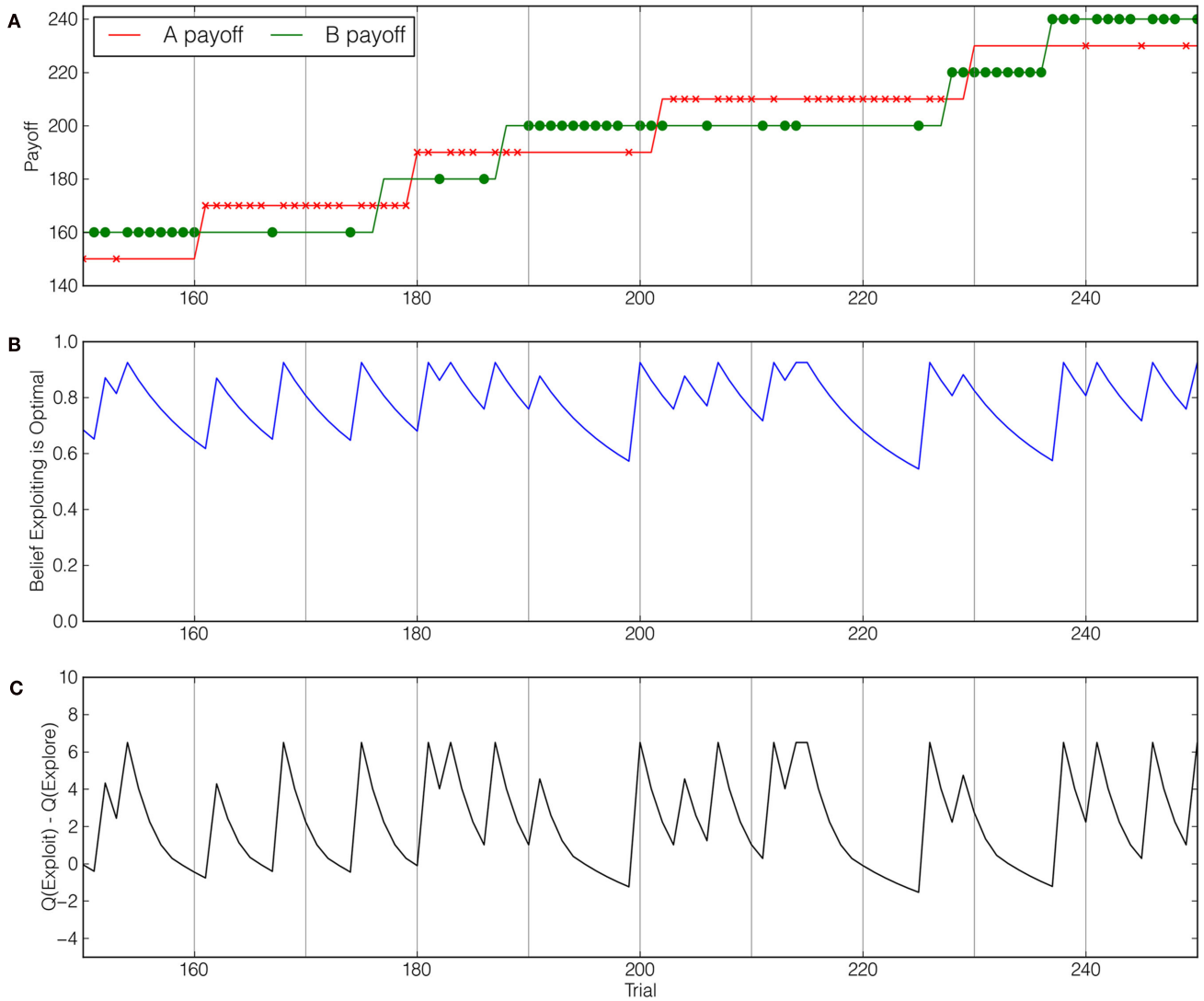


FIGURE 1 | Choice behavior and model-inferred beliefs and values in the Leapfrog bandit task. At each trial, there is a fixed probability, which we refer to as *volatility* or *P*(flip), that the inferior action increases its point payoff by 20 points. After the first "jump," Action A "pays" 30 points, superior to Action B's 20 points. After this jump, Action B's payoff will increase by 20 points with the same probability. **(A)** An example participant's sequence of choices over 100 trials in an environment with volatility rate *P*(flip) = 0.075. The two solid lines indicate the true payoffs for each option, and the × and • marks indicate the participant's choices among these options. **(B)** The Ideal Observer's belief that the exploitative options will yield larger the higher payoff – and thus the higher immediate reward – at each trial in the task instantiation at top. Note that the subject's certainty about the options' relative payoffs generally decreases during exploitation-only runs. **(C)** The relative long-term value of the exploitative option at each trial as determined by the Ideal Actor, using the beliefs from **(B)** and an optimal valuation function. Also note that changes in the relative value follow changes in belief in a non-linear fashion.

Figure 1B depicts beliefs as determined by the Ideal Observer for the actions and observations in **Figure 1A**. Similar Ideal Observer models have been employed in task domains such as visual search (Najemnik and Geisler, 2005) and prediction and change point detection (Steyvers and Brown, 2005). Of the two reflective models, only the Ideal Actor builds upon its optimal beliefs by considering the effect of exploration on reducing uncertainty in its future beliefs. Optimal beliefs about current payoffs and correct assessments of each action's informational value together yield a numeric expression of each action's overall value – expressed as *Q*-values – determining optimal choice behavior. The Ideal Actor's *Q*-values are calculated by converting the task to a POMDP and solving it in this form. In **Figure 1C**, the options' relative *Q*-values are shown as a function of the beliefs in **Figure 1B** and the number of remaining trials.

Because people may act noisily, in all three models we make choice a stochastic function of action values using the Softmax choice rule (Sutton and Barto, 1998), parameterizing the extent

to which the choice rule is sensitive to value differences using the inverse temperature parameter (henceforth the "Softmax parameter"). This constitutes the Naïve RL model's only free parameter, while the two reflective models have an additional parameter, $P(\text{flip})$, which represents the model's estimate of the environment volatility.

We rely on two complementary results to assess the belief-directed nature of subjects' choices in the Leapfrog task. First, we define a hazard rate metric elucidating the increasing likelihood of exploratory choice over time, for which reflective and reflexive models make clear and divergent predictions. Second, these qualitative results in turn motivate quantitative comparison of the extent to which these models characterize human choice. To foreshadow, we find that humans are best described by the reflective, but myopic, Belief model, suggesting that exploratory choice is not necessarily directed by a planning process that takes into account the value of future information yielded by actions. Finally, we analyze people's choice latencies in terms of the Ideal Actor's action prescriptions, observing that people exhibit larger latencies when they act suboptimally, demonstrating the Ideal Actor's potential as a tool for online, process-oriented analysis of exploratory choice behavior.

MATERIALS AND METHODS

SUBJECTS

A total of 139 undergraduates at the University of Texas participated in this experiment in exchange for course credit and a small cash bonus tied to proportion of trials for which the higher payoff option was chosen. The sample from which our sample was drawn is 54.3% female and 42.5% male, with 3.2% who declined to report their gender. The ages of participants in this pool ranged from 18 to 55 ($M = 19.08$, SD $= 1.76$). Participants were randomly assigned to three volatility level conditions, defined by the probability at each trial that the payoff ordering of options would flip, $P(\text{flip})$: low volatility [$P(\text{flip}) = 0.025$], medium volatility [$P(\text{flip}) = 0.075$], and high volatility [$P(\text{flip}) = 0.125$]. There were 51, 41, and 47 subjects in the low, middle, and high volatility conditions respectively.

MATERIALS AND PROCEDURE

The task instructions explained that one option was always worth 10 more points than the other option, that the superiority of the two options alternated over time, and that options always changed values by 20 points. Subjects were informed that their payment was tied to the number of times they chose the higher payoff option. Additionally, they were told at the outset which option, A or B, was initially superior at the start of the experiment and that the experiment consisted of 500 choices in total. The bandit task interface consisted of two buttons on a computer screen marked "OPTION A" and "OPTION B."

Prior to the main bandit task, subjects completed a number of training trials intended to familiarize them with the procedure and the volatility rate. In these training trials, participants first completed a passive viewing task in which they viewed 500 trials of the bandit task whose payoffs were randomly generated as previously described in the section on the Leapfrog task. To focus subjects' attention on the volatility rather than the true payoffs in the volatility-training phase, the payoffs for each option either read "SAME" or "CHANGED." Before each block of 100 trials, participants then provided an estimate of the number of flips they expected in the next block.

Following training, participants completed 500 trials of the main bandit task. On each trial, subjects saw the word "CHOOSE" and had 1.5 s to make a choice using the using the "Z" or "?" keys for the left and right options respectively. Following each choice, numerical feedback was provided for 1 s, indicating the number of points that resulted from the choice. When a response deadline was missed, the computer displayed the message "TOO SLOW" accompanied by a large red X for 1 s and the participant repeated that trial. Payoffs for options A and B started at 10 and 20 respectively and, as described above, alternated jumping by 20 points with probability governed by $P(\text{flip})$. An example instantiation of the payoffs is depicted in **Figure 1A** along with an example subjects' sequence of choices.

RESULTS

CHOICE BEHAVIOR

The primary dependent measure is whether subjects explored or exploited on a trial. We classified each choice made by a participant as either exploratory or exploitative based on their experienced payoffs up to that choice point: when the decision-maker chose the option with the highest-seen payoffs, that choice was considered an exploitative choice, and when they chose the other option, that was considered an exploratory choice (cf. Daw et al., 2006).

Figure 2A depicts the hazard rates of subjects' exploratory choice across the three conditions, calculated as the probability that an exploratory choice is made on trial t given that a payoff jump was observed on trial $t - n$, restricted to a five-trial window. In other words, this hazard rate is the probability of making an exploratory choice as a function of the number of consecutive exploitative choices. These hazard rates are calculated from 139 simulations – one for each subject in the experiment – of each model allocated across the three volatility conditions. Each model was "yoked" to a subject's particular instantiation of the Leapfrog payoff structure and, consequently, their environment volatility rate. To determine model choice behavior, we used the average of participants' best-fitting parameter values for each volatility condition and model (see **Table 1**; procedure described below). For each model, we calculated the hazard rate of exploration in the same way as subjects and report these rates in **Figures 2B–D**. It can be seen that subjects' rate of exploration increased monotonically over time, $F(1,137) = 5.96$, $p < 0.05$, contrasting with the predictions of the purely reflexive Naïve RL model but in accordance with the qualitative predictions of the reflective Belief and Ideal Actor models. Further, subjects in more volatile environments explored more frequently, $F(2,137) = 31.50$, $p < 0.001$, which is an intuitive result as beliefs about the relative expected payoffs of the options should change more rapidly in more volatile environments. There was a significant interaction between volatility and run length of exploitive trials, $F(2,137) = 4.47$, $p < 0.001$.

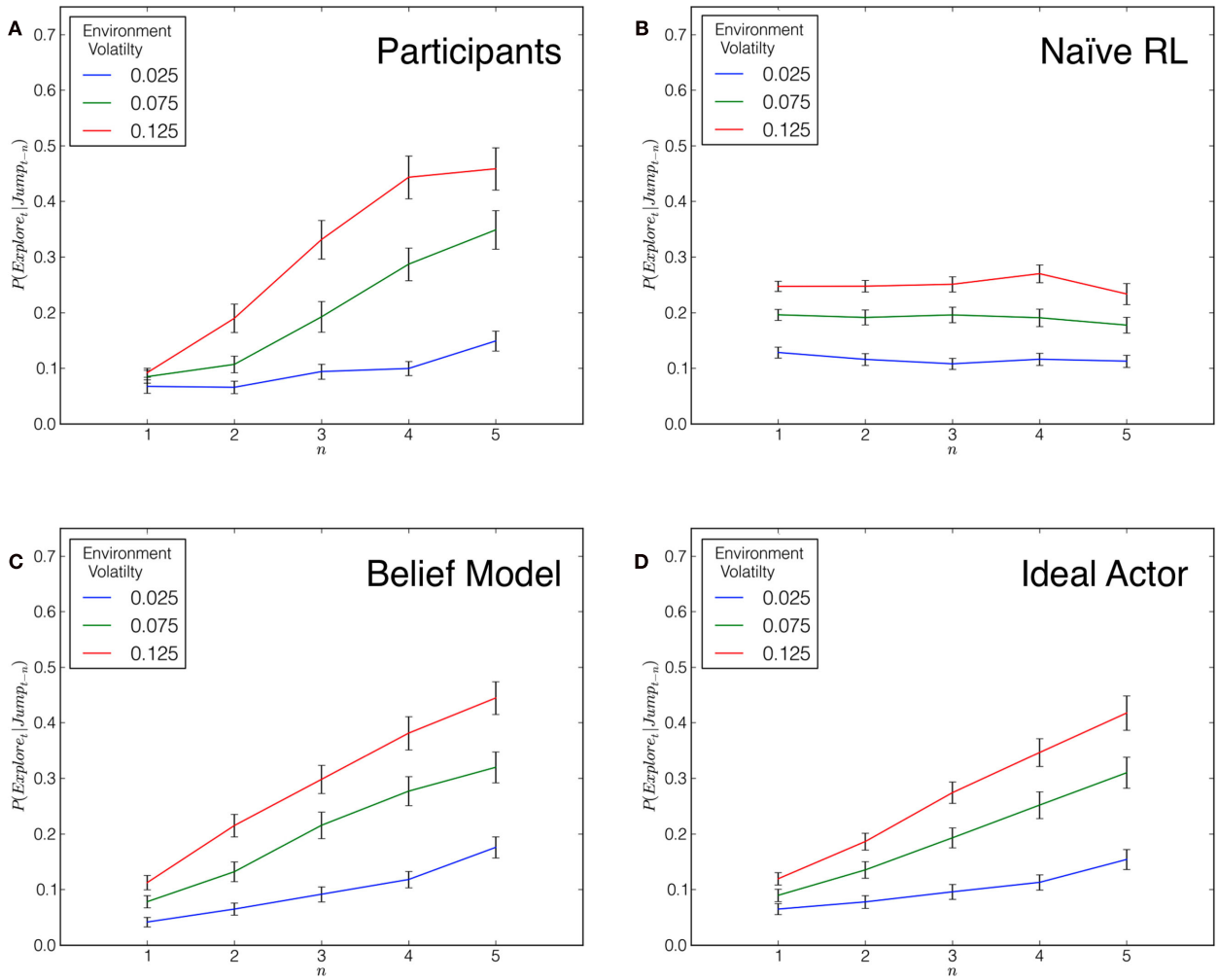


FIGURE 2 | "Hazard rate of exploration" for participants (A) and the three models under consideration (B–D), defined as the probability of exploring as a function of the number of consecutive exploitative choices after observing a payoff jump. Thus, for example, $n = 4$ refers to a situation where, beginning with the observation of a jump, three exploitative choices are made before an exploratory choice. Critically, the reflective nature of the Belief model and the Ideal Actor results in a monotonically increasing hazard rate as the models' certainty about the relative optimality of the options decays over time.

MODEL FITS

Having specified the three models computationally, we determined which model(s) best characterized participants' choices across the three volatility conditions. We used maximum likelihood estimation to find the set of parameters that maximized the likelihood of each model for each subject. To compare goodness of fit across models, we used the Bayesian information criterion (BIC: Schwartz, 1978) as the models have differing numbers of free parameters. Note that lower BIC values indicate better fit.

Across all three conditions, subjects were best fit by the Belief model as judged by log-likelihood scores (see **Figure 3A**). Though the Ideal Actor model fit worse overall (see **Figure 3B**), it provided the best fit for a considerable number of subjects. Very few subjects were best fit by the Naïve RL model. These results suggest that subjects' exploration manifests a reflective belief-updating rather than a reflexive process, but they do not appear to be optimally using these beliefs to conduct long-term planning.

CHOICE LATENCY

We also hypothesized that choice latencies (as measured by RTs) would provide an online assessment of a reflective and belief-driven decision process. We intuited that RTs would be larger in situations in which participants acted against their beliefs about the currently optimal action – that is, people would exhibit larger choice latencies when they made errors. Supporting this conjecture and following the data pattern of most studies of speeded choice in which response bias is minimal, leading models of choice predict that errors are associated with larger response times than are correct responses (cf. Ratcliff and Rouder, 1998). Accordingly, we factorially examined exploratory and exploitative choice RTs, classifying them as "explore optimal" or "exploit optimal," defining the two bins based on the Ideal Actor's choice prescription. To ensure that any effects of choice RT observed were not attributable to sequential effects such as response repetitions or switches (Walton et al., 2004) – which may be confounded with

Table 1 | Best-fitting parameter values by model and condition.

Condition	*P*(flip) (SD)	Softmax parameter (SD)	Total BIC
NAÏVE RL			
P(flip) = 0.025		1.9 (0.04)	16365
P(flip) = 0.075		1.4 (0.03)	16730
P(flip) = 0.125		1.1 (0.03)	21134
BELIEF MODEL			
P(flip) = 0.025	0.046 (0.033)	3.87 (1.40)	14757
P(flip) = 0.075	0.103 (0.059)	4.78 (2.02)	13668
P(flip) = 0.125	0.134 (0.069)	4.90 (2.28)	16807
IDEAL ACTOR			
P(flip) = 0.025	0.01 (0.01)	0.58 (0.23)	16535
P(flip) = 0.075	0.04 (0.05)	0.55 (0.23)	13995
P(flip) = 0.125	0.07 (0.07)	0.59 (0.25)	16914

exploratory choices – we first performed a regression to partial out these effects. This model assumed that choice RTs were a linear function of the response repetitions and switches (in relation to the present response) of the previous 10 trials. We then performed the analysis of interest on the resultant residual RTs. **Figure 4** depicts the average median reconstructed RTs across the three volatility conditions in the four unique situations described above.

It is apparent that, in the medium $[P(\text{flip}) = 0.075]$ and high $[P(\text{flip}) = 0.125]$ volatility conditions, participants exhibited larger choice latencies when they acted *against* the prescription of the Ideal Actor. A mixed-effects linear regression (Pinheiro and Bates, 2000) conducted on these residual RTs (random effects over subjects) revealed a significant interaction between chosen action (explore versus exploit, mirroring non-human primate results reported by Pearson et al., 2009) and prescribed optimal action (exploration-optimal versus exploitation-optimal), $\beta = -8.90$, SE $= 3.41$, $p < 0.01$. A full list of regression coefficients are provided in **Table 2**. It is important to note that these effects are prevalent even when explanations such as switch costs are taken into consideration. These patterns did not appear to vary significantly with volatility condition, $F = 0.18$, $p = 0.67$.

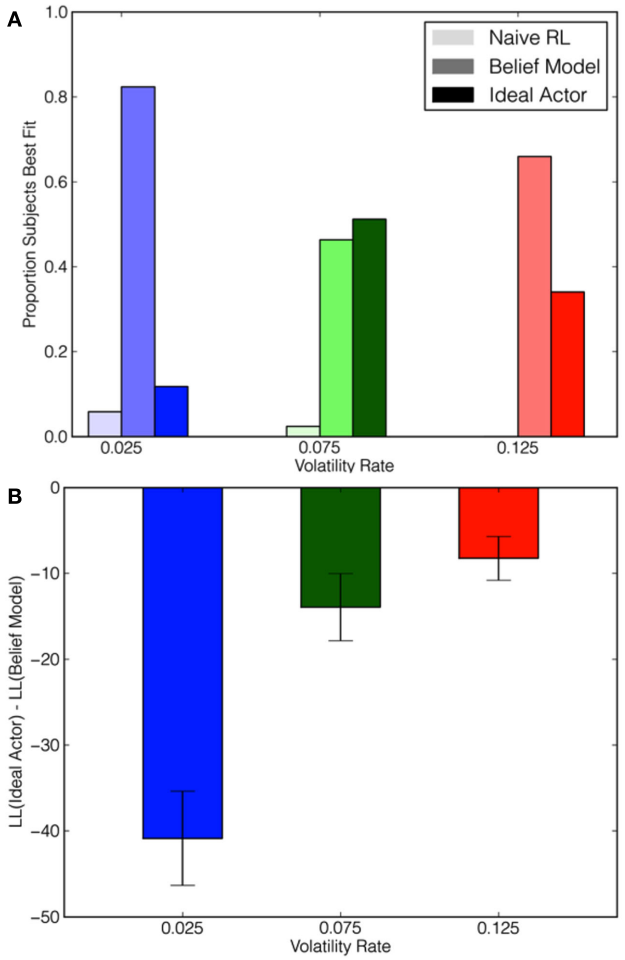


FIGURE 3 | Model goodness of fit measures as a function of model and volatility condition. (A) The proportion of participants that were best fit by each model in each condition. **(B)** Log-likelihood ratios comparing the Ideal Actor and the Belief model. Negative values indicate that participants were better described by the Belief model.

MODEL PERFORMANCE

To examine the importance of optimal planning (as opposed to myopic choice) in reflective belief models, we simulated deterministic versions of the Ideal Actor and the Belief model on 10,000 independent instantiations of the Leapfrog task. Rather than use a Softmax choice rule, these models act deterministically: the deterministic Ideal Actor always chooses the highest-valued action and the deterministic Belief model always exploits when the model's belief that the exploitative option yields higher payoff is greater than 0.5. Both models also employ optimal belief-updating by the Ideal Observer, using the condition's true *P*(flip) value. To examine how the addition of stochasticity might improve the Belief model's performance in this task, we simulated a Softmax variant of the Belief model using the true *P*(flip) values and Softmax parameters that were optimized to give the best performance. The results reported in **Table 3** confirm the importance of planning in the Leapfrog task and the benefit the Belief model derives from an element of stochastic (i.e., random) choice.

DISCUSSION

We examined whether human decision-makers approach exploratory choice in a reflective and belief-directed fashion as opposed to a stochastic and undirected fashion. Using a novel task that allowed for unambiguous identification of the two candidate strategies, we found that decision-makers appeared to be updating their beliefs about relative payoffs in a reflective manner – including knowledge about possible unseen changes in the task structure – but did not seem to be fully utilizing these beliefs by planning ahead with assessments of the informational value of actions. Indeed, for both subjects and reflective models, hazard rates reveal that the probability of exploratory choice increases with the number of immediately previous consecutive exploitive

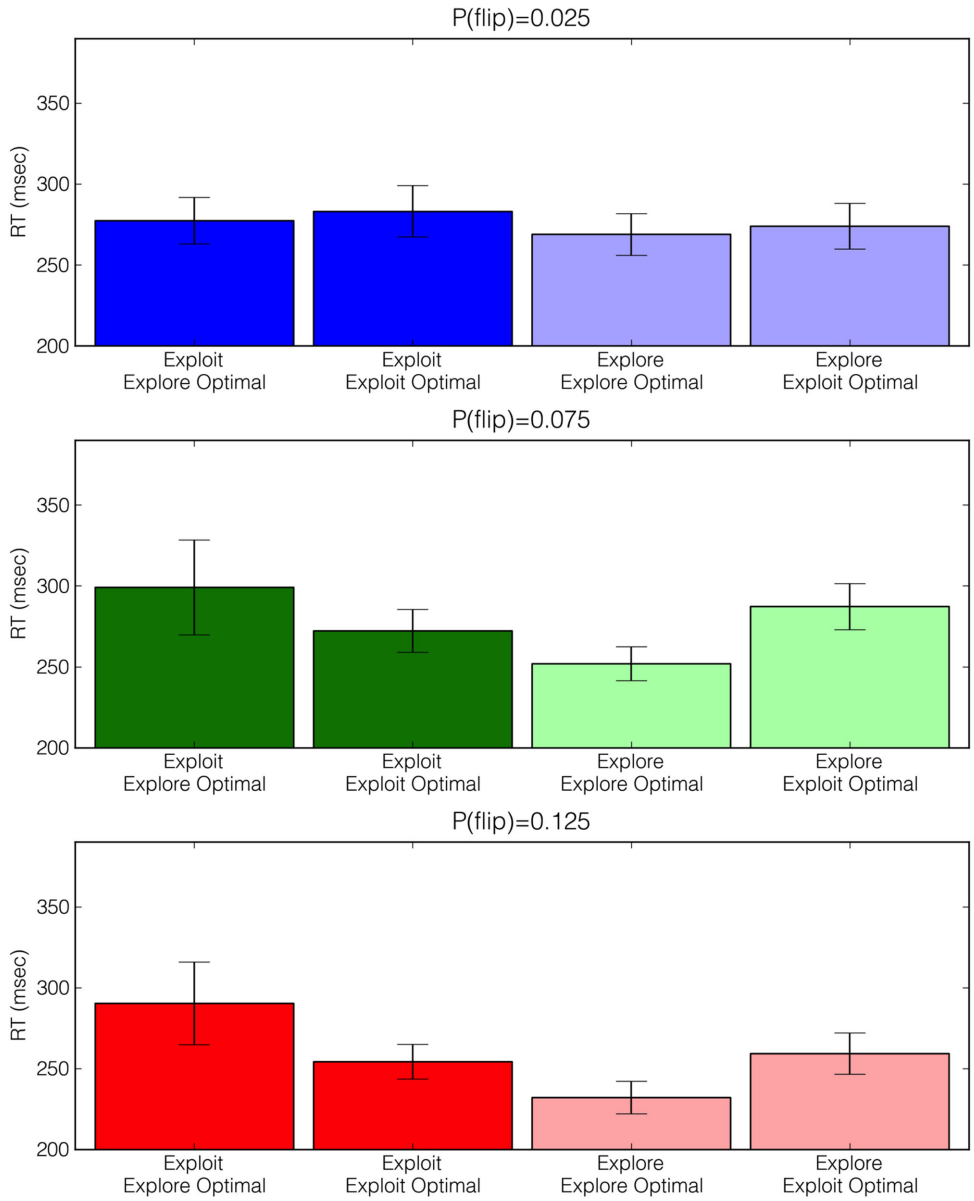


FIGURE 4 | Adjusted choice reaction times (RTs) across volatility conditions as a function of choice type (exploratory versus exploitative) and Ideal Actor choice prescription (exploit optimal versus explore optimal). For example, the far right bars represent situations in which participants explored when the Ideal Actor prescribed exploiting. Of note is that across the medium [P(flip) = 0.075] and high [P(flip) = 0.125] volatility conditions, participants exhibited larger RTs when they acted against the prescription of the Ideal Actor. See main text for analysis details.

choices (see **Figure 2**). This quality is not predicted by the reflexive Naïve RL model.

Given the reflexive and reflective models' qualitatively different predictions of sequential dependency, this comparison yields a strong test for determining which type of belief-updating better matches human behavior. Furthermore, our quantitative model comparisons revealed that the two reflective models clearly provide better fits than the purely reflexive Naïve RL model (see **Figure 3A**). These results suggest that people do exhibit marked sequential dependency and that their belief updates are reflective. Further, these results give credence to previous usage of reflective models of human choice behavior in bandit tasks (Daw et al., 2006; Boorman et al., 2009), which until now has not been empirically justified.

A number of related contributions dovetail with our reflexive versus reflective distinction. However, the tasks used in these studies differ in important ways. Recent work has sought to identify the contributions of model-based (i.e., reflective) and model-free (i.e., reflexive) strategies of choice in a multistep decision task (Daw et al., 2011). However, model-based behavior in Daw et al.'s study

Table 2 | Choice latency regression coefficients.

Coefficient	Estimate (SE)	*p*-Value
Volatility	−367.13 (185.93)	0.054
Choice type	5.96 (3.74)	0.1
Optimal choice	3.14 (4.26)	0.24
Volatility × choice type	0.06 (42.21)	0.76
Volatility × prescribed choice	35.06 (48.17)	0.4
Choice type × optimal choice	−8.87 (3.42)	0.006
Volatility × choice type × prescribed choice	7.89 (37.99)	0.71

Table 3 | Choice performance relative to deterministic Ideal Actor.

Condition	Deterministic Belief model	Stochastic Belief model
$P(\text{flip}) = 0.025$	0.854	0.988
$P(\text{flip}) = 0.075$	0.943	0.978
$P(\text{flip}) = 0.125$	0.946	0.954

did not entail updating beliefs about option payoffs across trials as uncertainty grew. Instead, a forward model of the environment was used to prospectively evaluate option values in accordance with the environment's transition structure. Related, Biele et al. (2009) found that a model prescribing use of higher-level strategies in a series of exploration–exploitation problems better predicted the patterns of sequential dependency in human behavior than a naïve sampling model that shared qualities with the Naïve RL model presented here

Critically, participants' differing levels of exploration across conditions could not be explained by the stochasticity with which they made choices: the best-fitting Softmax parameters did not decrease with environment volatility in either of the two models. Rather, the differences in rates of exploratory choice appear to be accounted for by the best-fitting $P(\text{flip})$ rates.

Notably, we also found that when decision-makers in medium and high volatility environments made sub-optimal decisions (insofar as the choices did not accord with the Ideal Actor's prescription), they exhibited larger choice RTs compared to when they made optimal choices. Since the Ideal Actor's choice prescriptions are a function of subjects' inferred trial-by-trial beliefs, these choice RTs provide another window into the belief-directed and reflective nature of their choices. Indeed, previous experimental work revealed that decision-makers exhibit greater choice latency when choosing options that will result in increased cognitive costs (Botvinick and Rosen, 2009) or when perceived logical conflict – and thus, the potential for making erroneous responses – is high (De Neys and Glumicic, 2008).

Quantitative comparison of the two reflective models – the Belief model and the Ideal Actor – favors the Belief model as a characterization of human choice behavior. Both models employ an Ideal Observer to maintain optimal beliefs about the expectation of immediate reward. However, the Ideal Actor also considers how an exploitative or exploratory action would inform its beliefs and precisely calculates the expected benefit of this information on future reward. Adding this value of information to the Ideal Observer's expectation of immediate reward, as the Ideal Actor does, decreases the model's ability to fit participants' choice behavior (shown in **Figure 3B**). Thus, it appears that people do not fully utilize these beliefs in a forward-planning way but, rather, appear to use the beliefs in a myopic fashion, in accordance with the Belief model.

In addition to providing a qualitative characterization of the structure of human exploratory choice, this paper contributes two tools for the study of exploratory decision-making. First, we present a task that allows us to disentangle reflective and belief-directed exploration from stochastic and undirected exploration, a feature absent in previous tasks used to examine exploratory choice (Worthy et al., 2007; Jepma and Nieuwenhuis, 2011). Further, the task is also sufficiently constrained to prescribe a statistically optimal pattern of choice, yielding the Ideal Actor. In turn, this model's belief-updating mechanism provides powerful tools for characterizing human choice behavior in this task, and its choice prescriptions afford the revelation of nuanced patterns of choice latencies that would be undetectable without such a model.

These formal models offer new ways of understanding what exploration is. The definition of exploratory choice depends on how one views the relationship between the actor and the structure of the environment or, even more abstractly, the relationship between the action and the hierarchical structure of the actor (Levinthal and March, 1993). In this paper, we define the exploitative action as a choice of the option that has yielded the highest-experienced payoff up to the time of choice. An alternative trial classification scheme could define an exploitative action as a choice of the option believed – according to a specific model – to give the highest payoff at the time of choice. However, we chose our definition because it avoids commitment to a particular model of choice; a choice is exploitative or exploratory regardless of any model under consideration. Under the alternative definition, exploration in our models could only arise from a purely stochastic process.

Simulations of the Ideal Actor and Belief model suggest an intriguing hypothesis, namely that stochastic behavior may be an adaptation to cognitive capacity limitations in long-range planning. As detailed in **Table 3**, performance of the Belief model (which does not plan) approaches that of the Ideal Actor (which does plan optimally) when stochasticity is incorporated into the Belief model's action selection. Consequently, we hypothesize that stochasticity in human decision-making may arise from a sub-optimal valuation process; the knowledge gained from potential exploration is not explicitly incorporated in valuation but is still obtained by random behavior. This hypothesis warrants further investigation through tasks and models that afford requisite discrimination.

ACKNOWLEDGMENTS

W. Bradley Knox was supported through much of this research by an NSF Graduate Student Research Fellowship. A. Ross Otto was supported by a Mike Hogg Endowment Fellowship. This research was also supported by AFOSR grant FA9550-10-1-0268, ARL grant W911NF-09-2-0038, and NIH grant MH091523 to Bradley C. Love and by NSF grant IIS-0917122, ONR grant N00014-09-1-0658, and FHA grant DTFH61-07-H-00030 to Peter Stone.

REFERENCES

Biele, G., Erev, I., and Ert, E. (2009). Learning, risk attitude and hot stoves in restless bandit problems. *J. Math. Psychol.* 53, 155–167.

Boorman, E. D., Behrens, T. E., Woolrich, M. W., andRushworth, M. F. (2009). How green is the grass on the other side? Frontopolar cortex and the evidence in favor of alternative courses of action. *Neuron* 62, 733–743.

Botvinick, M. M., and Rosen, Z. B. (2009). Anticipation of cognitive demand during decision-making. *Psychol. Res.* 73, 835–842.

Cassandra, A., Littman, M., and Zhang, N. (1997). "Incremental pruning: a simple, fast, exact method for partially observable Markov decision processes," in *Proceedings of the 13th Conference on Uncertainty in Artificial Intelligence*, Providence, 54–61.

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. D., Gershman, S. J., Seymour, B., Dayan, P., and Dolan, R. J. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron* 69, 1204–1215.

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.

De Neys, W., and Glumicic, T. (2008). Conflict monitoring in dual process theories of thinking. *Cognition* 106, 1248–1299.

Gureckis, T. M., and Love, B. C. (2009). Short-term gains, long-term pains: how cues about state aid learning in dynamic environments. *Cognition* 113, 293–313.

Jepma, M., and Nieuwenhuis, S. (2011). Pupil diameter predicts changes in the exploration-exploitation trade-off: evidence for the adaptive gain theory. *J. Cogn. Neurosci.* 23, 1587–1596.

Kaelbling, L., Littman, M., and Cassandra, A. (1998). Planning and acting in partially observable stochastic domains. *Artif. Intell.* 101, 99–134.

Lee, M. D., Zhang, S., Munro, M., and Steyvers, M. (2011). Psychological models of human and optimal performance in bandit problems. *Cogn. Syst. Res.* 12, 164–174.

Levinthal, D. A., and March, J. G. (1993). The myopia of learning. *Strategic Manage. J.* 14, 95–112.

Najemnik, J., and Geisler, W. S. (2005). Optimal eye movement strategies in visual search. *Nature* 434, 387–391.

Otto, A. R., Markman, A. B., Gureckis, T. M., and Love, B. C. (2010). Regulatory fit and systematic exploration in a dynamic decision-making environment. *J. Exp. Psychol. Learn. Mem. Cogn.* 36, 797–804.

Pearson, J. M., Hayden, B. Y., Raghavachari, S., and Platt, M. L. (2009). Neurons in posterior cingulate cortex signal exploratory decisions in a dynamic multioption choice task. *Curr. Biol.* 19, 1532–1537.

Pinheiro, J. C., and Bates, D. M. (2000). *Mixed-Effects Models in S and S-PLUS*. New York: Springer.

Ratcliff, R., and Rouder, J. (1998). Modeling response times for two-choice decisions. *Psychol. Sci.* 9, 347.

Schwartz, G. (1978). Estimating the dimension of a model. *Ann. Stat.* 6, 461–464.

Stankiewicz, B. J., Legge, G. E., Mansfield, J. S., and Schlicht, E. J. (2006). Lost in virtual space: studies in human and ideal spatial navigation. *J. Exp. Psychol. Hum. Percept. Perform.* 32, 688–704.

Steyvers, M., and Brown, S. (2005). "Prediction and change detection," in *Proceedings of Advances in Neural Information Processing Systems (NIPS)*, Vancouver.

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

Walton, M. E., Devlin, J. T., and Rushworth, M. F. S. (2004). Interactions between decision making and performance monitoring within prefrontal cortex. *Nat. Neurosci.* 7, 1259–1265.

Worthy, D. A., Maddox, W. T., and Markman, A. B. (2007). Regulatory fit effects in a choice task. *Psychon. Bull. Rev.* 14, 1125–1132.

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: 19 November 2011; paper pending published: 05 December 2011; accepted: 27 December 2011; published online: 31 January 2012.

Citation: Knox WB, Otto AR, Stone P and Love BC (2012) The nature of belief-directed exploratory choice in human decision-making. Front. Psychology ***2****:398. doi: 10.3389/fpsyg.2011.00398 This article was submitted to Frontiers in Cognitive Science, a specialty of Frontiers in Psychology.*

APPENDIX

THE LEAPFROG TASK

Here we give a more technical description of the Leapfrog task, using the framework of reinforcement learning (RL). We then briefly introduce partially observable Markov decision processes (POMDPs) and map the Leapfrog task to a highly compacted POMDP task, which we use in the following section to formulate the Ideal Observer, the belief-maintaining component of our reflective models, and the Ideal Actor itself.

We now describe the Leapfrog task more formally within the RL framework. Our task consists of two actions, A and B, two corresponding state variables, s_a or s_b, and reward is 1 for one action and 0 for the other. Given an action X, the reward is 1 if $s_X > s_Y$ and 0 otherwise. $s_a = 10$ and $s_b = 20$ when the task starts. At each trial, there is a fixed probability that the lower s_a and s_b increases by 20, switching the two actions' rewards and, consequentially, which action is optimal (thus the name Leapfrog). We call this probability *volatility* or $P(\text{flip})$. To create uncertainty that propagates from trial to trial – thus motivating exploration – the agent is not shown its reward. Rather, the agent only observes the value of the state variable that is tied to its action, s_a or s_b. The benefit of showing only s_a or s_b is that the lower-valued state variable can "jump" by 20 points and the agent will not know that the optimal actions are switched until it explores.

The task has a finite horizon (i.e., a set number of trials), and an agent's performance is evaluated by how much reward it accumulates. Note that, because of the limited observability and stochastic nature of the task, no decision-maker can guarantee to always choose the correct action. Therefore the Ideal Actor, in addition to its use as a model of the human, provides an upper bound on expected performance that facilitates assessment of how well people perform with various volatilities (and consequently, various difficulty levels).

Partially observable Markov decision processes

If the dynamics of a task are determined by a Markov decision process (see Sutton and Barto, 1998; for an introduction to MDPs), but the state cannot be directly observed by the agent – as is the case with the Leapfrog task – the task can often be modeled as a partially observable Markov decision process (POMDP).

To illustrate, consider a navigation task through a maze with a known map. When the decision-maker knows its exact location state at any time (e.g., via GPS), the task maps naturally to an MDP. On the other hand, a decision-maker without such global knowledge must use local features, such as corridors and corners, to localize itself. In this case, more than one location could share the same local features, and the decision-maker must use these observations along with its knowledge of recent movements and previous estimates of location to probabilistically estimate its location. In this case, the task maps well to a POMDP, as Stankiewicz et al. (2006) did for a similar navigation task.

More formally, a POMDP is defined by the set of variables $\{S, A, T, R, \Omega, O\}$ (Kaelbling et al., 1998). S and A are respectively the sets of states and actions. Given an action a_t and a current state s_t at time t, the state transitions to s_{t+1} at time $t+1$ with probability $T(s_t, a_t, s_{t+1})$ [i.e., $P(s_{t+1} \mid s_t, a_t)$]. At each time step, the agent also receives a real-valued reward, $r = R(s, a)$, and an observation O from the set Ω of possible observations. The probability of an observation o_t can be modeled equivalently as either $O(a_{t+1}, s_t, o_t)$ or $O(a_t, s_t, o_t)$ (Kaelbling et al., 1998). In the undiscounted case such as the Leapfrog task, an agent's goal in a POMDP is to choose actions that maximize return, defined as $E\left[\sum_{t=0}^{\infty} r_t\right]$.

Within a POMDP, optimal actions are determined not only by expectations of immediate reward and transitions to next states (as in MDPs), but also by the value of knowledge that actions yield. Therefore, an optimal action can have the sole purpose of gathering information about the true state. Note that a POMDP is a formal description of a task and is separate from a model of choice within the task.

Specifying the Leapfrog task as a three-state POMDP

As described, the leapfrog task can be specified as a POMDP with two state variables; s_A and s_B can respectively take values in $\{10, 30, \ldots, 10 + 10n\}$ and $\{20, 40, \ldots, 20 + (10n)\}$, where n is the number of trials.

However, we can specify the task much more compactly, reducing the belief space to three dimensions (one per state) from the n^2 dimensions it would otherwise have, making the Leapfrog task tractable to solve exactly. To justify this more compact representation, we first let s_H be the action-tied state variable (s_A or s_B) with the Highest observed number, which we call o_H, and $s_{\neg H}$ be the other action-tied state variable; H and $\neg H$ are their corresponding actions. Thus H frequently changes its mapping between the two possible actions in the task. Consequently, according to the definitions in the main text, choosing H would constitute an exploitative choice and choosing $\neg H$ would constitute an exploratory choice. At any trial, the agent knows the minimal states for $(s_H, s_{\neg H})$ are $(o_H, o_H - 10)$, based on the leapfrogging nature of the task.

For a trial with o_H, there are three possible action-tied state pairs $(s_H, s_{\neg H})$. The pairs $(o_H, o_H - 10)$ and $(o_H, o_H + 10)$ occur when there are zero and one unobserved jumps, respectively. When there are two unobserved jumps, resulting in the pair $(o_H + 20, o_H + 10)$, the agent is guaranteed to observe at least one jump regardless of its action. Since there is at most one new jump per trial, this guarantee of observing a jump makes it impossible to have more than two unobserved jumps. Note that action H (i.e., exploiting) receives a reward of 1 only when there are zero or two unobserved jumps. Therefore, the three possible states of our compacted POMDP are 0, 1, or 2 unobserved jumps. Additionally, a belief within this compacted POMDP is a vector of the probabilities that there are 0, 1, or 2 jumps, which necessarily sum to 1. Following this, the compacted observations are the number of previously unobserved jumps seen in a trial (0, 1, or 2), where payoffs $o_H - 10$ and o_H yield the same observation of 0 previously unseen jumps. In summary, the compacted POMDP has three states, three observations, two actions, and two possible rewards values.

MODELS OF HUMAN BEHAVIOR AND THE IDEAL OBSERVER

This section gives a full technical description of the Ideal Observer, which optimally maintains a distribution over possible state, and the Ideal Actor, the model that reflectively incorporates optimal beliefs along with an exact assessment of the information-based value of actions. On the explicative path to the Ideal Actor, we

comment on the other two models of human behavior used in our evaluations, Naïve RL and Belief.

All models choose options based on a Softmax choice rule that takes a model's valuation of each action as input. The probability of choosing option A at time t with payoff belief b_t, is

$$P_t(A) = \frac{e^{\gamma Q_t(b_t,A)}}{e^{\gamma Q_t(b_t,A)} + e^{\gamma Q_t(b_t,B)}}.$$

Here, $Q_t(b_t, \cdot)$ is the model's assessed value of option A or B, and γ is the Softmax inverse temperature parameter, the determination of which is described in the Model Fits section of this paper's body.

Naïve RL model

The simplest model reflexively maintains beliefs about payoffs based only on what it has seen. In other words, it believes the point payoffs for each action are those most recently observed. Therefore, the Naïve RL model assumes that action H and $\neg H$ respectively give rewards of 1 and 0. Its expectation of each action's reward, $Q(\cdot)$, is input into a Softmax action selector, giving it a constant probability of exploring or exploiting:

$$P(H) = \frac{exp[\gamma Q(H)]}{exp[\gamma Q(H)] + exp[\gamma Q(\neg H)]} = \frac{exp[\gamma]}{exp[\gamma] + 1}.$$

Here, γ is the Softmax inverse temperature parameter. In Softmax action selection, as this parameter rises, the probability that the highest-valued action (i.e., the greedy action) will be chosen increases. When the Softmax parameter approaches infinity, actions become deterministically greedy; at zero, the parameter creates uniformly random action selection. As mentioned in the main text, the Naïve RL model is equivalent to a memory-based RL agent, with a memory size of one, which is appropriate given the deterministic nature of the payoffs and that payoffs never return to previous values; a larger memory would not yield useful information for a reflexive model performing the Leapfrog task. Algorithmically, this model is equivalent to the Softmax model used in Worthy et al. (2007) and Otto et al. (2010), with a learning rate of 1.

Ideal Observer

An Ideal Observer uses past actions and observations optimally to update its belief distribution over the set of states. The Ideal Observer – agnostic to action selection – is used as a component of the belief and Ideal Actor models described below, providing correct beliefs at each time step. Because POMPDs by definition satisfy the Markov property (Kaelbling et al., 1998), belief updates can be performed with only the past belief, the last action, and the last observation, given knowledge of the observation and transition functions. In other words, the Ideal Observer can dispense with the remainder of its history of actions, beliefs, and observations.

Below, we show the derivation of our optimal Bayesian belief-updating procedure, which is specific to the case of POMDPs where observation o_t is a function of s_t and a_t, not the more typical s_t and a_{t-1},[1] since the number of unobserved jumps and the action determine the number of newly observed jumps. The final line of this derivation, a function of known distributions, is used to calculate the next state belief. In this notation, we put a_t and b_t after a ";" because they are fixed and known and are thus considered to parameterize the probability distributions.

$$P(s_{t+1} = i|o_t; a_t, b_t) = \frac{P(s_{t+1} = i|o_t; a_t, b_t)}{P(o_t; a_t, b_t)}$$

$$P(s_{t+1} = i|o_t; a_t, b_t) \propto P(s_{t+1} = i, o_t; a_t, b_t)$$

$$b_{t+1}(i) \propto P(s_{t+1} = i, o_t; a_t, b_t)$$

$$b_{t+1}(i) \propto \sum_j P\left(s_{t+1} = i, o_t, s_t = j; a_t, b_t\right)$$

$$b_{t+1}(i) \propto \sum_j P\left(o_t|s_{t+1} = i, s_t = j; a_t\right) P\left(s_{t+1} = i, s_t = j; a_t, b_t\right)$$

$\left(o_t \text{ is cond. indep. of } s_{t+1} \text{ given } s_t\right)$

$$b_{t+1}(i) \propto \sum_j P\left(o_t|s_t = j; a_t\right) P\left(s_{t+1} = i\,|s_t = j; a_t\right) P\left(s_t = j; b_t\right)$$

$$b_{t+1}(i) \propto \sum_j P\left(o_t|s_t = j; a_t\right) P\left(s_{t+1} = i|s_t = j; a_t\right) b_t(j)$$

Belief model

We can easily specify a model that values actions by their expected immediate rewards according to the Ideal Observer, creating a more sophisticated action selection technique than simply always choosing H. More precisely, this more sophisticated Belief model is more likely to choose action H than $\neg H$ when $b_{t+1}(0) + b_{t+1}(2) > 0.5$ – that is, when it believes that there are probably 0 or 2 unobserved jumps and thusly action H is expected to yield the higher immediate reward.

If the Belief model chooses the action deterministically, the model is optimal with respect to maximizing immediate reward. However, this model would not be fully optimal in the long-term because its choices fail to take into account the informational benefit of each action.

Ideal Actor model

An Ideal Actor uses beliefs reflectively provided by its Ideal Observer component to consider expected immediate reward, but it also evaluates an action's effect on its longer-term expectation of return caused by the change in its belief distribution. In other words, the Ideal Actor sometimes chooses actions with lower immediate rewards to increase its knowledge about the true state, facilitating more informed decisions in future trials. To implement the Ideal Actor, we employed the Incremental Pruning algorithm (using the POMDP-Solve library, Cassandra et al., 1997), an exact inference method that calculates action-value functions (i.e., Q-functions) for each time horizon (i.e.,

[1] By our subscripting, o_t occurs after s_t is set, immediately after a_t, and before s_{t+1}. This ordering is because the action is a causal factor of the observation, and the observation intuitively comes before the probabilistic jump that finally determines s_{t+1}. However, calling the observation o_{t+1} is an appropriate alternative.

number of trials remaining).[2] We used the implementation of Incremental Pruning from the POMDP-Solve library. This action-value function Q_t, where t is the horizon, takes as input a belief b_t and an action H or $\neg H$ and outputs a real-number value. The belief vector input to Q_t comes from the Ideal Observer's belief, making the Ideal Actor a reflective model. If acting optimally, the Ideal Actor deterministically chooses $\text{argmax}_a\ Q_t(b_t, a)$, where $a \in \{H, \neg H\}$. Unlike other actor models examined in this paper and in previous work, the Ideal Actor chooses actions based on both its belief about the immediate reward and the expected benefit from the knowledge gained by choosing each action. **Figure 1A** illustrates the trial-by-trial relative Q-values – that is, $Q_t(b_t, H) - Q_t(b_t, \neg H)$ – which are a function of the Ideal Observer's trial-by-trial belief (shown in **Figure 1B**).

[2]The time horizon affects the optimal action for a given belief vector because the value of knowledge changes as the final trial approaches.

MODEL-FITTING PROCEDURE

For each model, we sought parameter estimates that maximized the likelihood of each participant's observed choices given their previous history of choices and outcomes:

$$L_{\text{model}} = \prod_t P_{c,t}$$

where c,t reflects the choice made on trial t and $P_{c,t}$ is the probability of the model choosing c,t, informed by participant's choice and payoff experience up to trial t. We conducted an exhaustive grid search to optimize parameter values for each participant. To compare models, we utilized the Bayesian information criterion (BIC: Schwartz, 1978), which is calculated by

$$\text{BIC}_{\text{model}} = -2 \times \ln(L_{\text{model}}) + k_{\text{model}} \cdot \ln(n)$$

where k is the model's number of free parameters and n is the number of trials being fit (500 in all cases).

frontiers in PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 13 March 2012
doi: 10.3389/fpsyg.2012.00068

Prediction and control in a dynamic environment

***Magda Osman*[1,2]* *and Maarten Speekenbrink*[2]**

[1] *Biological and Experimental Psychology Centre, School of Biological and Chemical Sciences, Queen Mary College, University of London, London, UK*
[2] *Cognitive, Perceptual and Brain Sciences, Division of Psychology and Language Sciences, University College London, London, UK*

Edited by:
Erica Yu, University of Maryland, USA

Reviewed by:
Zheng Wang, Ohio State University, USA
Joachim Funke, Ruprecht-Karls-Universität Heidelberg, Germany

***Correspondence:**
Magda Osman, Biological and Experimental Psychology Centre, School of Biological and Chemical Sciences, Queen Mary College, University of London, Mile End, London E1 4NS, UK.
e-mail: m.osman@qmul.ac.uk

The present study compared the accuracy of cue-outcome knowledge gained during prediction-based and control-based learning in stable and unstable dynamic environments. Participants either learnt to make cue-interventions in order to control an outcome, or learnt to predict the outcome from observing changes to the cue values. Study 1 ($N = 60$) revealed that in tests of control, after a short period of familiarization, performance of Predictors was equivalent to Controllers. Study 2 ($N = 28$) showed that Controllers showed equivalent task knowledge when to compared to Predictors. Though both Controllers and Predictors showed good performance at test, overall Controllers showed an advantage. The cue-outcome knowledge acquired during learning was sufficiently flexible to enable successful transfer to tests of control and prediction.

Keywords: dynamic, decision making, learning, prediction, control

INTRODUCTION

Imagine a scenario in which after a visit to your doctor, you are told that your blood pressure is too high. Following the doctor's recommendation, you have decided to take up more exercise by jogging, and to change your diet by reducing your salt intake. In addition you have bought the latest mobile phone application which is a self monitoring device that can help you track and analyze your blood pressure. By steadily increasing exercise and reducing intake of salty foods (interventions), and measuring your blood pressure at regular intervals, the idea is to track the relationship between exercise and salt intake (cues) and blood pressure (outcome) to assess how effective the new healthy regime is. You notice that your blood pressure recordings fluctuate while you are resting (internal, or endogenous, changes in the outcome), as well as when you have gone for a jog, or eaten (external, or exogenous, changes in the outcome[1]). On some days your blood pressure does not reduce as substantially as on other days, and so on those occasions you decide to take more exercise.

The above example can be described as dynamic decision making, in which decisions (i.e., choosing a course of action) are adapted to the ongoing changes in the outcome. More specifically this type of dynamic decision making involves learning about the cue-outcome relations via cue-intervention (control-based decision making). A different approach may involve first monitoring for a period of time the cue-outcome relations by tracking changes in cues (i.e., observing your usual salt intake from day to day) from which one can make predictions about changes to the outcome. This is an example of prediction-based decision making in which cue-outcome relations are acquired via estimates of the expected outcome value. The difference between the two is that in the latter case there is no cue-intervention; instead, changes in cue and outcome values are used to adjust the predictions made.

Both prediction and control are methods of acquiring cue-outcome knowledge, and are examples of dynamic decision making. General models of learning (e.g., Reinforcement learning/reward-based learning, Schultz et al., 1997; Sutton and Barto, 1998; Schultz, 2006) would propose that both prediction and control are comparable ways of acquiring cue-outcome knowledge, because both rely on prediction errors signals arising from a comparison between predicted and actually obtained outcomes. One avenue that this present study explores is to directly compare the accuracy of cue-outcome knowledge when gained via prediction and when gained via control, in order to explore these general claims.

Dynamic decision making (hereafter DDM) is an area of decision making research that is growing in popularity (Brown and Steyvers, 2005; Osman et al., 2008; Lurie and Swaminathan, 2009; Osman, 2010a,b, 2011; Speekenbrink and Shanks, 2010). Typically, tasks designed to examine this type of process involve situations in which the decision maker has a clearly defined goal from the outset. Because the environment is probabilistic and/or dynamic, the desired goal cannot usually be achieved in one step. Therefore the decision maker must plan a series of actions that will help to incrementally move them closer to the goal. Thus, the process of decision making in dynamic environments is often described as goal directed, and the decisions themselves are usually inter-dependent across trials (Brehmer, 1992; Funke, 1992; Osman, 2010a).

Given this type of characterization, planning actions in order to obtain future outcomes is an important component of DDM, and therefore it is likely that some control-based decision are informed by predictions (Osman, 2010a,b). Moreover, the decision maker needs to be sensitive to possible changes in the environment in order to adapt their planned interventions accordingly. For

[1]"An endogenous variable [that] at time t has an effect of its own state at time $t + 1$ independent of exogenous influences that might add to the effect" (Funke, 1993, p. 322).

this reason DDM is also referred to as adaptive decision making (Cohen et al., 1996; Klein, 1997; Lipshitz et al., 2001). Typically in DDM contexts accurate estimates of the environment's future behavior are important for planning the best interventions that would achieve the right outcome. Consider the following engineering example, in which the goal is to maintain a supply of electricity from a power station that meets the consumers' needs. Since there are obvious changes in the demand according to daily as well as seasonal changes, this enables fairly stable predictions as to consumer use that enables matching the generator's output to load predictions. However, it is important to estimate future demands in order to adjust the system quickly enough as the changes arise (e.g., sudden unseasonably low temperatures). Predictions regarding fuel consumption are essential to informing the choice of intervention needed to control power supply, and research in the engineering domain supports this (e.g., Leigh, 1992). In fact, optimal control theory (Bryson and Ho, 1975) provides a formal basis for analyzing the changing states of a system in order to provide a complete stochastic description of control under uncertainty. Moreover, the theory has been implemented in psychological research on human motor control (e.g., Körding and Wolpert, 2006).

Many theoretical accounts of DDM (Vancouver and Putka, 2000; Burns and Vollmeyer, 2002; Bandura and Locke, 2003; Goode and Beckmann, 2010; Osman, 2010a,b) and formal descriptions of DDM (Gibson et al., 1997; Sun et al., 2001; Gibson, 2007) claim that controlling outcomes to a target goal involves decisions based on prediction. In their Social-Cognitive theory, Bandura and Locke (2003) refer to prediction as a feedforward process which has a motivational component attached to it. They propose that by estimating future outcomes and future success, when reinforced, successful predictions drive the decision maker to achieve even more accurate or more successful control of the environment. Burns and Vollmeyer (2002) describe DDM within the context of hypothesis testing. They posit that decision makers learn best by exploring the DDM environment. This entails generating expectations about the associations between their actions and the effects they will produce. The effects they achieve through their actions are used as feedback to update hypotheses about the relationship between cues and outcomes in the environment. Gibson's (Gibson et al., 1997; Gibson, 2007) neural network model also posits that DDM involves the development of hypotheses about cue-outcome associations, from which planned interventions are made. The model includes two submodels: a forward model and an action model. The action submodel decides what action to take based on the current state of the environment and the distance from the target goal. Each action generates an expected outcome which is then compared against the goal; from this the action which is most likely to minimize the distance between the expected state and the goal is chosen. The forward submodel takes as input an action that has been executed and compares its outcome with the goal. It then generates as output an expected outcome which is used to derive an error signal. Back propagation is then used to adjust the connection weights between action and predicted outcome, to improve the ability to predict the effects of actions on the environment.

Osman's (2010a,b, 2011) Monitoring and Control framework also proposes that there are two different judgments made regarding a DDM environment: its predictability and its controllability. People are sensitive to the endogenous as well as the exogenous changes in the environment. This is because they are repeatedly updating their expectancies of the outcomes that will occur in the environment, and this informs their subjective estimates of confidence in predicting outcomes of events in a dynamic environment (predictability of the environment). As well as developing expectancies about the outcomes that are likely to occur, the actions taken in the environment generate outcomes which can be fed back in order to update one's expectancies. This informs their subjective estimates of expectancy that an action executed will achieve a specific outcome in a control system (predictability of control).

In sum, the consensus amongst many theorists in the DDM research domain is that prediction forms a strong component of the decision making process needed to plan interventions in order to reliably achieve a target goal (Vollmeyer et al., 1996; Burns and Vollmeyer, 2002; Gibson, 2007; Osman, 2010b). However, there are alternative accounts of DDM that do not assign a role to prediction. For instance, Berry and Broadbent (1984, 1987, 1988) claim that people typically fail to provide a veridical verbal description of their acquired knowledge or to accurately predict different states of the DDM task because the cue-outcome associations are far too complex to learn explicitly. Accurate control of the outcome is achieved through active intervention, from which successful cue-interventions that generate desired outcomes are stored in memory and later recalled when faced with similar task situations. Because there is no need for predictive learning, there is no cue-abstraction, instead successful control performance is based solely on representations of action-outcome associations.

There is little empirical research that has investigated the contribution of prediction (i.e., estimating the outcome that will occur) to control (i.e., planning actions to achieve an outcome) in a dynamic environment. Thus, it is not clear whether controlling a dynamic environment relies on prediction, or for that matter, whether it is possible to learn about a dynamic environment from prediction alone. This is a limitation given the kinds of different theoretical claims made about the relevance of prediction to control. Therefore, the present study aims to separate prediction-based from control-based learning, and directly compare the accuracy of cue-outcome knowledge gained from both forms of learning.

There are two relevant literatures that have considered these issues in isolation, namely, multiple cue probability learning (MCPL) and complex dynamic control (CDC). Research on MCPL examines how we integrate information from different sources in order to learn to predict outcomes in probabilistic environments (Speekenbrink and Shanks, 2010). Typically, MCPL tasks involve presenting people with cues (e.g., symptoms – rash, fever, headaches) which are probabilistically associated with an outcome (e.g., disease – flu, cold). Participants are asked to predict the outcome (e.g., flu) for various cue combinations (e.g., rash, fever), and then receive outcome feedback on their prediction. Research on CDC has investigated the way in which decisions are formed over time in order to reliably control outcomes in dynamic and probabilistic environments (Osman, 2008a). People decide from a set of inputs (cues; e.g., drug A, drug B, drug C) actions that are relevant (e.g., selecting drug A at dosage X) for achieving and maintaining a particular output (outcome; e.g., reduce the spread

of disease Y). As with MCPL environments, input–output (hereafter: cue-outcome) associations are probabilistic, and also need to be learned. In addition, the environment can also be dynamic, in the sense that the outcome may change independently of actions made by the individual (e.g., disease Y spreads at a particular rate). In MCPL tasks, learning about the cue-outcome associations is indirect because only observations of the cue patterns are used to predict the events with the aim of reducing the discrepancy between expected and actual outcomes. By contrast, learning about the cue-outcome associations is direct in CDC tasks because the cues are manipulated and the change in outcome that follows serves as feedback which is used to update knowledge of cue-outcome associations.

We briefly review the MCPL and CDC literature in the next section. These research domains have remained relatively separate (Osman, 2010a). A secondary objective of this study, besides comparing prediction-based decision making with control-based decision making in a DDM environment, is to consider the potential common ground shared by the MCPL and CDC research paradigms.

Early MCPL studies were largely concerned with varying different properties of the task environment and examining the effects on predictive accuracy. Properties of the environment that were varied include: the number of cue-outcome associations (Slovic et al., 1971), the combination of continuous cue and binary outcomes (Vlek and van der Heijden, 1970), the presence of irrelevant cues (Castellan, 1973), the type of feedback presented (Björkman, 1971; Hammond et al., 1973; Muchinsky and Dudycha, 1975; Holzworth and Doherty, 1976), time constraints (Rothstein, 1986), and cue validities (Castellan and Edgell, 1973; Edgell, 1974). In the main, the evidence suggests that people show sensitivity to the cue validities and cue probabilities. Moreover, it has been suggested that people update their knowledge of the cue-outcome associations by developing hypothesis testing strategies.

More recently, the weather prediction task (WPT), developed by Knowlton et al. (1994), has become a popular paradigm for studying prediction-based decision making. It was originally designed to study incremental learning processes in clinical populations (e.g., amnesic patients). Knowlton et al. (1994) proposed that amnesics were able to successfully predict outcomes in the task through incidental acquisition of cue-outcome associations (procedurally) rather than explicit cue-abstraction (declaratively). That is to say, accurate cue-outcome knowledge of a probabilistic environment was acquired through processes that did not require deliberate evaluation of expected outcomes and actual outcomes. In support, others (e.g., Knowlton et al., 1996; Poldrack et al., 2001) have since shown that knowledge acquired procedurally is not available to conscious inspection, because of the implicit nature of learning of cue-outcome knowledge. However, critics of this position have investigated these claims in non-clinical (Price, 2009) as well as clinical populations (Speekenbrink et al., 2008) and have found evidence to suggest a correspondence between the accuracy of explicit cue-outcome knowledge and predictive accuracy, suggesting that people do have insight into the cue-outcome knowledge they have acquired. This has been taken to suggest that the strategies people develop to make their predictions are consciously accessible, and that there is an exhaustive evaluative process involved in comparing expected outcomes with actual outcomes on a trial by trial basis.

Research on CDC originated with the work of Dörner (1975), who used control tasks to simulate real world decision making in complex domains (e.g., maintaining ecologies). This line of research is now referred to as complex problem solving (Buchner, 1995). Better known however is Broadbent's early work on control tasks (Broadbent, 1977; Berry and Broadbent, 1984), which, like MCPL studies, demonstrated that procedural learning is the mechanism that supports control-based decision making. As with the WPT, much of the work that followed from Berry and Broadbent's (1984) pivotal study has examined the type of knowledge acquired under procedural learning. Broadbent's Exemplar theory proposes that while interacting with a CDC task, manipulations of specific inputs that lead to successful outcomes are stored as exemplars in a type of "look-up table." Controlling a CDC task is then based on a process of matching environmental cues to similar stored exemplars, rather than through following the rules acquired in cue-abstraction. In support, some studies have shown that there is limited transfer of control performance to versions that differ from the initial training task, and that people lack insight into the knowledge gained during learning (Dienes and Fahey, 1995, 1998; Gonzales et al., 2003).

In contrast, studies that have encouraged hypothesis testing behavior – either through explicit instruction, or task manipulations such as presenting a task history of action-outcome associations during learning – have shown that rule learning can lead to successful control performance. Moreover, under these conditions cue-outcome knowledge is transferable beyond the trained task environment to other variants as well as other goal criteria, and people show accurate reportable knowledge of cue-outcome associations (Sanderson, 1989; Burns and Vollmeyer, 2002; Osman, 2008a,b,c). In particular, two recent studies (Osman, 2008a,b) examined if actively engaging with a CDC task during learning is necessary for accurate control. Berry (1991) claimed that successful control is dependent on procedural learning, which does not generate cue-outcome knowledge. By extension, learning purely through observation would prevent the uptake of decision-action-outcome associations, and reduce control performance. Osman (2008a,b) used a yoking design to compare the effects of learning through action vs. learning through observation (i.e., simply observing cue changes and tracking their effects on the outcomes). The findings challenged Berry's original claims by demonstrating that learning through observation and learning through action generate equivalent declarative as well as procedural knowledge. Moreover, the findings demonstrated that rather than independent, accuracy of cue-outcome knowledge was positively associated with control performance. There is an emerging consensus from recent work with CDC tasks that people apply hypothesis testing strategies when planning actions which enables cue-abstraction (e.g., Sun et al., 2001; Burns and Vollmeyer, 2002; Osman, 2010a,b). Much like the MCPL studies, these finding suggest that learning to control outcomes involves a deliberate process of evaluating actions against expected outcomes. Though crucially in CDC tasks there is an additional step which involves selecting actions that are expected to reduce the discrepancy between an achieved and target outcome.

As discussed earlier, MCPL and CDC tasks share many common properties. In particular they are environments in which the decision maker is required to learn probabilistic relationships between cues and outcomes. Moreover, both predicting and controlling an outcome rely on inferring changes in outcome values from patterns of cue/input values. Besides the obvious difference that controlling an outcome involves direct manipulation of the inputs whereas predicting outcomes does not, the other key difference is the way in which the outcome is evaluated (Osman, 2010a). When controlling outcomes, the achieved outcome is evaluated with respect to the target goal. When predicting outcomes, the predicted outcome is evaluated with respect to the achieved outcome. While there are no direct comparisons of prediction- and control-based decision making in dynamic environments, Enkvist et al. (2006) compared the accuracy of cue-outcome knowledge of a group instructed to predict the outcome in a MCPL task with a group instructed to control the outcome to a specific criterion. They found that control-based decision making led to more accurate cue-outcome knowledge. However, this finding was reversed when binary rather than continuous cues were used. Also, as already mentioned, Berry (1991) and Osman (2008a,b,c) yoked active controllers with passive observers and compared both groups on their ability to control an outcome to criterion over successive trials. While Berry found an advantage for active controllers, Osman found no difference between passive and active learners. Osman noted that this divergence can be explained by a critical difference between the studies: in Osman's study, all participants were encouraged to learn through hypothesis testing, while in Berry's study participants were explicitly instructed not to engage in hypothesis testing.

We may conclude from these findings that, under certain conditions (i.e., in MCPL tasks when cue-outcome associations involve continuous variables, and in CDC tasks where instructions prevent hypothesis testing), cue-outcome knowledge is more accurate when learning is control-based as compared to prediction-based. However, this conclusion is only tentative, given that to date there have been no studies that compared both forms of learning within the same dynamic decision making task.

In the present study, we sought to induce prediction-based DDM by asking people to first learn to predict the effects of observed actions on the state of the environment. Later on, they were then asked to control the environment. We compared this condition to one in which people were instructed to achieve and maintain a specific target outcome from the outset. The present study used a yoking design such that Controllers and Predictors experienced exactly the same cue-outcome information. To our knowledge, this study is the first to allow such a direct comparison of prediction and control-based learning in a DDM environment. Based on reinforcement learning models, if accurate cue-outcome knowledge is dependent on the generation of prediction errors that are generated either through prediction or control, then performance at test should be equivalent regardless of training condition (Controllers, Predictors). If however, there are fundamental differences between prediction-based and control-based learning, then from previous results in the DDM and MCPL domain, we would expect that at Predictors should show an overall advantage in test of prediction, whereas Controllers should show an advantage on tests of control.

EXPERIMENT 1

In Experiment 1, we compared control performance between a group who learnt to control an outcome from the outset (Controllers) and a yoked group who first learnt to predict the outcome from the cue manipulations of those in the first group (Predictors). In addition, we manipulated the level of perturbation (i.e., noise) in the environment. A recent study (Osman and Speekenbrink, 2011) investigated DDM in tasks in which the stability of the environment was varied such that the autonomous changes to the outcome were relatively small (low noise) or large (high noise). The findings revealed that people were sensitive to the different rates at which the outcome fluctuated; accuracy of control performance suffered in high noise environments and more sub-optimal strategies were developed. By manipulating the level of noise in Experiment 1, we examined whether this effect on performance generalizes to prediction as well as control.

METHODS

Participants

The 60 graduate and undergraduate students who took part in the study were recruited from the University College London subject pool and were paid £6 for their participation. The assignment of participants to the four conditions was semi-randomized. There were a total of four groups (Control High Noise, Control Low Noise, Prediction High Noise, Prediction Low Noise[2]), with 15 participants in each. Pairs of participants (controller and yoked predictor) were randomly allocated to one of the two noise conditions[3].

DESIGN AND MATERIALS

Experiment 1 included two between subject manipulations, namely learning mode (prediction-based vs. control-based) and stability (high noise vs. low noise). Control performance was measured in two control tests. The task environment consisted of three cues and one outcome. One of the cues increased and one of the cues decreased the outcome. The third cue had no effect on the outcome. More formally, the task environment can be described as in the following equation

$$y(t) = y(t-1) + 0.65x_1(t) - 0.65x_2(t) + e(t)$$

in which $y(t)$ is the outcome on trial t, x_1 is the positive cue, x_2 is the negative cue, and e a random noise component, normally distributed with a zero mean and SD of 4 (Low Noise) or 16 (High Noise). The null cue x_3 is not included in the equation as it had no effect on the outcome.

PROCEDURE

Controllers were informed that as part of a medical research team they would be conducting tests in which they could inject a patient

[2]The assignment of noise to the system was first piloted in order to generate High variance (16 SD) and low variance (4 SD).

[3]In order to implement the yoking design in which a participant from each Prediction condition was yoked to a participant from the corresponding control condition, a control participant performed the experiment first in order to generate the learning trials presented to the yoked predictor.

with any combination of three hormones (labeled as hormones A, B, and C), with the aim of maintaining a specific safe level of neurotransmitter release. Predictors were assigned the same role, but were told that they would have to predict the level of neurotransmitter release by observing the level of hormones injected. **Figure 1** presents a screen-shot of a learning trial as experienced by Predictors and Controllers in the experiment. The task was performed on a desktop computer, using custom software written in C# for the .NET framework. The task consisted of a total of 80 trials, divided into two phases. The learning phase involved 40 learning trials and the test phase included two tests of 20 trials each.

Learning phase

On each trial, Controllers adjusted a slider to decide how much of each hormone to release (a value between 0 and 100). After confirming their decision, the effect was revealed visually on the outcome graph. On the next trial, the input values were reset to 0, but the outcome value was retained from the previous trial. Predictors were shown the input values chosen on that trial by the Controller they were yoked to. They were asked to predict the resulting outcome value by moving a slider. Once they were satisfied with their prediction, the actual outcome value was revealed alongside their prediction. On a subsequent trial, the previously predicted outcome value was omitted from the outcome graph, but the history of the actual outcome values on the previous five trials remained.

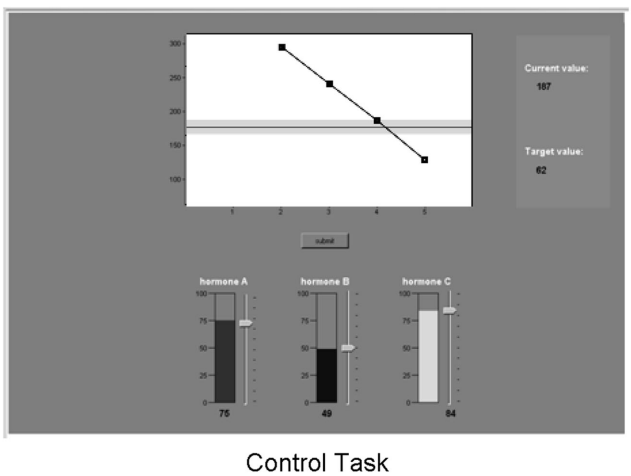


FIGURE 1 | Screen-shot of a learning trial for predictors and controller.

Test phase

The test phase was identical for Controllers and Predictors. Control Test 1 involved the same task as performed by Controllers in the learning phase; this was the first opportunity for Predictors to control the environment. Control Test 2 involved a different desired outcome level in order to examine success in controlling the system to a different criterion[4].

SCORING

Prediction performance was measured by a prediction error score $S_p(t)$ calculated as the absolute difference between predicted and expected outcome values:

$$S_p(t) = \left| P(t) - y(t-1) - 0.65x_1(t) + 0.65x_2(t) \right|,$$

in which $P(t)$ is a participant's prediction on trial t. We chose to compare predictions to expected rather than actual outcomes as the latter are subject to random noise, resulting in a biased comparison between the High and Low Noise conditions.

Control performance was measured by a control error score $S_c(t)$ calculated as the absolute difference between the expected and best possible outcome:

$$S_c(t) = \left| G(t) - y(t-1) - 0.65x_1(t) + 0.65x_2(t) \right|,$$

in which $G(t)$ is the goal on trial t: either the target outcome if achievable on that trial, or the closest achievable outcome.

RESULTS AND DISCUSSION

Learning phase: controllers

The learning phase was divided into four blocks of 10 trials each. For the following analyses, prediction and control error scores were averaged across each block for each participant; these are presented in **Figure 2**. A 4×2 ANOVA was conducted on the control scores, with Block (learning block 1, 2, 3, 4) as within-subject factor and Noise (high, low) as a between subject factor. As indicated in **Figure 2**, there was a main effect of Block, $F(3,84) = 28.89$, $p < 0.001$, partial $\eta^2 = 0.508$. To explore the possibility that accuracy of control-based decisions improved over blocks of trials, t-tests revealed that error scores were lower in Blocks 2, 3, and 4 as compare to Block 1 ($t = 3.72$, $p = 0.036$, $t = 3.76$, $p = 0.013$, $t = 3.04$, $p = 0.013$), no other differences reached significance. A main effect of Noise, $F(1,28) = 9.48$, $p = 0.004$, partial $\eta^2 = 0.253$, suggests that control performance was poorer in the High compared to the Low Noise condition. **Figure 2** also suggests more

[4]In the learning phase and in Control Test 1, the starting value was 178 and the target value throughout was 62. Participants were instructed to maintain the outcome within a safe range (± 10) of the target value. In Control Test 2, the starting value was 156, the target value was 74, and the safe range ± 5 from the target value.

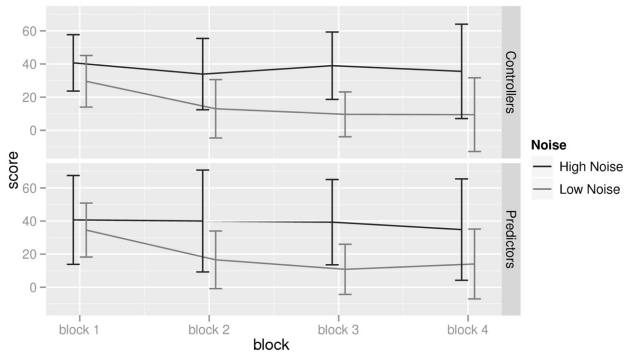


FIGURE 2 | Error scores (± SE) in the learning phase of Experiment 1. For Controllers, these are control error scores, and for Predictors, these are predictive error scores.

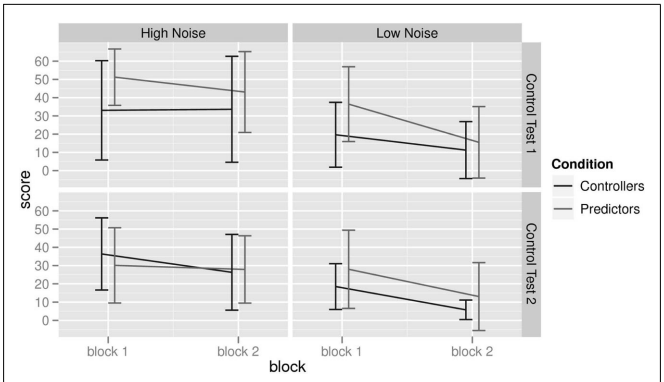


FIGURE 3 | Control error scores (± SE) of predictors and controllers in the test phase of Experiment 1.

pronounced learning in the Low Noise compared to the High Noise condition, which was supported by a significant Noise × Block interaction, $F(3,84) = 3.93$, $p = 0.011$, partial $\eta^2 = 0.123$.

Learning phase: predictors

A similar ANOVA on the prediction error scores showed a main effect of Noise, $F(1,28) = 6.72$, $p = 0.015$, partial $\eta^2 = 0.193$, confirming that Predictors in the Low Noise condition outperformed those in the High Noise condition. The main effect of Block, $F(3,84) = 2.95$, $p = 0.037$, partial $\eta^2 = 0.095$. To examine if differences in performance across-blocks reflected learning, t-test comparisons were conducted, and revealed that prediction error scores were lower in Blocks 2, 3, and 4 as compare to Block 1 ($t = 2.02$, $p = 0.036$, $t = 2.63$, $p = 0.013$, $t = 2.63$, $p = 0.013$), no other differences reached significance. To summarize, both Predictors and Controllers showed clear evidence of learning, and performance in both groups was negatively affected by increasing the level of noise in the system.

Test phase

The test phase provided the opportunity to compare the four conditions on an equivalent measure of performance. Each test was divided into two blocks of 10 trials each. The following analyses are based on participants' average control error scores in each block and test, as presented in **Figure 3**. Control error scores were first analyzed with a 2 × 2 × 2 × 2 ANOVA with Test (Control Test 1, 2) and Block (block 1, 2) as within-subject factors, and Learning Mode (prediction, control), and Noise (high, low) as between subject factors. The main effect of Test, $F(1,56) = 14.94$, $p < 0.001$, partial $\eta^2 = 0.211$, suggests a general practice effect: even though it included an unfamiliar goal, participants performed better in Control Test 2 than in Control Test 1. There were also practice effects within tests, as shown by a main effect of Block, $F(1,56) = 56.99$, $p < 0.001$, partial $\eta^2 = 0.504$. A significant Test × Learning Mode interaction, $F(1,56) = 6.00$, $p < 0.017$, partial $\eta^2 = 0.097$, indicates that Predictors improved their performance more over tests than Controllers. Controllers performed better than Predictors in Control Test 1, $t(58) = 2.11$, $p = 0.038$, but there was no difference in Control Test 2, $t(58) = 0.68$, $p = 0.51$. In addition, there was a significant Block × Task × Learning Mode interaction, $F(1,56) = 6.38$, $p = 0.014$, partial $\eta^2 = 0.102$. Follow-up t-tests indicated that Controllers outperformed Predictors only in the first block of Control Test 1, $t(58) = 3.14$, $p < 0.003$; there was no significant difference elsewhere (all p's > 0.29).

There were two significant effects involving noise. Overall, participants in the low noise conditions outperformed those in the high noise conditions, $F(1,56) = 14.52$, $p < 0.001$, partial $\eta^2 = 0.206$. In addition, a Block × Noise interaction, $F(1,56) = 13.31$, $p < 0.001$, partial $\eta^2 = 0.192$, indicates that this difference was particularly strong in the second block. To explore the difference between Controllers and Predictors in the first block of Control Test 1 in more detail, we compared Controllers' and Predictors' performance when they were first exposed to the control task. That is, we compared Controllers' performance in the first block of the learning phase with Predictors' performance in the first block of Control Test 1. The aim of this across-blocks comparison was to assess the benefit for prediction training prior to encountering the control task. The analysis revealed no significant difference, $F(1,28) = 0.44$, $p = 0.51$, partial $\eta^2 = 0.016$, suggesting that Predictors did not benefit from their four blocks of prediction learning. In contrast, the same analysis comparing Controllers performance in the first block of learning with their performance in the first block of Control Test 1, $F(1,28) = 17.38$, $p = 0.0002$, partial $\eta^2 = 0.375$, suggested that their prior experience was beneficial.

The results from this first experiment can be summarized as follows: given that the accuracy of controlling an environment is dependent to a greater extent on the way the outcome fluctuates (either in a less noisy or more noisy manner) rather than the mode of learning of the environment (either predicting or controlling the outcome), the results are broadly consistent with the view that control and prediction involve similar processes. Moreover, the evidence suggests that Osman and Speekenbrink's (2011) findings generalize to prediction as well as control. In addition, Predictors required a short period of familiarization with controlling the system in the first 10 trials of Control Test 1, but after that were able to control the environment as well as the Controllers who had been doing so from the outset. However, in this initial first block of testing, Predictors' performance was no different to

Controllers' performance in the first block of learning, suggesting that prediction-based learning about cue-outcome relations, regardless of environmental stability, was not directly transferable to control. This raises the question whether control-based learning generates flexible cue-outcome knowledge that would facilitate performance on tests of prediction.

EXPERIMENT 2

Experiment 1 indicated a small and transient difference in performance between Controllers and Predictors. In Experiment 2 we sought to further examine this by replicating our findings and also examining whether control-based learning generates flexible knowledge which can be used to predict as well as control the outcome.

METHOD

Participants

The 30 graduate and undergraduate students that took part were recruited from Queen Mary College and were paid £6 for their participation. The assignment of participants to the conditions followed the same yoking design as Experiment 1. There were two conditions (Controllers, Predictors), each with 15 participants.

DESIGN AND PROCEDURE

Experiment 2 was identical to Experiment 1 with two exceptions. First, the SD of the noise was kept at a single intermediate value of 8 (between the High and Low conditions of Experiment 1). Second, tests of cue-outcome associations (Insight Tests) were presented directly after Control Test 1 and again after Control Test 2. In the Insight Tests participants were asked to predict the value of the outcome or one of the cues, given the values of the other variables. No feedback was presented in either Insight Test. Each test included 16 trials which were divided into 8 old and 8 new trials, each set containing 2 trials to predict each target (the outcome, positive, negative, and null cue). Old trials were randomly selected from the learning phase (for Controllers these were trials that they had generated themselves, for Predictors these were the same yoked learning trials). The eight new trials were designed prior to the experiment, so neither group had prior experience of them. The presentation of the 16 trials in each Insight Test was randomized for each participant. Performance on the insight tests was measured similarly to the prediction error scores in Experiment 1 as the absolute difference between predicted and expected value.

RESULTS AND DISCUSSION

Learning phase

The average control and prediction error scores by block (4 blocks of 10 trials each) in the learning phase are presented in **Figure 4**. For the Controllers, a one-way ANOVA on control scores showed a significant effect of Block, $F(3,42) = 21.93$, $p < 0.001$, partial $\eta^2 = 0.610$. Further t-test comparisons were conducted and revealed that control error scores were lower in Blocks 2, 3, and 4 as compare to Block 1 ($t = 6.67$, $p < 0.005$, $t = 5.90$, $p < 0.005$, $t = 11.76$, $p < 0.005$). A similar analysis on the prediction scores for Predictors showed no effect of Block, $F(3,42) = 0.36$, $p = 0.78$, partial $\eta^2 = 0.025$. Again, t-tests were conducted to examine if performance improved across-block. Analyses revealed that compared with Block 1, prediction error scores were lower in Block 2, 3, and 4 ($t = 2.95$, $p = 0.011$, $t = 3.88$, $p = 0.002$, $t = 4.18$, $p = 0.001$), no other comparisons were significant.

Test phase

The following analyses were based on the mean control scores by Block (block, 1, 2) and Test (Control Test 1 and 2) as presented in

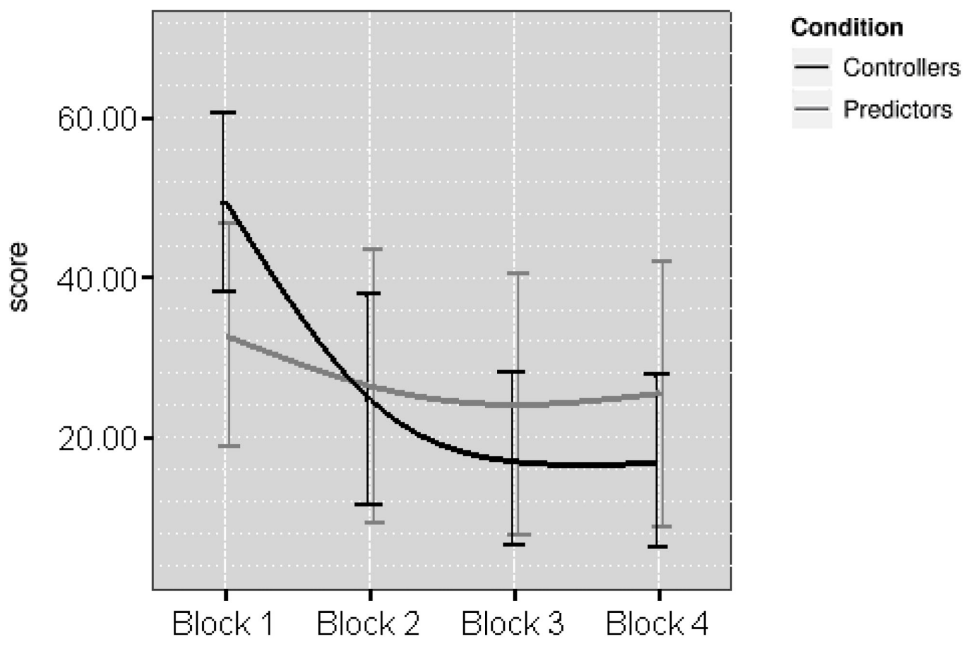


FIGURE 4 | Error scores (± SE) in the learning phase of Experiment 2. For Controllers, these are control error scores, and for Predictors, these are predictive error scores.

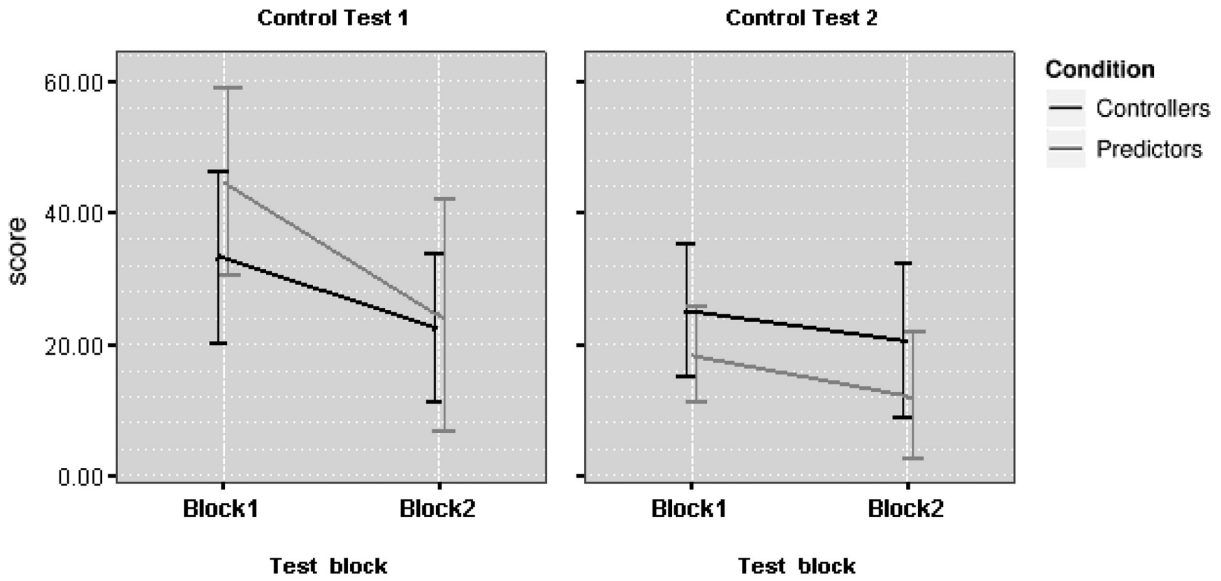


FIGURE 5 | Control error scores (± SE) of predictors and controllers in the test phase of Experiment 2.

Figure 5. A 2 × 2 × 2 ANOVA was conducted with Test and Block as within-subject factors, and Learning Mode (prediction, control) as between subject factor. As in Experiment 1, there was a main effect of Block, $F(1,28) = 26.01$, $p < 0.001$, partial $\eta^2 = 0.482$. There was a trend toward an effect of Test, $F(1,28) = 3.66$, $p = 0.066$, partial $\eta^2 = 0.115$; while failing to reach significance, this effect is consistent with Experiment 1. No other main effects were significant. As with Experiment 1, we used an across-blocks comparison to assess the benefit for prediction training prior to encountering the control task. As in Experiment 1, Predictors did not show an advantage when they first came to control the outcome compared to Controllers, $F(1,28) = 1.57$, $p = 0.221$, partial $\eta^2 = 0.057$. However, when comparing Controllers' performance between block 1 of the learning phase and block 1 of Control Test 1, their prior experience did facilitate control accuracy at test, $F(1,28) = 3.73$, $p = 0.064$, partial $\eta^2 = 0.152$.

Insights tests

Experiment 2 also included tests that enabled comparisons of groups on their ability to accurately predict the states of the system. In all analyses reported below, predictions for the null cue were not taken into account; as this cue had no effect on the outcome, its value was effectively unpredictable (it was only included in the set of questions not to alert participants to this fact). We computed error scores as the absolute difference between participants' predictions and optimal predictions; these are presented in **Figure 6**. A 2 × 2 × 3 × 2 ANOVA was conducted with Insight Test (first, second), Trial Type (old, new) and Target (positive cue, negative cue, outcome) as within-subject factors, and Learning Mode (prediction, control) as a between subject factor. This analysis showed only a significant effect of Trial Type, $F(1,28) = 14.51$, $p < 0.001$, partial $\eta^2 = 0.341$, indicating better performance for old than for new items. Other effects were not significant. In particular, there were no effects involving condition (all p's > 0.10). Comparing the error scores for predictions of the outcome to the prediction error scores of Predictors in the learning phase showed no difference for either Predictors, $t(14) = 1.06$, $p = 0.31$ (dependent samples), or Controllers, $t(28) = 0.17$, $p = 0.87$ (independent samples). Thus, Controllers were able to predict the outcome with a similar accuracy as Predictors who had learned to do so from the outset.

In sum, the findings from Experiment 2 suggest that control-based learning generates cue-outcome knowledge that is sufficiently flexible to enable transfer from control tasks to tests of cue-outcome knowledge. Moreover, the findings indicate that regardless of mode of learning, cue-outcome knowledge is accessible, to the extent that it can be reported in measures of task insight.

GENERAL DISCUSSION

Many accounts of DDM have referred to the importance of prediction in processes designed to control outcomes in a dynamic environment (Gibson et al., 1997; Osman, 2010a). In fact cognitive architectures such as ACT-R have been developed to model dynamic decision making, and variants of the framework also include components that assume a prediction-based process (Gonzales et al., 2003). In addition, general models of learning (reinforcement learning models/reward-based learning) claim that prediction-based and control-based decisions generate equivalent knowledge in dynamic learning environment (Schultz et al., 1997; Schultz, 2006). However, to date, there has been little empirical support for this claim in DDM research.

The aim of this investigation was to examine the accuracy of cue-outcome knowledge when applied to tests of control and prediction, by comparing prediction-based learning and control-based learning in the same dynamic environment. We developed a task in which Controllers and Predictors experienced identical cue-outcome information in a dynamic environment. Both

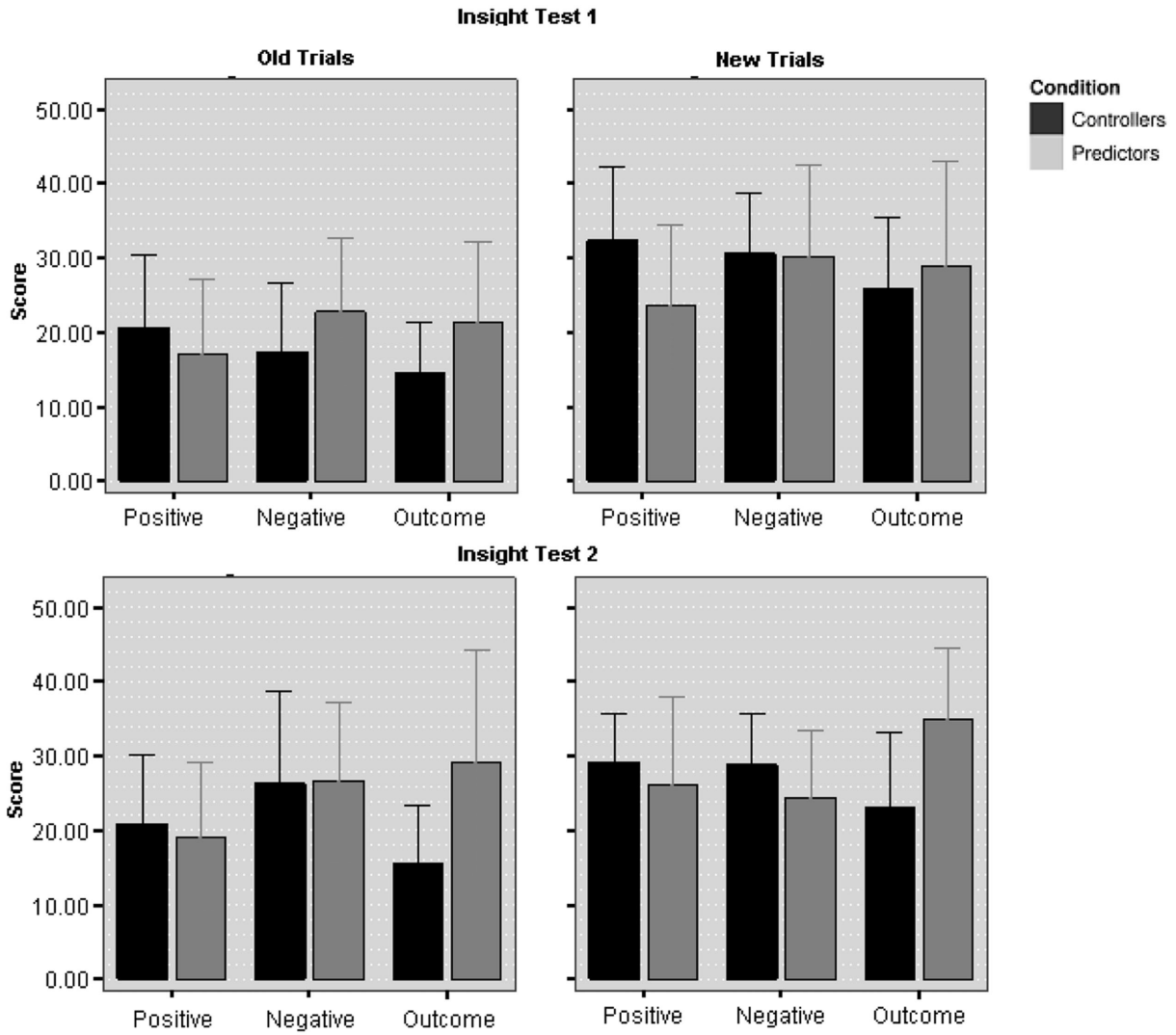


FIGURE 6 | Insight scores (± SE) for predictors and controllers for the insight tests of Experiment 2.

Predictors and Controllers were sensitive to the stability of the DDM environment, suggesting that there is greater accuracy in controlling and predicting outcomes in stable than in unstable environments. In both experiments we showed that predictive learning is sufficient to enable accurate control of an outcome in dynamic environments that varied according to stability. After an initial period of familiarization (approximately 10 trials) with the control test, Predictors' control accuracy was equivalent to that of Controllers. However, in both experiments Predictors' experience during learning did not generate cue-outcome knowledge that initially facilitated control performance. In contrast, we found that control-based experience during learning enabled successful transfer to predictive tests of cue-outcome associations as well as to tests of control. The discussion focuses on two questions: Did Predictors and Controllers learn in the same way? and How do prediction and control in the present study relate to conceptualizations of prediction and control in other decision making and learning domains?

DID PREDICTORS AND CONTROLLERS LEARN IN THE SAME WAY?

We propose that fundamentally prediction and control share basic properties when it comes to evaluating the outcome. In line with Gibson et al.'s (1997) action model, control-based learning involves an online comparison between the expected outcome and achieved outcome, as well as an online comparison between the achieved outcome and the target outcome. Both comparisons are used to form a judgment as to which action to take in order to reach and maintain the target outcome in dynamic tasks with a non-independent trial structure (Gibson et al., 1997; Gibson, 2007). We also propose that, in line with Gibson's action model, prediction-based learning involves an online comparison between an expected outcome and an achieved outcome, whereas control-based decisions are based on a comparison between the achieved and target outcome. Thus, prediction-based and control-based DDM are similar because both evaluate the achieved outcome according to their expected outcome. Given that in general predictive accuracy was equivalent for Predictors and Controllers,

the comparison between actual outcome and expected outcome appears to be sufficient to enable the acquisition of accurate cue-outcome knowledge. However, in both experiments the findings suggest that this comparison is not sufficient to facilitate initial transfer to tests of control, instead training that involves decisions based on comparisons between expected outcome and target outcome is needed. These comparisons are required for planning and our results suggest that such comparisons may need some time to develop, even when cue-outcome knowledge itself is already accurate. Learning cue-outcome relations involves updating expected outcomes in light of achieved outcomes. Learning to control involves updating the expected distance from the goal to the actually achieved distance from the goal.

We have suggested that the critical factor differentiating Predictors from Controllers is related to the type of comparison made (comparing expected to achieved outcomes, and comparing achieved outcomes to the goal outcome). However, there were other factors that may have contributed to the difference between Predictors and Controllers in initial tests of control. In the present study both types of learners could be differentiated in the following three ways. For instance, Controllers intervened on the cues directly, while Predictors could not. Some in the CDC research domain treat this as an important difference given that CDC tasks lend themselves to procedural rather than declarative forms of learning (Berry, 1991; Lee, 1995). In addition, Controllers received different feedback to Predictors. Controllers received feedback with respect to the deviation of their achieved outcome from the target outcome on a trial by trial basis. While this information was also presented to Predictors during learning, it was not directly relevant for their decisions. Only Predictors received direct feedback about the difference between their predicted (expected) outcome and the outcome actually achieved. However, we would argue that Controllers also formed expectations of the outcome value based on their cue-interventions and re-evaluated these predictions in light of to the actual achieved outcome. Finally, Controllers' actions had direct effects on the outcome in the actual DDM environment, whereas for Predictors an action (prediction) had no consequence on the outcome in the DDM environment. It may be the case that the advantage for Controllers stems from a more thorough evaluation of their decisions because errors not only affect the current outcome, but also future actions and outcomes. For instance, by making a poor cue-intervention which increases the deviation between achieved outcome and target outcome, future actions will have to compensate, and several more actions may be needed to reach the goal. Although Predictors may feel intrinsically motivated to make accurate predictions, and so an error would generate some personal dissonance, predictions do not require the type of planning behaviors needed in control. Further investigation is needed to examine these potential factors, for instance by varying feedback (e.g., positive, negative feedback) and reward (e.g., incentivizing accurate decisions).

HOW DOES PREDICTION AND CONTROL IN THE PRESENT STUDY RELATE TO CONCEPTUALIZATIONS OF PREDICTION AND CONTROL IN OTHER DECISION MAKING AND LEARNING DOMAINS?

MCPL vs. CDC tasks

With respect to the CDC research community the paradigm we used to examine prediction and control relates directly to other dynamic decision making tasks (Berry and Broadbent, 1987; Burns and Vollmeyer, 2002; Gonzales et al., 2003; Osman, 2008a). Crucially, the findings are consistent with the claim that Controllers base their decisions on predictions (Burns and Vollmeyer, 2002; Gibson et al., 1997; Goode and Beckmann, 2010; Osman, 2010a,b; Sun et al., 2001). In the present study the task environment in which prediction was examined involved cumulative changes across trials, whereas in typical MCPL tasks there is no dependency between trials. If prediction and control are determined by the type of environment in which they are tested, it may be that this qualitative difference between MCPL and CDC tasks does not allow one to draw strong conclusions about the similarities between prediction and control, or to generalize our findings to MCPL research. However, if MCPL tasks were designed with a dependent trial structure then our findings suggest that in stable and unstable dynamic environments, cue-outcome knowledge via prediction does not prevent accurate control-based decisions, although it does not facilitate initial transfer of knowledge to control tasks. Further research is required to examine the extent to which these findings generalize to task environments that are traditionally studied in MCPL tasks. If the trial structure makes little difference, then based on the findings from this experiment, we would still predict that overall cue-outcome knowledge for Controllers over Predictors should be equivalent.

Observation vs. intervention

Prediction- and control-based learning can also be viewed within the context of causal reasoning research which contrasts learning via intervention and observation (Meder et al., 2008; Hagmayer et al., 2010). Prediction and observation are both indirect ways of learning about an environment, whereas control and intervention involve direct interaction with the environment while learning (Lagnado and Sloman, 2004; Osman, 2010b). The distinction between these forms of interaction with the environment has been made using formal models (Spirtes et al., 1993; Pearl, 2000) which capture probabilistic dependencies in a given set of data as well as their relationship to the causal structures that could have generated the data. These models provide a strong theoretical basis for arguing that intervention is a crucial component in the acquisition of causal structures.

Studies that examine causal structure learning often require participants to either infer the causal structure from observing causes and effects, or from active learning by manipulating a candidate cause and observing the effects that follow. The evidence suggests that causal knowledge is more accurate when making interventions on causes than when observing causes and their effects (Steyvers et al., 2003; Lagnado and Sloman, 2004). Comparisons between observation and intervention have typically been conducted in static environments (for an exception see Hagmayer et al., 2010), and observation-based learning involves tracking causes and their effects as well as predicting changes in effects given specific causes. Though not commonly referred to, there is a close association between causal reasoning and DDM according to the ways in which prediction (observation) and control (intervention) generate cue-outcome knowledge. Funke and Müller (1988) found that while control performance was marginally impaired when learning was observation-based, causal knowledge of the environment was generally not influenced by mode of learning,

and accuracy of casual knowledge predicted control performance. Moreover, recent work that has explored different types of intervention suggests that it is not always the case that intervention is superior to observation. There is converging evidence that atomic interventions (i.e., ones in which acting on the environment will clearly reveal some aspect of the causal structure) lead to superior casual knowledge as compared with uncertain interventions (i.e., ones in which actions may not fix the state of a variable quite so clearly; e.g., Hagmayer et al., 2010; Meder et al., 2008). While interventions are generally helpful in uncovering casual structures, there is evidence to suggest that observation-based learning can generate more accurate causal knowledge than uncertain interventions for particular environments (e.g., Meder et al., 2008). Viewed from this perspective, the findings from the present study suggest that the accuracy of casual knowledge is not impaired when learning is observation-based rather than intervention-based, assuming that parallels can be drawn between observation and prediction. However, again, to forge links between CDC research and causal learning, research is needed to explore the extent to which the types of interventions involved in controlling a dynamic environment are similar to the kinds of interventions used to uncover causal structures, where the latter are typically studied in static environments.

Model-free vs. model-based learning

It is worth examining the possible relationship between prediction and control in the present study with a distinction made in machine learning concerning model-free and model-based reinforcement learning systems (e.g., Sutton and Barto, 1998; Dayan and Daw, 2008; Gläscher et al., 2010). Both model-free and model-based learning involve predictions about the expected reward from future actions, but the mechanism by which this is acquired is different. Model-based learning involves constructing a model of the task environment, which informs the agent about the expected consequences of actions taken in particular situations. Actions can move the environment from one state to another and the agent's aim is to move the environment to the state with the highest reward. Thus, the agent's learning involves the full sequence from action to state to reward whereas in model-free learning the agent learns action-reward pairings directly. Model-based methods are more flexible and can more easily accommodate changes in the state-reward pairings. Adaptation to changes in the reward structure is much slower with model-free methods. In our study Predictors were encouraged to learn the cue-outcome associations before being tested on their ability to decide upon actions to move the system to a desired state. In this way, Predictors could be described as model-based learners. The Controllers on the other hand were not explicitly encouraged (though not discouraged) to learn about the structure of the system. But given their performance in tests of prediction, we would speculate that their learning was also model-based. While the initial differences in control performance imply that Predictors and Controllers may have been using different strategies, further research is needed to explore the association between learning strategies (model-free vs. model-based) and modes of learning (prediction vs. control).

CONCLUSION

While prediction is crucial to control, and both tend to be described as complementary processes, in this article we considered the question whether learning to control a dynamic environment involves decisions based on prediction, and whether learning via prediction is sufficient to enable accurate control. The evidence from two experiments suggests that accurate cue-outcome knowledge is gained via prediction and control, and in general both forms of learning enable transfer to tests of prediction and control. However, Predictors required some familiarization with control whereas Controllers were able to transfer their knowledge to tests of prediction without familiarization to the tests. We propose that Predictors and Controllers evaluated the outcome according to the discrepancy between expected and actual outcomes, whereas Controllers also evaluate the outcome with respect to an intended goal. While there are processes shared by prediction and control, the critical difference between them appears enough to generate more flexible cue-outcome knowledge for Controllers than Predictors.

ACKNOWLEDGMENTS

The support of the ESRC Research Centre for Economic Learning and Human Evolution is gratefully acknowledged. Preparation for this research project was supported by the Economic and Social Research Council, ESRC grant RES-062-23-1511 (Maarten Speekenbrink), and the Engineering and Physical Sciences Research Council, EPSRC grant – EP/F069421/1 (Magda Osman).

REFERENCES

Bandura, A., and Locke, E. A. (2003). Negative self-efficacy and goal effects revisited. *J. Appl. Psychol.* 88, 87–99.

Berry, D. (1991). The role of action in implicit learning. *Q. J. Exp. Psychol.* 43, 881–906.

Berry, D., and Broadbent, D. E. (1984). On the relationship between task performance and associated verbalizable knowledge. *Q. J. Exp. Psychol.* 36, 209–231.

Berry, D., and Broadbent, D. E. (1987). The combination of implicit and explicit knowledge in task control. *Psychol. Res.* 49, 7–15.

Berry, D. C., and Broadbent, D. E. (1988). Interactive tasks and the implicit-explicit distinction. *Br. J. Psychol.* 79, 251–272.

Björkman, M. (1971). Policy formation as a function of feedback in a non-metric CPL-task. *Umeå Psychol. Reports*, 49.

Brehmer, B. (1992). Dynamic decision making: human control of complex systems. *Acta Psychol. (Amst.)* 81, 211–241.

Broadbent, D. E. (1977). Levels, hierarchies and the locus of control. *Q. J. Exp. Psychol.* 32, 109–118.

Brown, S., and Steyvers, M. (2005). The dynamics of experimentally induced criterion shifts. *J. Exp. Psychol. Learn. Mem. Cogn.* 31, 587–599.

Bryson, A. E., and Ho, Y. C. (1975). *Applied Optimal Control.* New York: Wiley.

Buchner, A. (1995). "Basic topics and approaches to the study of complex problem solving," in *Complex Problem Solving: The European Perspective*, eds P. A. Frensch and J. Funke (Hillsdale: Lawrence Erlbaum), 27–63.

Burns, B. D., and Vollmeyer, R. (2002). Goal specificity effects on hypothesis testing in problem solving. *Q. J. Exp. Psychol.* 55, 241–261.

Castellan, N. J. Jr. (1973). Multiple-cue probability learning with irrelevant cues. *Organ. Behav. Hum. Perform.* 9, 16–29.

Castellan, N. J. Jr., and Edgell, S. E. (1973). An hypothesis generation model for judgment in non-metric multiple-cue probability learning. *J. Math. Psychol.* 10, 204–222.

Cohen, M. S., Freeman, J. T., and Wolf, S. (1996). Meta-cognition in time stressed decision making: recognizing, critiquing and correcting. *Hum. Factors* 38, 206–219.

Dayan, P., and Daw, N. (2008). Connections between computational and neurobiological perspectives on decision making. *Cogn. Affect. Behav. Neurosci.* 4, 429–453.

Dienes, Z., and Fahey, R. (1995). Role of specific instances in controlling

a dynamic system. *J. Exp. Psychol. Learn. Mem. Cogn.* 21, 848–862.
Dienes, Z., and Fahey, R. (1998). The role of implicit memory in controlling a dynamic system. *Q. J. Exp. Psychol.* 51, 593–614.
Dörner, D. (1975). Wie menschen eine welt verbessern wollten und sie dabei zerstörten [How people wanted to improve the world]. *Bild Wiss.* 12, 48–53.
Edgell, S. E. (1974). *Configural Information Processing in Decision Making.* Indiana Mathematical Psychology Program, Report No. 74–4.
Enkvist, T., Newell, B., Juslin, P., and Olsson, H. (2006). On the role of causal intervention in multiple-cue judgment: positive and negative effects on learning. *J. Exp. Psychol. Learn. Mem. Cogn.* 32, 163–179.
Funke, J. (1992). Dealing with dynamic systems: research strategy, diagnostic approach and experimental results. *German J. Psychol.* 16, 24–43.
Funke, J. (1993). "Microworlds based on linear equation systems: a new approach to complex problem solving and experimental results," in *The Cognitive Psychology of Knowledge*, eds G. Strube and K.-F. Wender (Amsterdam: Elsevier), 313–330.
Funke, J., and Müller, H. (1988). Eingreifen und prognostizieren als determinanten von systemidentifikation und systemsteuerung [Intervention and prediction as determinants of system identfication and system control]. *Sprache Kogn.* 7, 176–186.
Gibson, F., Fichman, M., and Plaut, D. C. (1997). Learning in dynamic decision tasks: computational model and empirical evidence. *Organ. Behav. Hum. Perform.* 71, 1–35.
Gibson, F. P. (2007). Learning and transfer in dynamic decision environments. *Comput. Math. Organ. Theory* 13, 39–61.
Gläscher, 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.
Gonzales, C., Lerch, J. F., and Lebiere, C. (2003). Instance-based learning in dynamic decision making. *Cogn. Sci.* 27, 591–635.
Goode, N., and Beckmann, J. (2010). You need to know: there is a causal relationship between structural knowledge and control performance in complex problem solving tasks. *Intelligence* 38, 345–352.
Hagmayer, Y, Meder, B, Osman, M., Mangold, S., and Lagnado, D. (2010). Spontaneous causal learning while controlling a dynamic system. *Open Psychol. J.* 3, 145–162.
Hammond, K. R., Summers, D. A., and Deane, D. H. (1973). Negative effects of outcome feedback in multiple-cue probability learning. *Organ. Behav. Hum. Perform.* 9, 30–34.
Holzworth, R. J., and Doherty, M. E. (1976). Feedback effects in a metric multiple-cue probability learning task. *Bull. Psychon. Soc.* 8, 1–3.
Klein, G. (1997). Developing expertise in decision making. *Think. Reason.* 3, 337–352.
Knowlton, B. J., Mangels, J. A., and Squire, L. R. (1996). A neostriatal habit learning system in humans. *Science* 273, 1399–1402.
Knowlton, B. J., Squires, L. R., and Gluck, M. A. (1994). Probabilistic category learning in amnesia. *Learn. Mem.* 1, 106–120.
Körding, K., and Wolpert, D. (2006). Bayesian decision theory in sensorimotor control. *Trends Cogn. Sci.* 7, 319–326.
Lagnado, D., and Sloman, S. A. (2004). The advantage of timely intervention. *J. Exp. Psychol. Learn. Mem. Cogn.* 30, 856–876.
Lee, Y. (1995). Effects of learning contexts on implicit and explicit learning. *Mem. Cogn.* 23, 723–734.
Leigh, J. R. (1992). *Control Theory a Guided Tour.* London: Peter Peregrinus.
Lipshitz, R., Klein, G., Orasanu, J., and Salas, E. (2001). Taking stock of naturalistic decision making. *J. Behav. Decis. Mak.* 14, 332–351.
Lurie, N. H., and Swaminathan, J. M. (2009). Is timely information always better? The effect of feedback frequency on decision making. *Organ. Behav. Hum. Decis. Process* 108, 315–329.
Meder, B., Hagmayer, Y., and Waldmann, M. R. (2008). Inferring interventional predictions from observational learning data. *Psychon. Bull. Rev.* 15, 75–80.
Muchinsky, P. M., and Dudycha, A. L. (1975). Human inference behavior in abstract and meaningful environments. *Organ. Behav. Hum. Perform.* 13, 377–391.
Osman, M. (2008a). Evidence for positive transfer and negative transfer/anti-learning of problem solving skills. *J. Exp. Psychol. Gen.* 137, 97–115.
Osman, M. (2008b). Observation can be as effective as action in problem solving. *Cogn. Sci.* 32, 162–183.
Osman, M. (2008c). Seeing is as good as doing. *J. Problem Solving* 2, 1.
Osman, M. (2010a). Controlling uncertainty: a review of human behavior in complex dynamic environments. *Psychol. Bull.* 136, 65–86.
Osman, M. (2010b). *Controlling Uncertainty: Decision Making and Learning in Complex Worlds.* Oxford: Wiley-Blackwell.
Osman, M. (2011). "Monitoring, control and the illusion of control," in *Self Regulation*, eds M. Fitzgerald and T. Jane, Nova Publishers.
Osman, M., and Speekenbrink, M. (2011). Information sampling and strategy development in complex dynamic control environments. *Cogn. Syst. Res.* 355–364.
Osman, M., Wilkinson, L., Beigi, M., Parvez, C., and Jahanshahi, M. (2008). The striatum and learning to control a complex system? *Neuropsychologia* 46, 2355–2363.
Pearl, J. (2000). *Causality: Models, Reasoning, and Inference.* New York: Cambridge University Press.
Poldrack, R., Clark, J., Paré-Blagoev, E. J., Shahomy, D., Creso Moyano, J., Myers, C., and Gluck, M. (2001). Interactive memory systems in the human memory. *Nature* 414, 546–550.
Price, A. (2009). Distinguishing the contributions of implicit and explicit processes to performance on the weather prediction task. *Mem. Cogn.* 37, 210–222.
Rothstein, H. G. (1986). The effects of time pressure on judgment in multiple cue probability learning. *Organ. Behav. Hum. Decis. Process* 37, 83–92.
Sanderson, P. M. (1989). Verbalizable knowledge and skilled task performance: association, dissociation, and mental models. *J. Exp. Psychol. Learn. Mem. Cogn.* 15, 729–747.
Schultz, W. (2006). Behavioral theories and the neurophysiology of reward. *Annu. Rev. Psychol.* 57, 87–115.
Schultz, W., Dayan, P., and Montague, R. (1997). A neural substrate of prediction and reward. *Science* 275, 1593–1599.
Slovic, P., Rorer, L. G., and Hoffman, P. J. (1971). Analyzing use of diagnostic signs. *Invest. Radiol.* 6, 18–26.
Speekenbrink, M., Channon, S., and Shanks, D. R. (2008). Learning strategies in amnesia. *Neurosci. Biobehav. Rev.* 32, 292–310.
Speekenbrink, M., and Shanks, D. R. (2010). Learning in a changing environment. *J. Exp. Psychol. Gen.* 139, 266–298.
Spirtes, P., Glymour, C., and Scheines, P. (1993). *Causation, Prediction, and Search.* New York: Springer Verlag.
Steyvers, M., Tenenbaum, J. B., Wagenmakers, E. J., and Blum, B. (2003). Inferring causal networks from observations and interventions. *Cogn. Sci.* 27, 453–489.
Sun, R., Merrill, E., and Peterson, T. (2001). From implicit skills to explicit knowledge: a bottom-up model of skill learning. *Cogn. Sci.* 25, 203–244.
Sutton, R. S., and Barto, A. G. (1998). *Reinforcement Learning.* Cambridge: MIT press.
Vancouver, J. B., and Putka, D. J. (2000). Analyzing goal-striving processes and a test of the generalizability of perceptual control theory. *Organ. Behav. Hum. Decis. Process* 82, 334–362.
Vlek, C. A. J., and van der Heijden, L. H. C. (1970). Aspects of suboptimality in a mutidimensional probabilistic information processing task. *Acta Psychol.* 34, 300–310.
Vollmeyer, R., Burns, B. D., and Holyoak, K. J. (1996). The impact of goal specificity and systematicity of strategies on the acquisition of problem structure. *Cogn. Sci.* 20, 75–100.

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 November 2011; paper pending published: 22 December 2011; accepted: 21 February 2012; published online: 13 March 2012.
Citation: Osman M and Speekenbrink M (2012) Prediction and control in a dynamic environment. Front. Psychology ***3****:68. doi: 10.3389/fpsyg.2012.00068*
This article was submitted to Frontiers in Cognitive Science, a specialty of Frontiers in Psychology.

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 05 October 2012
doi: 10.3389/fpsyg.2012.00381

The influence of initial beliefs on judgments of probability

Erica C. Yu* and David A. Lagnado

Cognitive, Perceptual and Brain Sciences, Division of Psychology and Language Sciences, University College London, London, UK

Edited by:
Konstantinos Tsetsos, Oxford University, UK

Reviewed by:
Ingmar Visser, Universiteit van Amsterdam, Netherlands
Magda Osman, Queen Mary Univesity of London, UK

****Correspondence:***
Erica C. Yu, Department of Psychology, University of Maryland, College Park, MD 20742, USA.
e-mail: erica.yu@umd.edu

This study aims to investigate whether experimentally induced prior beliefs affect processing of evidence including the updating of beliefs under uncertainty about the unknown probabilities of outcomes and the structural, outcome-generating nature of the environment. Participants played a gambling task in the form of computer-simulated slot machines and were given information about the slot machines' possible outcomes without their associated probabilities. One group was induced with a prior belief about the outcome space that matched the space of actual outcomes to be sampled; the other group was induced with a skewed prior belief that included the actual outcomes and also fictional higher outcomes. In reality, however, all participants sampled evidence from the same underlying outcome distribution, regardless of priors given. Before and during sampling, participants expressed their beliefs about the outcome distribution (values and probabilities). Evaluation of those subjective probability distributions suggests that all participants' judgments converged toward the observed outcome distribution. However, despite observing no supporting evidence for fictional outcomes, a significant proportion of participants in the skewed priors condition expected them in the future. A probe of the participants' understanding of the underlying outcome-generating processes indicated that participants' judgments were based on the information given in the induced priors and consequently, a significant proportion of participants in the skewed condition believed the slot machines were not games of chance while participants in the control condition believed the machines generated outcomes at random. Beyond Bayesian or heuristic belief updating, priors not only contribute to belief revision but also affect one's deeper understanding of the environment.

Keywords: model-based learning, belief revision, priors, gambling, probability judgment

INTRODUCTION

When a decision making process occurs over a temporally extended interval, new evidence may accumulate over time that warrants the updating of initial beliefs. Particularly in novel environments where initial beliefs may be based on only scant information, how should we integrate our initial beliefs about the environment with subsequent observations?

Combining a prior description with experience is an everyday activity, as mundane as judging the likelihood that it will rain later in the afternoon, given a weather forecast and one's own observations from looking out the window. Consider, for example, slot machines. At most slot machines today, the only information known to players before the start of a game is a long-run payout percentage and a succinct payout table that lists which outcomes are associated with which combinations of symbols (or, as some might have you believe: which combinations of symbols *cause* which outcomes). This incomplete description of the environment lacks probability information for the listed outcomes. Players must repeatedly play the machines to learn about the missing probability distributions and learn which machines may be most valuable to play at. In other words, the slot machine is an inductive inference problem. Gambles are of particular interest to the study of inductive inference because they are so similar to everyday reasoning and yet very dissimilar at the same time. Due to the underlying randomness in many gambles (such as slot machines), people who fare well in everyday reasoning can fail in the face of gambles. Indeed, it is still a mystery why people can be so good at some gambles (such as poker) and yet so bad at others (roulette). How can we account for this?

In this paper, we investigate the impact of initial descriptive information about an environment of uncertainty on judgments about the environment's structure. Our study extends previous work on the topic by investigating the impact of induced priors as well as evidence on judgments and beliefs about structure. To accompany the slot machines in our full-feedback paradigm, we provide payout tables, thereby controlling the initial beliefs of which outcomes are possible. However, we do not provide the probability distribution nor set one machine to be more profitable. Instead, we manipulate whether the evidence is congruent or incongruent with prior beliefs. How will participants update their prior beliefs given this evidence? To anticipate our results, we find that neither simple Bayesian nor heuristic-based accounts of belief revision can explain our findings; instead, a model-based framework is proposed for a descriptive account of decision making over time.

Prior knowledge is a central factor of decision making theories, across the spectrum from heuristic (gambler's fallacy: a prior regarding representativeness of outcomes; Tversky and Kahneman, 1971; Kahneman and Tversky, 1972; anchoring and adjustment: insufficiently adjusting estimates from a reference point;

Tversky and Kahneman, 1974) to rational (game theory: consistency of prior beliefs between players in a game; Harsanyi, 1967). Likewise, research has demonstrated that different assumptions for priors can lead to different posterior beliefs (Troutman and Shanteau, 1977; Koehler, 1993) as can prior outcomes affect subsequent judgments and decisions (Thaler and Johnson, 1990). Decision makers may maintain their initial hypotheses by dismissing disconfirming evidence (Klayman and Ha, 1987) or even inappropriately using disconfirming evidence to support initial hypotheses (Snyder and Swann, 1978; Doherty et al., 1979; Fischhoff and Beyth-Marom, 1983). Under uncertainty, prior beliefs can have a cascading effect on subsequent judgment and decision making that is deeper than mere belief-adjustment.

The model-based approach to understanding learning uniquely captures this relationship between prior and posterior beliefs. It derives psychological validity from theories of internal representations of external events or ideas. Well-known examples of internal representations in the literature include schemas, structures of knowledge that include concepts of components, attributes, and relationships between specific instances (Simon, 1957; Schank and Abelson, 1977; Bower, 1981; Pearson et al., 1984), and mental models, internal symbolic understanding of the external world (Johnson-Laird, 1983). Models in this sense are small-scale representations of reality, neglecting facts, and relationships that are outside one's scope of knowledge with cascading effects on reasoning. As in previous theories of decision making that use models, this approach includes representations of the states of the environment, actions, and rewards (as opposed to model-free learning, which assumes no representation of structure; Dayan and Niv, 2008); however, the present model-based approach focuses on *inference* in contrast to optimized decision making. A close and very interesting example of model-based inference can be found in Lopes (1976) though here, too, the emphasis is on how to optimize betting rather than how models influence understanding of the structure of the game. Although computational model-based learning considers the value of these representations in the construction of a model of the environment, the present perspective emphasizes representations of explanatory, causal, and goal-directed beliefs about the relationships between these components.

The top-down structure enables the agent to reason broadly and make inferences about classes and categories of events and relationships using prior knowledge as well as in a more detailed manner about the specifications of the current problem using data. The revision process, in which models are updated as new evidence is acquired, utilizes the structure of the models to exploit previous experience and knowledge, including information about the relationship between events such as data-generating processes, not just the events themselves (Sutton and Barto, 1998). The learner evaluates new information against his or her prior beliefs, much like Edwards (1968) and the belief-adjustment model of Hogarth and Einhorn (1992) in which new evidence is added or averaged with previous information using anchoring and adjustment. However, within the model-based framework, new evidence can be observed without requiring adjustment to the model. The overall process can be depicted as a feedback loop whereby the outcomes of the learner's actions are fed back into his or her beliefs about the problem. **Figure 1** illustrates this loop: the consequence of an outcome after an agent's action may be to update the policy directly or to feed back to the agent's model and be integrated with prior beliefs. For example, fluctuations in outcome within a range of expected possible outcomes, such as those in scenarios with random processes, need not change the learner's model of the environment (Yu and Dayan, 2005). A player at a slot machine may avoid a machine after experiencing losses or, depending on his model of the outcome-generating process, stay at the machine because a win is more likely to occur after a string of losses. A poker player may fold after observing another player raise or, depending on his model of the player, re-raise because he believes based on previous games with that player the raise to be a bluff. Because models of the environment are highly structured and include histories of previously held beliefs, new information can go beyond adjustments to cause qualitative and systemic changes in models.

Critically, as a consequence of these two dissociable components of model and belief revision, an individual may appear to behave irrationally while implementing rational inference. If the player at the slot machine holds a model of the game whereby a string of losses is certainly followed by a win, then he acts rationally when he continues to play despite losing. If the model of the problem is inaccurate, it may be the case that new evidence does not correct errors despite the correct implementation of the updating process. The individual may persist in believing inaccurate information.

The model-based learning framework is not commonly cited in psychological research on decision making. Researchers more often appeal to rational Bayesian accounts (Edwards et al., 1963; Steyvers et al., 2003; Tenenbaum et al., 2006). Bayes' theorem, a general mathematical rule commonly used for belief updating using evidence, results in a posterior probability expressing the degree of belief about the likelihood of a hypothesis being true after observing data. However, while Bayesian inference itself may be straightforward, the assumptions made about the hypothesis sets and approximation algorithms are less clear, and sometimes lacking in theoretical and empirical grounding (Jones and Love, 2011). The psychological implications as regards to the cognitive capacity required to consider all evidence even-handedly, generate exhaustive sets of hypotheses, and calculate likelihoods are out of

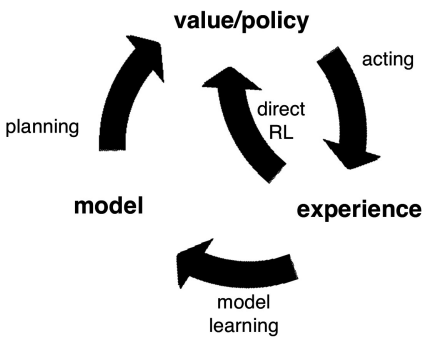


FIGURE 1 | The cycle of model revision over time (Sutton and Barto, 1998, p. 231).

reach for most people, including experts (Meehl, 1954; Fischhoff et al., 1978; Fischhoff and Beyth-Marom, 1983).

To study the integration of prior beliefs with evidence accumulated over time, we use a slot machine modification of the "computerized money machine" typically used in decision-from-experience tasks (e.g., Barron and Erev, 2003). Two slot machines each present a "button" to push that offer probabilistic payouts based on distributions unknown to the participant at the start. One difference to previous experiments is that, in the present task, the machines are identical. As we are primarily interested in how the descriptive information is combined with experience, we did not complicate the task further. A second critical difference is that the slot machine provides additional descriptive information to the participant in the form of a payout table that lists the space of potential outcomes. This will be discussed in greater detail below.

Previous work using this decision making over time paradigm has focused on finding differences between problems based on description and those based on experience (Barron and Erev, 2003; Hertwig and Erev, 2009; Ungemach et al., 2009). This paper focuses instead on an area slightly outside of that dialog: the combined effect of description and experience. Newell and Rakow (2007) discuss this research question with a binary prediction game using dice for which participants are given outcome probabilities and the opportunity to experience outcomes over time. When given only abstract, descriptive information about the outcomes and probabilities, participants were likely to choose the optimal strategy; when given both abstract information and experience, participants were more likely to sub-optimally probability match. Although this finding reveals the combined effect of descriptive and experiential information on risky choice, it does not provide a deeper insight into the cognitive processes participants may be using in such tasks. Strategy probes questioned participants only at the end of games and did not capture the reasons behind response choices. Critically, with dice stimuli that represent a definitively random outcome-generating process with known probabilities, questions of how participants learn about structure and probability distributions under uncertainty are outside the scope of that study.

To test the effects of prior knowledge on learning, we manipulate the content of the prior. Some participants were induced with the "control" condition prior, which accurately reflected the abstract outcome information for the task to be played: the slot machine payout table included all valid outcomes only. In contrast, other participants were induced with a skewed prior, which reflected the same information as controls with the addition of fictional higher outcomes. All participants played the same machines and observed effectively the same evidence. What effect would the different priors have on beliefs about structure and probability as new evidence was attained?

MATERIALS AND METHODS

PARTICIPANTS

Fifty-three participants were recruited to participate in paid studies on gambling from the University College London Psychology Department Subject Pool, a popular online UK notice board, and local newspapers. All individuals gave their informed consent to participate in the experiment, as approved by the Department of Psychology at University College London.

DESIGN

The design comprised one within-subjects factor over time with three levels (before the start of the game, after 30 trials, after 80 trials) and one between-subjects factor with two levels (congruency of payout table and outcomes, "skewed" in congruency). Participants were randomly assigned to conditions, resulting in a final sample of 27 participants in the congruent payout table group and 26 participants in the skewed payout table group. Participants played 80 mandatory pulls on the machines and answered questions before the start of play, after 30 arm pulls, and at the end after 80 arm pulls; machine choice and pace were up to participants but the total number of pulls played was fixed.

All participants played the same task with equivalent slot machines that sampled from the same underlying distribution of outcomes: 0, 2, 3, 4, 5, and 10. Participants in the control condition were shown a payout table before the start of the game that displayed the outcome values of the game and their associated reel symbol combinations (i.e., congruent with the observed evidence: 0, 2, 3, 4, 5, and 10). The other group of participants was shown the same payout table, but now with fictional higher-value outcomes (i.e., incongruent with the observed evidence: 0, 2, 3, 4, 5, 10, 15, 20, 25, and 100). Despite the different payout tables, both groups played the same task and observed random draws from the same underlying distribution. No participants were shown the probabilities associated with the outcomes. The skewed table suggested an initial belief that overlapped with the control participants' but also included fictional higher-value outcomes that would not be observed during the experiment[1].

Participants were randomly assigned to either the experimental skewed or the control condition and remained unaware of any alternative task specifications. In each condition, participants were informed that their compensation would depend on the bank's value at the end of the task.

MATERIALS

The slot machines required a five pence stake for each play, which was taken from the £3.00 bank endowed by the experimenter to the participant at the start of the task. The machines used a random process to select outcomes from a fixed distribution: an outcome of 0, 2, 3, 4, or 5 pence with 17.4% probability each or 10 pence with 13.0%. The expected value of a play at any machine was 3.9 pence, at a loss of 1.1 pence given the cost to play.

The experiment interface used fruit graphics, which are highly associated with slot machines imagery, and animations such as spinning reels and moving levers to simulate the appearance of real-world slot machines. The size of the screen display allowed the participant to see up to three symbols on each reel depending on the reel position. The machines had a single payout line (combinations of symbols must fall on the payout line to qualify for winnings) through the middle; symbols above and below (near

[1] Although the higher-value outcomes were fictional in this instance, the distribution is not dissimilar from real-world slot machines for which the desirable high-value outcomes are generated at infinitesimally low probabilities.

misses) were also viewable but randomized. When a participant clicked on a machine to play it, the reels appeared to spin, and then slow to a rest after 3 s, displaying the final screen of fruit symbol graphics that matched the outcome received on that trial. The final screen and numerical outcome value (e.g., four pence) was shown on the screen for 1.5 s. The screen then returned to the initial state and participants could click on the machine they wished to play next. No bank or cumulative total information was shown to the participant at any time during the task except after completion.

When asked to illustrate their beliefs about what outcomes the machines paid out and how likely those outcomes were, participants were given blank pie chart templates and pens. Each template sheet had one blank circle for the machine on the left and another identical one for the machine on the right, each with an indicator for the center of circle to aid in drawing. Most participants readily understood this instruction but all watched the experimenter create an example pie chart and had the opportunity to ask questions. This method was chosen both for its familiarity for participants and for its convenience for eliciting comparable judgments between groups of the size of the outcome space. By providing blank response templates, this design requires participants to generate the outcome space and associated probabilities without being prompted by the experimenter for different outcome values. Alternative methods of elicitation of probabilities, such as prompting responses one outcome at a time, or listing the full space of outcome values for the participant, may have defined the outcome space for the participant and consequently rendered their subjective probability distributions invalid. Although phenomena such as sub additivity cannot be investigated in this paradigm, the benefits for the relevant hypotheses being tested are greater than this limitation.

PROCEDURE

At the start of the task and at two times during the task, participants were asked to answer questions. Before beginning machine play but after being shown the playing environment including the machines and payout tables, participants were asked to give their best guess as to the hidden probabilities associated with the outcome distribution. After completing a pie chart for the machine on the left of the screen and another for the machine on the right, participants then began to play the machines. After 30 trials, the program automatically stopped and prompted participants to respond to questions. The experimenter presented pen and clean paper and asked the participant to again illustrate the different payouts they believed the machines generally paid out, and how likely those payouts were. After 80 trials, the prompts were repeated. The timing of the prompts after 30 and 80 trials was not known to the participants. The payout tables that displayed the outcome space were visible during the first response time before play had begun but were not visible during the latter two judgment collections; participants completed these pie charts from memory.

After completing the 80 trials, participants responded to a forced choice question about the machines' outcome-generating processes. Participants were asked which statement most closely matched their belief: "playing required skill to avoid bad luck or bad streaks at machines" or "it did not matter what I did or how I played."

RESULTS

In this experiment, a mixed design compared two groups' changing beliefs over time about a hidden outcome distribution, measured by probability judgments and responses to direct questions after completion of the task. Neither group was compared to the true underlying distribution because an infinite number of processes might have produced the sequences observed by the participants; only summary statistics (expected value) of the judged probability distribution were compared to the observed distribution. The experience of the two groups varied only on the range of outcomes listed in the payout tables shown during the task; all participants experienced slightly different sequences of outcomes due to random sampling but no significant differences in the final sum of outcomes received were found between groups, $t(52) = 1.23$, ns.

ESTIMATES OF MEAN PAYOUT

Participants in both experimental groups made probability distribution judgments before play began, after 30 trials, and after 80 trials. At these collection times, participants were asked to report what payouts they thought the machine paid out in general, and how likely were those payouts. An example of a participant's response in pie chart form is shown in **Figure 2**.

The first analysis of these data is of the raw mean estimates calculated using participants' pie charts, as shown in **Figure 3**. By measuring each pie chart segment, we were able to assess how likely the participant believed each outcome to be and therefore the participant's subjective expected value for a play of the slot machine. In the example shown in **Figure 2**, the participant expresses a belief that the probability of an outcome of 0 pence is 50% and the subjective expected value of a play of the machine is 7.74 pence. Estimates of the left and right machines were averaged (no differences between expected value estimates given for left and right machines (paired samples t tests, all $ps > 0.24$), resulting in a single data point for each subject at each time point. Analysis of these data finds that there are significant differences between the two experimental groups [$F(1,49) = 29.95$, $p < 0.001$, $\eta_p^2 = 0.38$], demonstrating that, overall, the payout table with fictional higher-value outcomes led to higher valuations

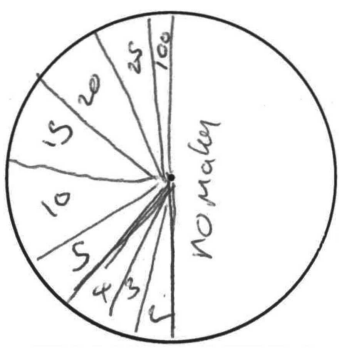


FIGURE 2 | Example of a participant's hand-drawn initial judgment of a machine's underlying probability distribution of outcomes. This participant indicated that the possible outcome space included 0, 2, 3, 4, 5, 10, 15, 20, 25, and 100 pence outcomes and that the most likely outcome would be of no matching symbols, or 0 pence.

of the slot machines. Significant effects were also found within groups over the three judgment times [$F(2,98) = 4.18$, $p < 0.02$, $\eta_p^2 = 0.08$]; *post hoc* Bonferroni-adjusted comparisons show the difference lies between the first and third judgment times only ($p < 0.01$; between first and second: $p < 0.92$ and between second and third: $p < 0.09$). In other words, valuations gradually decreased. The payout table and time interaction was also significant [$F(2,102) = 4.12$, $p < 0.02$, $\eta_p^2 = 0.08$] with follow-up tests showing that the payout tables with the fictional higher-value outcomes led to higher valuations compared to the control group at all judgment times (all ps < 0.02). These findings suggest that participants who were shown skewed payout tables consistently gave higher estimates of the expected value of a machine play than control participants, but showed a trend of converging toward observed values as a function of number of trials played. Splitting those participants who were shown skewed payout tables further into those who persisted in representing fictional higher-value outcomes throughout the duration of the game (shown in **Figure 3**) suggests that this difference is primarily driven by the representation of higher-value outcomes rather than misestimates of the probability of the true observed outcomes.

Remembering that each participant observed a different randomly generated sequence of outcomes, the next analysis illustrated in **Figure 4** examines each participant's pie estimates of the expected value given their unique observed sequence of outcomes to assess whether group differences in estimates are due to differences in observed sequences. Their observed values were compared to their reported estimates from collection times after 30 and 80 trials, with the left and right machine estimates averaged for each participant, to calculate an error measure; initial judgments are not included in this analysis because participants did not observe any outcomes before making initial judgments. These data confirm the analysis of the raw estimates: there are significant differences between the two experimental groups [$F(1,51) = 7.06$, $p < 0.01$, $\eta_p^2 = 0.12$] and within groups over the three judgment times [$F(1,51) = 6.45$, $p < 0.01$, $\eta_p^2 = 0.11$]. The interaction of the two factors is also significant [$F(1, 51) = 6.29$, $p < 0.02$, $\eta_p^2 = 0.11$].

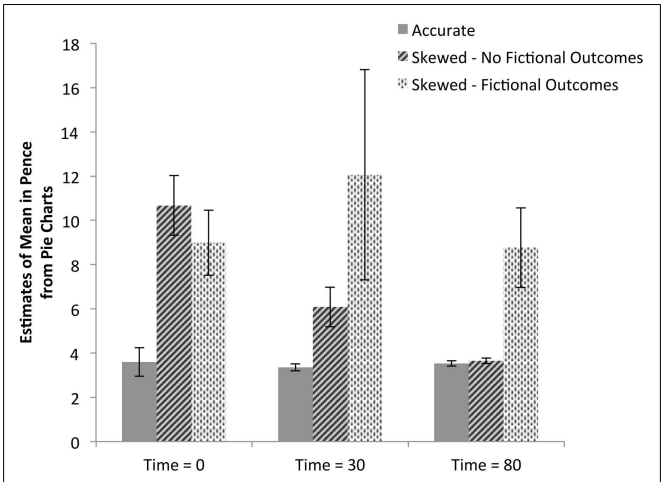


FIGURE 3 | Implied estimates of mean payout across time. Data from participants' pie charts were used to calculate each participants implied estimate of mean payout. Data shown in solid gray bars are from participants who were shown accurate payout tables including only valid outcomes. The data are split based on whether the participant persisted throughout the game in including the fictional higher-value outcomes of 15, 20, 25, and 100 pence; 8 of 26 participants given skewed payout tables included the fictional outcomes; no participants given accurate payout tables did so. Responses for the left and right machines at each judgment collection time were averaged resulting in one response per participant for each of the three judgment times.

ACCURACY OF PIE ESTIMATES

As shown in **Figure 4**, control participants who were shown an accurate payout table made precise estimates not significantly different from the observed data after 80 trials [$M = -0.09$, $SE = 0.13$; one-sample t-test against 0: $t(26) = 0.67$, ns] and after only 30 trials [$M = -0.10$, $SE = 0.13$; one-sample t-test against 0: $t(26) = 0.77$, ns]. This accuracy provides support for the validity of this method of subjective probability elicitation to capture sensible data.

REPRESENTATION OF FICTIONAL HIGHER-VALUE OUTCOMES

Although it is evident that the participants of the two groups perceive the expected value of each machine play differently, further analysis of the pie charts may explain this difference. Estimates of the means alone cannot distinguish overestimation of the likelihood of observed outcomes (e.g., believing the 5 or 10 pence outcome happen more frequently than the observed data suggest) from categorically representing higher-value outcomes with any degree of likelihood (e.g., believing the 100 pence outcome is possible with 1% probability). The pie charts show that the primary source of the overestimation comes from maintaining a belief in the likelihood of the fictional higher-value outcomes. After 30 trials, 61.54% of participants (16 of 26 participants) in the skewed payout table group continued to maintain the unsupported belief of at least one higher-value outcome while only 3.70% (1 of 27 participants) indicated the same in the control group ($p < 0.001$, Fisher's exact test). After even 80 trials, the difference in number of participants maintaining beliefs in higher-value outcomes remains significant (Skewed: 30.77% or 8 of 26 participants; Controls: 0%; $p < 0.01$, Fisher's exact test). This pattern shows that participants converged toward the observed data and no participants developed a skewed belief of unsupported fictional higher-value outcomes after having expressed a belief reflecting the observed outcome values only.

BELIEFS ABOUT THE NATURE OF THE OUTCOME-GENERATING PROCESS

A direct question probed participants for their beliefs about the nature of the underlying outcome-generating process. Although participants were probed only once after the completion of the task, this assessment enables us to infer, assuming no differences in beliefs about the outcome-generating processes of slot machines before beginning the task, whether the task environment changed participants' beliefs. It was hypothesized that participants who were shown the skewed payout table and expected to receive higher-value outcomes may believe that they are performing poorly on the task when they do not receive the expected outcomes. When asked whether the outcomes were generated by

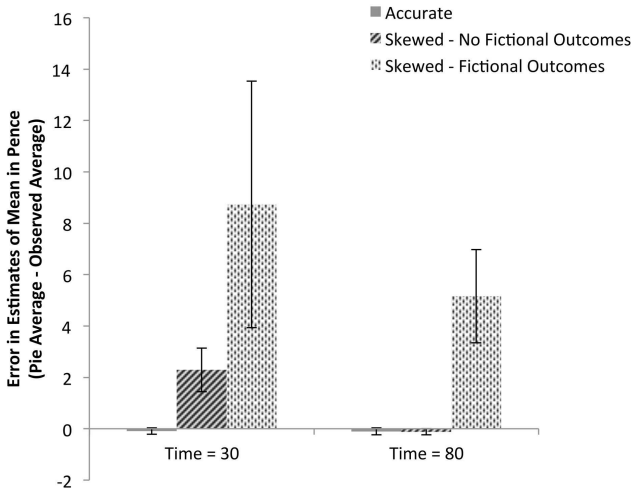


FIGURE 4 | Mean error in payout estimation based on condition and belief in fictional outcomes. Each participant's mean observed payout for each machine was subtracted from each participant's pie chart estimate of mean payout for that machine to calculate errors in estimation. The data are split based on whether the participant persisted throughout the game in including the fictional higher-value outcomes of 15, 20, 25, and 100 pence; 8 of 26 participants given skewed payout tables included the fictional outcomes; no participants given accurate payout tables did so. Values greater than zero indicate overestimation and values close to zero indicate accurate estimation. Errors for both the left and right machines were averaged resulting in one error measure per participant per judgment during machine play.

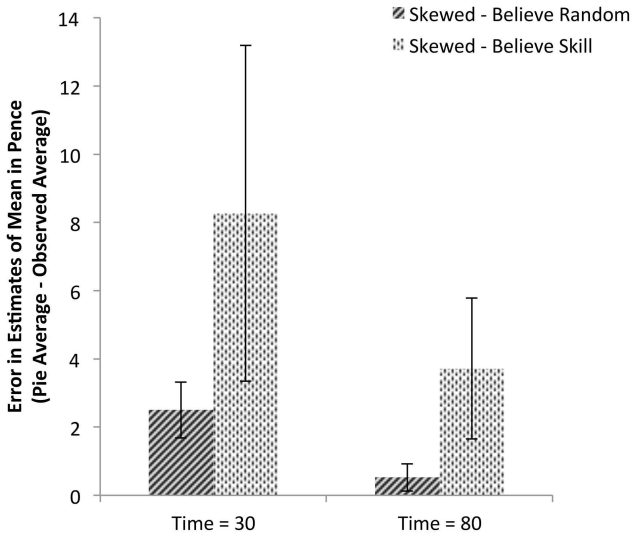


FIGURE 5 | Mean error in payout estimation based on structural beliefs. Mean payout estimate errors are shown for those participants given skewed payout tables, split based on whether the participant responded with the belief that the machine play was random or based on skill (8 of 26 participants); data from the control group of participants shown accurate payout tables is not included because only 1 of 27 participants expressed belief that machine play was based on skill. Values greater than zero indicate overestimation and values close to zero indicate accurate estimation. Errors for both the left and right machines were averaged resulting in one error measure per participant per judgment during machine play.

a random process (a choice between the two statements: "playing required skill to avoid bad luck or bad streaks at machines" or "it did not matter what I did or how I played"), only 3.84% (1 of 26 participants) in the control group indicated that skill was a significant factor while 30.77% (8 of 26 participants) indicated this from the group who viewed skewed payout tables. This analysis suggests that the different payout tables influenced the participants' beliefs about the outcome-generating process. This difference between groups within the Skewed condition is shown in **Figure 5**: those participants who believe skill to be involved in machine play estimate the machines to be significantly more valuable than the machines actually are. Indeed, the two are highly related: a logistic regression predicting belief type finds that those participants who categorically represent at least one fictional higher-value outcome are 8.00 times more likely to also believe the outcome-generating process is based on skill ($B = 2.08$, $p < 0.02$). Although these correlations cannot provide directional explanations for participants' responses, they support the hypothesis that beliefs about an underlying outcome distribution and associated underlying outcome-generating process may be related.

GENERAL DISCUSSION

In this study, we induced different initial beliefs about the task environment and observed how the accumulation of evidence changed those beliefs. The initial beliefs informed participants of the outcome space in the environment; some participants were given information congruent with what they would subsequently observe while others were given incongruent information (the description of the outcome space included fictional higher-value outcomes). Participants who predicted that the machines would produce fictional higher-value outcomes did so despite never observing their occurrence, and ultimately attributed the absence of those outcomes to (lack of) skill. From this experiment, there are two broad findings: participants integrated observed evidence with their initial beliefs for probability judgments but relied heavily on initial information to understand the structure of the environment. In other words, prior beliefs are more than just initial information; priors are the basis of internal models that affect beliefs about the structure of the data-generating processes.

When participants created pie charts from blank templates, they had to express not only the probabilities they thought were associated with outcomes but also the outcomes themselves. Despite the limitations of this study due to possible error in manual measurement of participants' responses, the pattern of categorical responses is clear. The pie charts showed that many participants who initially believed that higher-value outcomes would occur continued to believe so even after playing 30 and 80 trials and never seeing those outcomes occurring. Observed evidence alone cannot account for this. The results are contrary to most memory-based theories that would predict the non-observance of these outcomes resulting in their absence from the subjective probability distributions, including those based on availability

(Tversky and Kahneman, 1973). Similarly, model-free theories, in which judgment is based solely on experience, are also unable to account for this. Alternatives such as Bayesian accounts are more successful in explaining the role of prior beliefs in these judgments. Bayesian belief revision corresponds with the convergence of judgments over time and the variable weight on prior beliefs relative to observed evidence. From this perspective, participants' skewed judgment in persistently predicting the fictional higher-value outcomes despite their non-occurrence can be seen as a rational belief: not enough evidence had been presented to outweigh the initial belief in the higher-value outcomes.

Up to this point in the analysis, Bayesian accounts are consistent with the behavior we found in this task. However, basic Bayesian theories do not specify the underlying cognitive process and have no explanation for the differences in the beliefs of randomness in the outcome-generating process required to explain the critical results. The evident difference in beliefs about how the outcomes are generated dictates that more is needed to explain the overall results – beliefs about the *structure of the environment* is the critical element. The model-based approach, which relies on the player's representation and understanding of the structure of the game, tolerates the updating of probabilities based on evidence in parallel with the persistent representation of higher-value outcomes on the basis of no observed evidence at all. Although we probed participants only once for their attitudes regarding the data-generating process, those responses considered along with the probability judgments illustrate that initial beliefs have qualitative effects on subsequent judgments of probability and understanding of the structure of the environment. Future research should aim to capture a richer understanding of belief revision and decision making strategies by using additional measures and analysis methods, such as measuring probability judgments more often or such methods as those used in Jansen et al. (2012). This experiment demonstrates that theories of decision making must account for beliefs about the underlying data-generating process. In this broad sense, both heuristic and Bayesian theories that either approximate or explicitly include such causal beliefs may adequately describe the behavior in this task (Krynski and Tenenbaum, 2007). Future research should seek to refine and enrich our understanding of model structure.

Gambles are effectively inference problems in which the player must learn about hidden underlying outcome distributions and outcome-generating processes by generalizing from samples. Slot machines are a paradigm in which decision makers must learn over time about the value of their actions. While a die has six sides each with equal likelihood of landing face-up, a slot machine has an unknown number of symbols in unknown locations and ratios on each reel and an unknown algorithm determining the outcomes. Ultimately, slot machines pose the highest risk among games of chance because the unknown probabilities permit players to persist in believing (Griffiths, 1990). A question often raised is, "How many times did the individual continue to bet despite losing?" But it is the nature of the data-generating process that informs the predictability of the game and the rationality of a wager. It is the player's representation of the game and how outcomes are generated that determines whether he succeeds in learning or persists in failing in the face of uncertainty. The models internal to the player are critical to understanding why people can persist in gambling despite losses (Gilovich, 1983; Walker, 1992). The conclusions from this research could be applied to future treatment research, improving targeted efforts to modify beliefs about the data-generating processes underlying gambles. Cognitive-based treatments have shown therapeutic gains (Bujold et al., 1994; Ladouceur et al., 1998); however, our research suggests that further efficacy gains may be made by focusing on the patient's internal model of the gamble rather than teaching general principles of randomness. Similarly, simple changes to game infrastructure, such as displaying the probabilities of all outcomes rather than only the long-run expected value, may reduce inappropriate beliefs about the games. Gambles can be such difficult inference problems precisely because the player has so little information about his environment but must use his model nonetheless. Under uncertainty, the focus of our research should be on the cognitive process as well as the outcomes.

ACKNOWLEDGMENTS

This research was partly funded by a scholarship to Erica C. Yu granted by the Responsible Gambling Fund, a UK charity. A portion of this study was presented at the Subjective Probability, Utility and Decision Making conference in Rovereto, Italy in August 2009. The authors wish to thank Nick Chater, Nigel Harvey, Gordon D. A. Brown, and Yaron Shlomi for their comments on earlier versions of this manuscript and two anonymous reviewers.

REFERENCES

Barron, G., and Erev, I. (2003). Small feedback based decisions and their limited correspondence to description based decisions. *J. Behav. Decis. Mak.* 16, 215–233.

Bower, G. H. (1981). *The Psychology of Learning and Motivation: Advances in Research and Theory*. San Diego: Academic Press.

Bujold, A., Ladoucer, R., Sylvain, C., and Boisvert, J. M. (1994). Treatment of pathological gamblers: an experimental study. *J. Behav. Ther. Exp. Psychiatry.* 25, 275–282.

Dayan, P., and Niv, Y. (2008). Reinforcement learning: the good, the bad and the ugly. *Curr. Opin. Neurobiol.* 18, 1–12.

Doherty, M. E., Mynatt, C. R., Tweney, R. D., and Schiavo, M. D. (1979). Pseudodiagnosticity. *Acta Psychol. (Amst.)* 43, 111–121.

Edwards, W. (1968). "Conservatism in human information processing" in *Formal Representation of Human Judgment*, ed. B. Kleinmuntz (New York: Wiley), 17–52.

Edwards, W., Lindman, H., and Savage, L. J. (1963). Bayesian statistical inference for psychological research. *Psychol. Rev.* 70, 193–242.

Fischhoff, B., and Beyth-Marom, R. (1983). Hypothesis evaluation from a Bayesian perspective. *Psychol. Rev.* 90, 239–260.

Fischhoff, B., Slovic, P., and Lichtenstein, S. (1978). Fault trees: sensitivity of estimated failure probabilities to problem representation. *J. Exp. Psychol. Hum. Percept. Perform.* 4, 330–344.

Gilovich, T. (1983). Biased evaluation and persistence in gambling. *J. Pers. Soc. Psychol.* 44, 1110–1126.

Griffiths, M. (1990). Fruit machine gambling: the importance of structural characteristics. *J. Gambl. Stud.* 9, 101–120.

Harsanyi, J. C. (1967). Games with incomplete information played by "Bayesian" players, I-III. *Manage. Sci.* 14, 159–182.

Hertwig, R., and Erev, I. (2009). The description – experience gap in risky choice. *Trends Cogn. Sci. (Regul. Ed.)* 13, 517–523.

Hogarth, R. M., and Einhorn, H. J. (1992). Order effects in belief updating: the belief-adjustment model. *Cogn. Psychol.* 24, 1–55.

Jansen, B. R. J., van Duijvenvoorde, A. C. K., and Huizenga, H. M. (2012). Development of decision making: sequential versus integrative rules. *J. Exp. Child. Psychol.* 111, 87–100.

Johnson-Laird, P. N. (1983). *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness.* Cambridge: Cambridge University Press.

Jones, M., and Love, B. C. (2011). Bayesian fundamentalism or enlightenment? On the explanatory status and theoretical contributions of Bayesian models of cognition. *Behav. Brain Sci.* 34, 169–231.

Kahneman, D., and Tversky, A. (1972). Subjective probability: a judgment of representativeness. *Cogn. Psychol.* 3, 430–454.

Klayman, J., and Ha, Y. W. (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Psychol. Rev.* 94, 211–228.

Koehler, J. J. (1993). The influence of prior beliefs on scientific judgments of evidence quality. *Organ. Behav. Hum. Decis. Process.* 56, 28–55.

Krynski, T. R., and Tenenbaum, J. B. (2007). The role of causality in judgment under uncertainty. *J. Exp. Psychol. Gen.* 136, 430–450.

Ladouceur, R., Sylvain, C., Letarte, H., Giroux, I., and Jacques, C. (1998). Cognitive treatment of pathological gamblers. *Behav. Res. Ther.* 36, 1111–1119.

Lopes, L. L. (1976). Model-based decision, and inference in stud poker. *J. Exp. Psychol. Gen.* 105, 217–239.

Meehl, P. (1954). *Clinical versus Statistical Prediction: A Theoretical Analysis and A Review of the Evidence.* Minneapolis: University of Minnesota Press.

Newell, B. R., and Rakow, T. (2007). The role of experience in decisions from description. *Psychon. Bull. Rev.* 14, 1133–1139.

Pearson, P. D., and Barr, R., Kamil, M. L., and Mosenthal, P. (eds). (1984). *Handbook of Reading Research.* New York: Longman, Inc.

Schank, R., and Abelson, R. (1977). Scripts, plans, goals, and understanding: an inquiry into human knowledge structure. Hillsdale, NJ: Lawrence Erlbaum Associates.

Simon, H. A. (1957). *Models of Man: Social, and national.* New York: Wiley.

Snyder, M., and Swann, W. B. (1978). Behavioral confirmation in social interaction: from social perception to social reality. *J. Exp. Soc. Psychol.* 14, 148–162.

Steyvers, M., Tenenbaum, J. B., Wagenmakers, E.-J., and Blum, B. (2003). Inferring causal networks from observations and interventions. *Cogn. Sci.* 27, 453–489.

Sutton, R. S., and Barto, A. G. (1998). *Introduction to Reinforcement Learning.* Cambridge: MIT Press.

Tenenbaum, J. B., Griffiths, T. L., and Kemp, C. (2006). Theory-based Bayesian models of inductive learning and reasoning. *Trends Cogn. Sci. (Regul. Ed.)* 10, 309–318.

Thaler, R. H., and Johnson, E. J. (1990). Gambling with the house money and trying to break even: the effects of prior outcomes on risky choice. *Manage. Sci.* 36, 643–660.

Troutman, C. M., and Shanteau, J. (1977). Inferences based on nondiagnostic information. *Organ. Behav. Hum. Perform.* 19, 43–55.

Tversky, A., and Kahneman, D. (1971). Belief in the law of small numbers. *Psychol. Bull.* 76, 105–110.

Tversky, A., and Kahneman, D. (1973). Availability: a heuristic for judging frequency and probability. *Cogn. Psychol.* 5, 207–232.

Tversky, A., and Kahneman, D. (1974). Judgment under uncertainty: heuristics and biases. *Science* 27, 1124–1131.

Ungemach, C., Chater, N., and Stewart, N. (2009). Are probabilities overweighted or underweighted, when rare outcomes are experienced (rarely)? *Psychol. Sci.* 20, 473–479.

Walker, M. B. (1992). Irrational thinking among slot machine players. *J. Gambl. Stud.* 8, 245–261.

Yu, A., and Dayan, P. (2005). Uncertainty, neuromodulation, and attention. *Neuron* 46, 681–692.

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 March 2012; accepted: 13 September 2012; published online: 05 October 2012.

*Citation: Yu EC and Lagnado DA (2012) The influence of initial beliefs on judgments of probability. Front. Psychology **3**:381. doi: 10.3389/fpsyg.2012.00381*

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

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 06 June 2012
doi: 10.3389/fpsyg.2012.00177

The role of inertia in modeling decisions from experience with instance-based learning

***Varun Dutt*[1]* and *Cleotilde Gonzalez*[2]**

[1] School of Computing and Electrical Engineering and School of Humanities and Social Sciences, Indian Institute of Technology, Mandi, India
[2] Dynamic Decision Making Laboratory, Department of Social and Decision Sciences, Carnegie Mellon University, Pittsburgh, PA, USA

Edited by:
Konstantinos Tsetsos, Oxford University, UK

Reviewed by:
Christian C. Luhmann, Stony Brook University, USA
Adrian R. Camilleri, Duke University, USA

***Correspondence:**
Varun Dutt, School of Computing and Electrical Engineering and School of Humanities and Social Sciences, Indian Institute of Technology, Mandi, PWD Rest House, Near Bus Stand, Mandi – 175 001, Himachal Pradesh, India.
e-mail: varundutt@yahoo.com

One form of inertia is the tendency to repeat the last decision irrespective of the obtained outcomes while making decisions from experience (DFE). A number of computational models based upon the Instance-Based Learning Theory, a theory of DFE, have included different inertia implementations and have shown to simultaneously account for both risk-taking and alternations between alternatives. The role that inertia plays in these models, however, is unclear as the same model without inertia is also able to account for observed risk-taking quite well. This paper demonstrates the predictive benefits of incorporating one particular implementation of inertia in an existing IBL model. We use two large datasets, estimation and competition, from the Technion Prediction Tournament involving a repeated binary-choice task to show that incorporating an inertia mechanism in an IBL model enables it to account for the observed average risk-taking and alternations. Including inertia, however, does not help the model to account for the trends in risk-taking and alternations over trials compared to the IBL model without the inertia mechanism. We generalize the two IBL models, with and without inertia, to the competition set by using the parameters determined in the estimation set. The generalization process demonstrates both the advantages and disadvantages of including inertia in an IBL model.

Keywords: decisions from experience, instance-based learning, binary-choice, inertia, risk-taking, alternations

INTRODUCTION

People's reliance on inertia, the tendency to repeat the last decision irrespective of the obtained outcomes (successes or failures), has been documented in literature concerning managerial and organizational sciences as well as behavioral sciences (Samuelson, 1994; Reger and Palmer, 1996; Hodgkinson, 1997; Tripsas and Gavetti, 2000; Gladwell, 2007; Biele et al., 2009; Gonzalez and Dutt, 2011; Nevo and Erev, 2012). For example, inertia acts like a *status quo* bias and helps to account for the commonly observed phenomenon whereby managers fail to update and revise their understanding of a situation when it changes, a phenomenon that acts as a psychological barrier to organizational change (Reger and Palmer, 1996; Tripsas and Gavetti, 2000; Gladwell, 2007). In these situations, inertia is generally believed to have a negative effect on decision making (Sandri et al., 2010).

Inertia has also been incorporated to account for human behavior in existing computational models of decisions from experience (DFE). DFE are choices that are based on previous encounters with one's alternatives; as opposed to decisions from description, which are based on summary descriptions detailing all possible outcomes and their respective likelihoods of each option (Hertwig and Erev, 2009). In DFE, researchers have studied both the risk-taking behavior and alternations between alternatives in repeated binary-choice tasks, where decision makers consequentially choose between risky and safe alternatives repeatedly (Samuelson, 1994; Börgers and Sarin, 2000; Barron and Erev, 2003; Erev and Barron, 2005; Biele et al., 2009; Hertwig and Erev, 2009; Erev et al., 2010a; Gonzalez and Dutt, 2011; Nevo and Erev, 2012). The alternations explain how individuals search information and how this search pattern changes over repeated trials. Thus, alternations tell us about the information-search patterns and learning in DFE (Erev et al., 2010a). Accounting for both risk-taking and alternations helps to develop a complete understanding about how decision makers reach certain long-term outcomes, which cannot be determined by solely studying one of these measures in the isolation of the other (Gonzalez and Dutt, 2011).

Most recently, models based upon the Instance-Based Learning Theory (IBLT; and "IBL models" hereafter), a theory of dynamic DFE, have shown to account for both the observed risk-taking and alternations in a binary-choice task better than most of the best known computational models. A number of these IBL models have incorporated some form of the inertia mechanism (Gonzalez and Dutt, 2011; Gonzalez et al., 2011), while others have not incorporated inertia and still accounted for the risk-taking behavior (Lejarraga et al., 2012). For example, Lejarraga et al. (2012) have shown that a single IBL model, without inertia, is able to explain observed risk-taking and generalize across several variants of the repeated binary-choice task. Therefore, it appears that inertia may not be needed in computational models to account for the observed risk-taking. However, Lejarraga et al. (2012) model does not demonstrate how alternations are accounted for or how alternations and risk-taking are accounted for simultaneously. As discussed above, people's experiential decisions may likely rely on inertia, and computational models might need some form of inertia to account for both observed risk-taking and alternations. Yet, the role that inertia mechanisms play in existing

computational models is unclear and needs to be systematically investigated.

In this paper, we evaluate the role of an inertia mechanism in an IBL model. We evaluate a model with inertia and another without inertia for their ability to account for observed risk-taking and alternation behaviors. In order to evaluate the inertia mechanism, we use two large human datasets that were collected in the Technion Prediction Tournament (TPT) involving the repeated binary-choice task (Erev et al., 2010b). In what follows, we first discuss the current understanding of the role of inertia in accounting for DFE. Next, we present the results of calibrating two existing IBL models, with and without inertia, in the TPT's estimation dataset and evaluate the added value and contribution of including inertia. Finally, we present the results that generalize these models into the TPT's competition dataset. We close this paper by discussing our results and highlighting some future directions in this ongoing research program.

THE ROLE OF INERTIA IN DECISIONS FROM EXPERIENCE

Inertia may be a psychological barrier to changes in an organization if decision makers fail to update their understanding of a situation when it changes (Reger and Palmer, 1996; Hodgkinson, 1997; Tripsas and Gavetti, 2000; Gladwell, 2007). For example, Tripsas and Gavetti (2000) provided a popular example of inertia in a managerial setting concerning the Polaroid Corporation. Polaroid believed that it could only make money by producing consumables and not the hardware. Thus, it decided to stick to producing only consumables. This decision led the company to neglect the growth in digital imaging technologies. Because of the prevailing inertial "mental model" of their business, the corporation failed to adapt effectively to market changes. Furthermore, Gladwell (2007) has suggested that inertia is one powerful explanation as to why established firms are not as innovative as young, less established firms. For example, as an established firm, Kodak's management is reported to have suffered from a *status quo* bias due to inertia: They believed that what has worked in the past will also work in the future (Gladwell, 2007).

In judgment and decision making, inertia has been shown to play a role in determining the proportion of risk-taking due to the timing of a descriptive warning message (Barron et al., 2008). Barron et al. (2008) compared the effect of a descriptive warning received *before* or *after* making risky decisions in a repeated binary-choice task. In this task, participants made a choice between a safe option with a sure gain and a risky option with the possibility of incurring a loss or a gain such that the probability of incurring the loss was very small ($p = 0.001$). Thus, most of the time, the task offered gains for both safe and risky choices. These authors show that when an early warning coincides with the beginning of a decision making process, the warning is both weighted more heavily in future decisions and induces safer behavior (i.e., a decrease in the proportion of risky choices), which becomes the *status quo* for future choices. Thus, although the proportion of risk-taking is lower for an early warning message compared to a late warning message, the risky and safe choices in both cases show excessive reliance on inertia to repeat the last choice made. Here, inertia acts like a double-edged sword: It is likely to encourage or discourage ongoing risky behavior depending upon the timing of a warning.

Some researchers have depicted inertia as an irrational behavior in which individuals hold onto choices that clearly do not provide the maximizing outcome for too long (Sandri et al., 2010). However, these authors have only shown that behavior may be inconsistent with one specific rational model of maximization, which may be an arbitrary standard that is difficult to generalize to other rational models of maximization. There are certain other situations where inertia is likely to produce positive effects as well. In psychology, inertia is also believed to be a key component of love, trust, and friendship (Cook et al., 2005). If evidence shows that a friend is dishonest, then the decision to mistrust the friend in future interactions would demand much more instances of dishonesty from the friend than that required to form an opinion about a stranger. Thus, the inertia of continuing to trust the friend makes it difficult to break the friendship.

Inertia has been incorporated in a number of existing cognitive models of DFE. It is believed that inertia helps these models account for both observed risk-taking and alternations in the repeated binary-choice (Samuelson, 1994; Börgers and Sarin, 2000; Biele et al., 2009; Erev et al., 2010a; Gonzalez and Dutt, 2011; Nevo and Erev, 2012). For example, Erev et al. (2010a) observed that in the repeated binary-choice task, participants selected the alternative that led to an observed high outcome in the last trial in 67.4% of the trials, while they repeated their last choice for an alternative, irrespective of it being high or low, in 75% of the trials. These observations suggest that participants tend to repeat their last choice even when it does not agree with the high outcome in their last experience, exhibiting robust reliance on inertia that seems to be independent of observed outcomes. Some researchers have suggested that in situations where estimating the choice that yields high outcomes from observation is costly, difficult, or time consuming, relying on inertia might be the most feasible course of action (Samuelson, 1994). But other researchers have found this inertia effect even when the forgone outcome (i.e., what respondents would have gotten had they chosen the other alternative) is greater than the obtained outcome (Biele et al., 2009).

In order to account for these observations, recent computational models of DFE have explicitly incorporated three different forms of inertia as part of their specification (Erev et al., 2010a; Gonzalez and Dutt, 2011; Gonzalez et al., 2011). In the first form, inertia increases over time as a result of a decrease in surprise, where surprise is defined as the difference in expected values of the two alternatives (Erev et al., 2010a). This definition of inertia has been included in the Inertia Sampling and Weighting (I-SAW) model. The I-SAW model was designed for a repeated binary-choice market-entry task, and it distinguishes between three explicit response modes: exploration, exploitation, and inertia (Erev et al., 2010a; Chen et al., 2011). The I-SAW model also provides reasonable predictions in the repeated binary-choice task (Nevo and Erev, 2012). Inertia is represented in this model with the assumption that individuals tend to repeat their last choice, and the probability of inertia in a trial is a function of surprise. Surprise is calculated as the difference in the expected value of the two alternatives due to the observed outcomes in each alternative in previous trials. The probability of inertia is assumed to increase over trials, as surprise decreases over trials. This definition based upon surprise incorporates the idea of learning over repeated trials

of game play where, due to repeated presentations of the same set of outcomes, participants tend to get increasingly less surprised and begin to stick to an option that they prefer (i.e., show inertia in their decisions).

In the second form (that is similar to the first form), inertia increases over time as a result of a decrease in surprise, which is based upon the difference in blended values (a measure of utility of alternatives based on past experience in Gonzalez et al., 2011 model). This definition of inertia has been included in the IBL model that was runner-up in the market-entry competition (Gonzalez et al., 2011). This model includes an inertia mechanism that is driven by surprise like in the I-SAW model; however, surprise here is calculated as the difference between the blended values of two alternatives.

In the third and simpler form, inertia is a probabilistic process that is triggered randomly over trials, where the random occurrences of inertia are based upon a calibrated probability parameter, *pInertia* (Gonzalez and Dutt, 2011). This definition of inertia is the one we evaluate in this paper, as it was recently included in an IBL model that produced robust predictions superior to many existing models (Gonzalez and Dutt, 2011). According to Gonzalez and Dutt (2011), the IBL model with the *pInertia* parameter accounts for both observed risk-taking and alternations simultaneously in different paradigms of DFE and performs consistently better than most existing computational models of DFE that competed in the TPT.

Although computational models have included inertia in several forms, Lejarraga et al. (2012) have recently shown that a single IBL model without any inertia assumption is also able to account for the observed risk-taking behavior in different tasks that included probability-learning, binary-choice with fixed probability, and binary-choice with changing probability. Although the use of some form of inertia seems necessary in many computational models of DFE (Erev et al., 2010a; Chen et al., 2011; Gonzalez and Dutt, 2011; Gonzalez et al., 2011; Nevo and Erev, 2012), its role in accounting for risk-taking and alternations in DFE is still unclear and a systematic investigation of its role in computational models is needed.

Given the wide use of inertia in computational models, it is likely that incorporating inertia assumptions might make them more ecologically valid. That seems likely because if a model accounts for risk-taking behavior already, then incorporating a form of inertia in its specification might directly influence its ability to account for alternations as well. However, we currently do not know how inertia in a model might impact its ability to account for both the risk-taking behavior and the alternations simultaneously. The incorporation of inertia in a model is likely to be beneficial only if it improves the model's ability to account for both risk-taking and alternations, and not solely one of these measures.

MATERIALS AND METHODS

RISK-TAKING AND ALTERNATIONS IN THE TECHNION PREDICTION TOURNAMENT

The TPT (Erev et al., 2010b) was a modeling competition organized in 2008 in which different models were submitted to predict choices made by human participants. Competing models were evaluated following the generalization criterion method (Busemeyer and Wang, 2000), by which models were fitted to choices made by participants in 60 problems (the estimation set) and later tested in a new set of 60 problems (the competition set) with the parameters obtained in the estimation set. Although the TPT involved three different experimental paradigms, here we use data from the E-repeated paradigm that involved consequential choices in a repeated binary-choice task with immediate feedback on the chosen alternative. We use this dataset to evaluate the inertia mechanism in an IBL model.

The TPT dataset's 120 problems involved a choice between a safe alternative that offered a medium (M) outcome with certainty; and a risky alternative that offered a high (H) outcome with some probability (pH) and a low (L) outcome with the complementary probability. The M, H, pH, and L were generated randomly, and a selection algorithm assured that the 60 problems in each set were different in domain (positive, negative, and mixed outcomes) and probability (high, medium, and low pH). The positive domain was such that each of the M, H, and L outcomes in a problem were positive numbers (>0). The mixed domain was such that one or two of the outcomes among M, H, and L (but not all three) in a problem were negative (<0). The negative domain was such that each of the M, H, and L outcomes in a problem were negative numbers (<0). The low, medium, and high probability in a problem corresponded to the value of pH between 0.01–0.09, 0.1–0.9, and 0.91–0.99, respectively. The selection algorithm ensured that there were 20 problems each for the three domains and about 20 problems each for the three probability values in the estimation and the competition sets. The resulting set of problems in the three domains and the three probability values was large and representative. For each of the 60 problems in the estimation and competition set, a sample of 100 participants was randomly assigned into 5 groups, and each group completed 12 of the 60 problems. Each participant was instructed to repeatedly and consequentially select between two unlabeled buttons on a computer screen in order to maximize long-term rewards for a block of 100 trials per problem (the end point on trials was not provided or known to participants). One button was associated with a risky alternative and the other button with a safe alternative. Clicking a button corresponding to either the safe or risky alternative generated an outcome associated with the selected button (i.e., there was only partial feedback and participants were not shown the foregone outcome on the unselected button). The alternative with the higher expected value, which could be either the safe or risky, could maximize a participant's long-term rewards. Other details about the E-repeated paradigm are reported in Erev et al. (2010b).

The models submitted to the TPT were not provided with the alternation data (i.e., the A-rate), and they were evaluated only according to their ability to account for risk-taking behavior (i.e., the R-rate; Erev et al., 2010b). Gonzalez and Dutt (2011) had calculated the A-rate for analyses of alternations from the TPT datasets and we followed the exact same procedures in this paper. First, alternations were either coded as 1 s (a respondent switched from making a risky or safe choice in the last trial to making a safe or risky choice in the current trial) or as 0 s (the respondent repeated the same choice in the current trial as that in the last trial). Then, the A-rate is computed as the proportion of alternations in each

trial starting in trial 2 (the A-rate in trial 1 is undefined as there is no preceding trial to calculate alternations). The proportion of alternations in each trial is computed by averaging the alternations over 20 participants per problem and 60 problems in each dataset. The R-rate is the proportion of risky choices (i.e., choices of the risky alternative) in each trial averaged over 20 participants per problem and 60 problems in each dataset.

Figure 1 shows the overall R-rate and A-rate over 99 trials from trial 2 to 100 in the estimation and competition sets. As seen in both datasets, the R-rate decreases slightly across trials, although there is a sharp decrease in the A-rate. The sharp decrease in the A-rate shows a change in the exploration (information-search) pattern across repeated trials. Overall, the R-rate and A-rate curves suggest that participants' risk-taking behavior remains relatively steady across trials, while they learn to alternate less and choose one of the two alternatives more often. Later in this paper, we evaluate the role of inertia mechanism to account for these R- and A-rate curves in **Figure 1** in a computational IBL model.

AN INSTANCE-BASED LEARNING MODEL OF REPEATED BINARY-CHOICE

Instance-Based Learning Theory has been used for developing computational models that explain human behavior in a wide variety of dynamic decision making tasks. These tasks include dynamically complex tasks (Gonzalez and Lebiere, 2005; Gonzalez et al., 2003; Martin et al., 2004), training paradigms of simple and complex tasks (Gonzalez et al., 2010), simple stimulus-response practice and skill acquisition tasks (Dutt et al., 2009), and repeated binary-choice tasks (Lebiere et al., 2007; Gonzalez and Dutt, 2011; Gonzalez et al., 2011; Lejarraga et al., 2012) among others. Its applications to these diverse tasks illustrate its generality and its ability to explain DFE in multiple contexts.

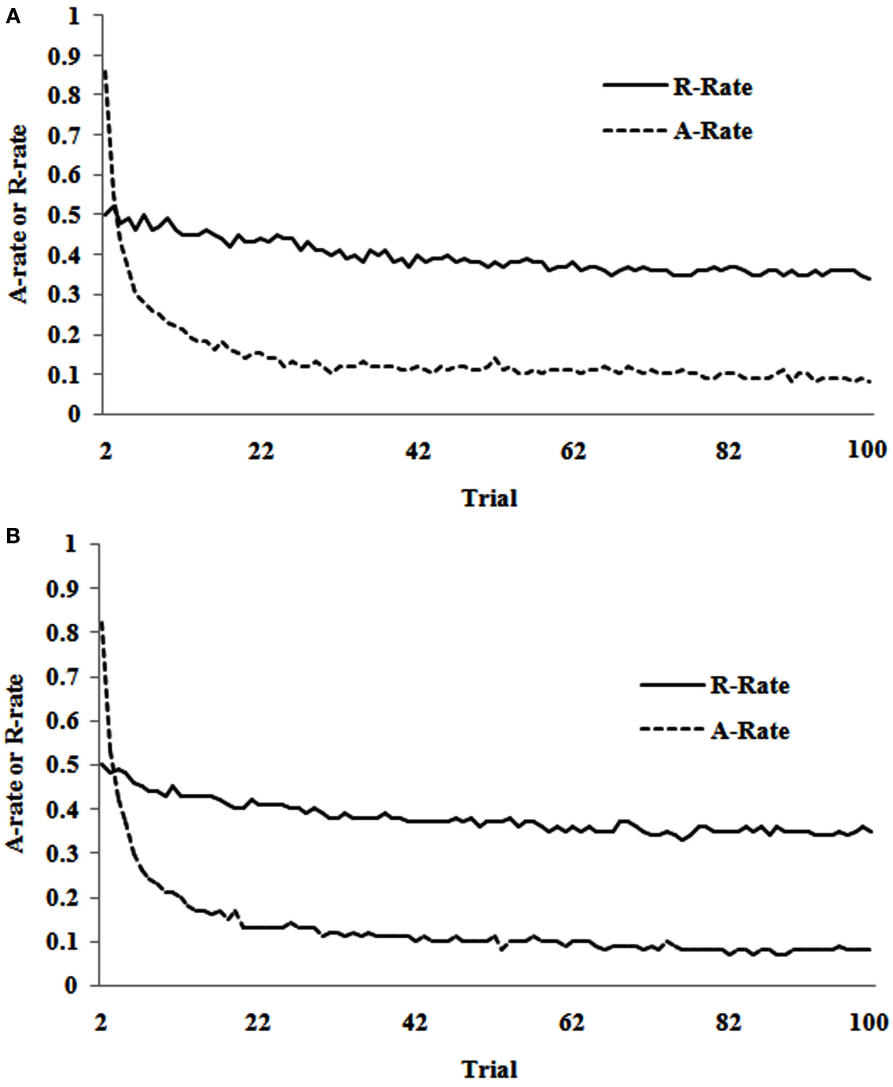


FIGURE 1 | (A) The R-rate and A-rate across trials observed in human data in the estimation set of the TPT between trial 2 and trial 100. **(B)** The R-rate and A-rate across trials observed in human data in the competition set of the TPT between trial 2 and trial 100.

Here, we briefly discuss an IBL model that has shown to successfully account for both risk-taking and alternation behaviors in DFE (Gonzalez and Dutt, 2011). This model assumes reliance on recency, frequency, and random inertia to make choice selections. Here, we evaluate how the same IBL model, with and without the random inertia mechanism, can simultaneously account for risk-taking and alternation in repeated binary-choice. This evaluation will enable us to better understand the role of this particular simpler formulation of inertia in computational IBL models.

IBL model

All IBL models propose an "instance" as a key representation of cognitive information. An instance is a representation of each decision alternative and consists of three parts: a situation (a set of attributes that define the alternative), a decision for one of the many alternatives, and an outcome resulting from making that decision. An IBL model of repeated binary-choice has assumed a simple instantiation of inertia mechanism (Gonzalez and Dutt, 2011): A free parameter, called *pInertia*, determines the repetition of the previous choice in the current decision according to this probabilistic mechanism (see the Appendix for the formal definitions of all the mechanisms of the IBL model for repeated binary-choice and the *pInertia* parameter). If a uniformly distributed random number is less than the probability *pInertia*, then the model repeats its last choice; otherwise, the model compares blended values for the risky and safe alternatives, and makes a choice for the alternative with the higher blended value.

In this paper, we call this IBL model implementation with the random inertia mechanism, the "IBL-Inertia model." In addition, we consider another version of the same model, but without this inertia mechanism (Lejarraga et al., 2012) as a baseline to compare against the IBL-Inertia model. We call this implementation without inertia, the "IBL model." In the absence of inertia, this model relies solely on the comparison between the blended values for the risky and safe alternatives to make choice selections in each trial (the IBL-Inertia model also compares blended values to make choice selections; however, blended values are used only in the IBL-Inertia model when a random number is greater than or equal to the *pInertia* parameter in a trial). With the exception of the presence of *pInertia* in the IBL-Inertia model and its absence in the IBL model, both models are identical in all other respects.

Blending, as proposed in both model implementations, is a function of the probability of retrieving instances from memory multiplied by their respective outcomes stored in instances (Lebiere, 1999; Gonzalez and Dutt, 2011; Lejarraga et al., 2012). Each instance consists of a label that identifies a decision alternative in the task and the outcome obtained. For example (A, \$32) is an instance where the decision was to choose the risky alternative (A) and the outcome obtained was \$32. The probability of retrieving an instance from memory is a function of its activation (Anderson and Lebiere, 1998). A simplified version of the activation mechanism, which relies on recency and frequency of using instances and noise in retrieval, has been shown to be sufficient to capture human choices in several repeated binary-choice and probability-learning tasks (Gonzalez and Dutt, 2011; Lejarraga et al., 2012). The activation is influenced by the decay parameter d that captures the rate of forgetting or reliance on recency. The higher the value of the d parameter, the greater is the model's reliance on recent experiences. The activation is also influenced by a noise parameter s that is important for capturing the variability in human behavior from one participant to another.

For the first trial, both model implementations, IBL-Inertia and IBL, have no instances in memory from which to calculate blended values. Therefore, these implementations make a selection between instances that are pre-populated in their memory. We used a value of +30 in the outcome slot of the two alternatives' instances (Gonzalez and Dutt, 2011). The +30 value is arbitrary, but most importantly, it should be greater than any possible outcomes in the TPT problems to trigger an initial exploration of the two alternatives. For the first trial, the choice between the two alternatives in both implementations is solely based on the blended values. From the second trial onward, the inertia mechanism is used along with blending in IBL-Inertia model and only blending is used in the IBL model.

RESULTS

MODEL CALIBRATION AND EVALUATION OF INERTIA

The IBL model is compared with the IBL-Inertia model for their ability to account for both the proportion of risk-taking (R-rate) and alternations (A-rate) across trials. We will first calibrate the models' shared parameters, noise s and decay d, to the data in the TPT's estimation set. Then, we explore the role of adding the *pInertia* parameter to the IBL model (i.e., the IBL-Inertia model) by recalibrating all its parameters. Then, we generalize both the calibrated models, IBL and IBL-Inertia, to the TPT's competition set.

Calibrating a model to human data means finding the parameter values that minimize the mean-squared deviation (MSD) between the model's predictions and the observed human performance on a dependent measure. We used a genetic algorithm program to calibrate the model's parameters. The genetic algorithm tried different combinations of parameters to minimize the sum of MSDs between the model's average R-rate per problem and the average A-rate per problem measures and the corresponding values in human data (we call this sum as the combined R-rate and A-rate measure). Calibrating on the combined R-rate and A-rate measure is expected to produce the best account for both measures in human data compared to using only one of these measures (Dutt and Gonzalez, under review). Also, calibrating on the combined R-rate and A-rate measure allows us to test the IBL model's maximum potential to account for both these measures.

In order to compare results on the R-rate and A-rate during calibration, we use the AIC (Akaike Information Criterion) measure in addition to the MSD (mean-squared deviation) measure. The AIC definition takes into account both a model's complexity (estimated by the number of free parameters in the model), as well as its accuracy (estimated by G^2, defined the "lack of fit" between model and human data; Pitt and Myung, 2002; Busemeyer and Diederich, 2009). The AIC definition and the computation procedures used here are the same as those used by Gonzalez and Dutt, 2011; for more details on the AIC definition refer to the Appendix). The use of AIC during calibration is relevant because the IBL and IBL-Inertia models are hierarchical (or nested) models (Maruyama, 1997; Loehlin, 2003; Kline, 2004) and they differ only in terms of

the inertia mechanism. Thus, the IBL model can be simply derived from the IBL-Inertia model by restricting the *pInertia* parameter's value to 0 during model calibration. Furthermore, in order to capture the trend of R-rate and A-rate from a model over trials, we used the Pearson's correlation coefficient (r) between model and human data across trials (for the A-rate we used trials 2–100 and for the R-rate we used trials 1–100; the A-rate is undefined for trial 1). Also, we computed the MSE (mean-squared error) between model and human data across trials. For the MSE, we averaged the R-rate and A-rate in model and human data across all participants and problems in a dataset for each trial. Then, we calculated the mean of the squared differences between model and human data for each trial. Because the MSE is computed across trials, it measures the distance between the model and human data curves trial-by-trial (for more details on the MSE definition refer to the Appendix).

For the purpose of calibration, the average R-rate per problem and the average A-rate per problem were computed by averaging the risky choices and alternations in each problem over 20 participants per problem and 100 trials per problem (for the A-rate per problem, only 99 trials per problem were used for computing the average). Later, the MSDs were calculated across the 60 problems by using the average R-rate per problem and the average A-rate per problem measures from the model and human data in the estimation set. Some researchers suggest calibrating models to the data of each participant per problem rather than to aggregate measures (Pitt and Myung, 2002; Busemeyer and Diederich, 2009); however, the calibration to aggregate behavior is quite common in the cognitive and behavioral sciences (e.g., Anderson et al., 2004; Erev et al., 2010a; Gonzalez and Dutt, 2011; Lejarraga et al., 2012). In fact, calibrating to aggregate measures is especially meaningful when the participant-to-participant variability in the dependent measure is small compared to the value of the dependent measure itself (Busemeyer and Diederich, 2009). In the estimation and competition sets, the standard deviations for the A-rate and R-rate were similar and very small (∼0.1) compared to the values of the R-rate (∼0.5) and the A-rate (∼0.3) measures themselves. Thus, we use the average dependent measures for the purposes of model calibration in this paper.

For calibrating the models, both the s parameter and the d parameters were varied between 0.0 and 10.0, and the *pInertia* parameter was varied between 0.0 and 1.0. Although the genetic algorithm can continue to indefinitely optimize parameters, it was stopped when there was no change in the parameter values obtained for a consecutive period of 200 generations. The assumed range of variation for the *pInertia*, s, and d parameters, and the decision process to stop the genetic algorithm are expected to provide good optimal parameter estimates (Gonzalez and Dutt, 2011). Also, the large range of parameters' variation ensures that the optimization process does not miss the minimum sum of MSDs (for more details about genetic algorithm optimization, please see Gonzalez and Dutt (2011).

We calibrated both the IBL and IBL-Inertia models to the combined R-rate and A-rate in TPT's estimation set. The purpose of the calibration was to obtain optimized values of d and s parameters in the IBL model and *pInertia*, d, and s parameters in the IBL-Inertia model. Later, keeping d and s parameters at their optimized values in the IBL-Inertia model, we varied the *pInertia* parameter from 0.0 to 1.0 in increments of 0.05. By only varying the *pInertia* parameter and keeping the other parameter values fixed at their optimized values, we were able to determine the inertia mechanism's full contribution in the model.

Table 1 shows the values of calibrated parameters, MSD, r, AIC, and MSE compared to baseline for IBL and IBL-Inertia models in TPT's estimation set. First, both models' d and s parameters have values in the same range as those reported by Lejarraga et al. (2012). Lejarraga et al. (2012) reported $d = 5$ and $s = 1.5$ for a MSD = 0.0056 calibrated on R-rate using the IBL model. As documented by Lejarraga et al. (2012), the values of both d and s reported in **Table 1** are high compared to the ACT-R default values of $d = 0.5$ and $s = 0.25$ (the default values were reported by Anderson and Lebiere (1998, 2003). A high d value points to a quick decay in memory and a strong dependence on recently experienced outcomes (i.e., reliance on recency). The high s value allows the model to exhibit participant-to-participant variability in capturing the R-rate and A-rate. The *pInertia* value in IBL-Inertia model (=0.62) is high and it shows that on a trial, this model is likely to repeat its previous choice with a 62% chance. In general, the results from both models are generally good (MSDs <0.05 and MSEs <0.05), where both models perform slightly better at capturing the human A-rate than the human R-rate.

Secondly, the individual MSDs, MSEs, and AICs on the R-rate and A-rate in the IBL model are larger than those in the IBL-Inertia model. For example, in the IBL-Inertia model, the MSDs for the R-rate, A-rate, and the sum of R-rate and A-rate are consistently smaller than those in the IBL model (0.008 < 0.016, an improvement of +0.008; 0.003 < 0.005, an improvement of +0.002; and, 0.011 < 0.021, an improvement of +0.010). Also, the relative AIC in the IBL-Inertia model is negative (i.e., better) for both the R-rate and the A-rate. Thus, even with an extraparametric complexity (the *pInertia* parameter), the IBL-Inertia model performs more accurately compared to the IBL model. Although the MSE in the

Table 1 | The values of calibrated parameters for IBL and IBL-Inertia models and the MSD, *r*, AIC, and MSE in TPT's estimation set.

Model	Calibrated parameters	MSD	*r*	AIC	MSE
IBL (calibrated upon R-rate + A-rate)	$d = 8.31$; $s = 1.26$	0.005 (A-rate) 0.016 (R-rate) 0.021 (R-rate + A-rate)	0.95 (A-rate) 0.94 (R-rate)	−479.2 (A-rate) −546.3 (R-rate)	0.0076 (A-rate) 0.0041 (R-rate)
IBL-Inertia (calibrated upon R-rate + A-rate)	$d = 6.71$; $s = 1.40$; *pInertia* = 0.62	0.003 (A-rate) 0.008 (R-rate) 0.011 (R-rate + A-rate)	0.85 (A-rate) 0.92 (R-rate)	−561.3 (A-rate) 680.0 (R-rate)	0.0032 (A-rate) 0.0010 (R-rate)

IBL model is larger than that in the IBL-Inertia model for both R-rate and A-rate; however, as is also shown in **Table 1**, the IBL-Inertia model does not account for the trends in the R-rate and the A-rate across trials compared with the IBL model (the *r* in the IBL model is greater than that in the IBL-Inertia model for both the R-rate and the A-rate).

Figure 2 presents the R-rate and A-rate across trials predicted by the calibrated IBL and IBL-Inertia models and that observed in human data in the TPT's estimation set. In general, these results reveal that both models generate good accounts for both observed risk-taking and alternation behaviors. The IBL model is able to capture the gradual decreasing trend in the A-rate as well as the slightly decreasing trend in risk-taking across trials. However, the model's account for the R-rate exhibit as lightly greater decrease compared with that observed in human data across increasing number of trials. Also, the model's account for the A-rate shows more alternations during about the first half of the trials than that observed in human data. This latter observation is likely due to the +30 pre-populated instances initially put in model's memory, which make it explore both options for a longer time and causes a higher A-rate in the first few trials. However, with increasing trials, the activation of these pre-populated instances becomes weak (as these values are not observed in the problems) and their influence on the A-rate diminishes, causing the A-rate to decrease sharply and meet the human data.

As shown in the bottom graphs of **Figure 2**, the IBL-Inertia model corrects for the under-estimation and over-estimation in the R-rate and A-rate. However, because of the *pInertia* parameter, the model is unable to account for the initial decrease in the A-rate in the first few trials as well as the IBL model, which does so naturally. A likely reason is the high calibrated value of *pInertia* parameter (=0.62) that overshadows the effect of pre-populated instances in the first few trials. Also, it seems that the random effect of *pInertia* across trials causes disruptions in IBL-Inertia model's R-rate trends over trials. Overall, these observations explain why the IBL-Inertia accounts for overall behavior better than the IBL model, but it does not account for the trends in the R-rate and the A-rate.

EVALUATING THE INERTIA MECHANISM

Although the analyses above provide some benefits of including *pInertia* in the IBL model, one would like to understand these benefits more thoroughly for different values of the *pInertia* parameter over its entire range. If including *pInertia* in the IBL model is beneficial, then we should observe smaller MSDs on the R-rate and A-rate across a large part of the parameter's range of variation compared with the IBL model without *pInertia*. Also, this analysis is important because the calibrated value of *pInertia* in the IBL model was found to be high (=0.62), minimizing the role of the blending mechanism.

For this investigation, we used the IBL-Inertia model with its optimized parameters calibrated on the combined R-rate and A-rate measure (i.e., $d = 6.71$; $s = 1.40$) and varied the *pInertia* parameter from 0.0 to 1.0 in increments of 0.05 in TPT's estimation set. Varying *pInertia* like so allows us to determine the range of values for which the sum of the MSDs computed on the average R-rate per problem and the average A-rate per problem are minimized.

Figure 3 shows the MSDs for the IBL-Inertia model calibrated on the combined R-rate and A-rate as a function of *pInertia* values in the estimation set. It also shows the three corresponding MSDs from the original IBL model (shown as dotted lines in **Figure 3**) for

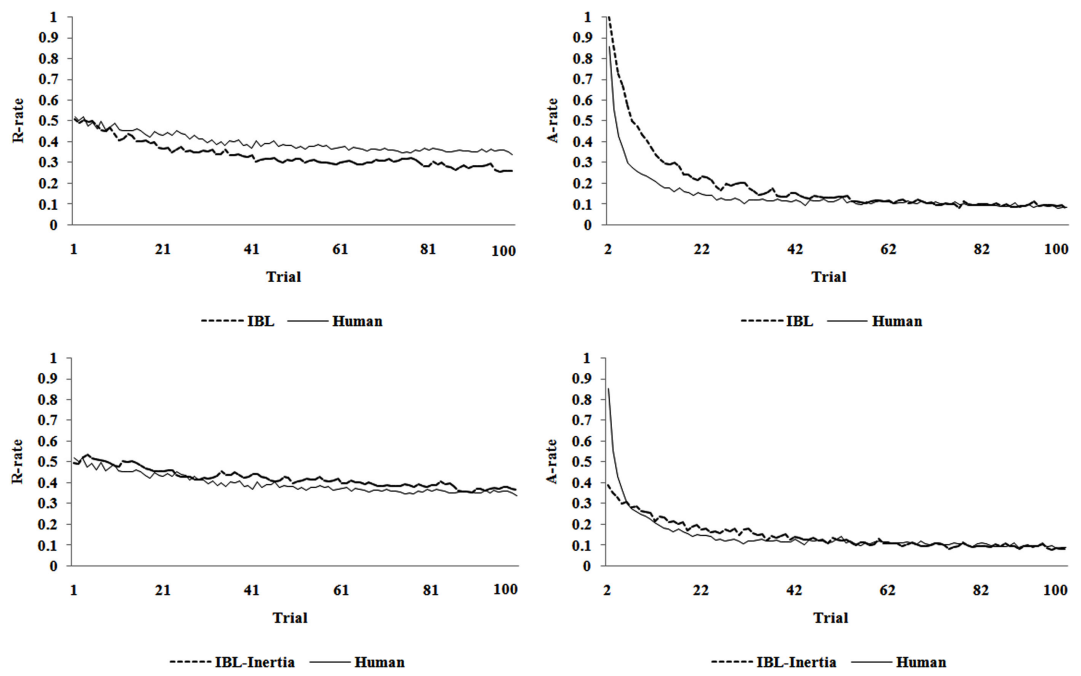


FIGURE 2 | The R-rate and A-rate across trials predicted by the IBL and IBL-Inertia models and that observed in human data in the TPT's estimation set.

comparison purposes (these MSDs are also reported in **Table 1**). The MSDs for the R-rate, the A-rate, and the sum of the MSDs for the R-rate and A-rate in the IBL-Inertia model are below the corresponding MSDs in the IBL model for all values of *pInertia* greater than 0.05 and less than 0.90. Thus, including inertia in the IBL model and calibrating all model parameters improves the model's ability to account for the average R-rate and A-rate compared with the IBL model without inertia. Also, the advantages of including *pInertia* parameter seem to be present over a large range of this parameter's variation.

GENERALIZING THE IBL MODELS TO THE COMPETITION SET

A popular method of comparing models of different complexity is through models' generalization in novel conditions (Stone, 1977; Busemeyer and Wang, 2000; Ahn et al., 2008). In generalization, the calibrated models with different complexities (number of free parameters) are run in novel conditions to compare their performance. The novel conditions would minimize any advantage the model with more parameters has over the model with fewer parameters. In fact, TPT also accounted for model complexity among submitted models by generalization, i.e., by running models in the new competition set with the parameters obtained in the estimation set (Erev et al., 2010b). We used the same procedures as used in the TPT and generalized the calibrated IBL and IBL-Inertia models to TPT's competition set.

In related research, we have claimed that the TPT's estimation and competition data sets are too similar, raising questions regarding the value of using the competition set for generalization (Gonzalez and Dutt, 2011; Gonzalez et al., 2011). These similarities arise because the problems used in the estimation and competition sets were generated by using the same algorithm. However, given that the TPT competition set was collected in a new experiment, involving new problems, and involving a different set of participants from that of the estimation set, testing the models in the competition set is still a relevant exercise to determine the robustness of the models. This generalization further helps us to take into account both models' complexity (number of parameters) and their accuracy of predictions (MSDs; Busemeyer and Diederich, 2009).

The IBL model and IBL-Inertia model were run in the TPT's competition set problems using the parameters determined in the estimation set: $d = 6.71$, $s = 1.40$, and $pInertia = 0.62$ (the *pInertia* parameter is only for the IBL-Inertia model). As previously mentioned, these parameter values had resulted in the lowest MSDs on the combined R-rate and A-rate measure for the two models in the estimation set. **Table 2** shows the values of MSD, *r*, and MSE for the IBL and IBL-Inertia models upon their generalization in TPT's competition set. The IBL-Inertia model's predictions resulted in overall MSDs and MSEs for the R-rate and the A-rate that were smaller than those for the IBL model. Like in the estimation set, however, the IBL-Inertia model did not account for the over trial trend in the R-rate and the A-rate compared with the IBL model (demonstrated by the *r* calculations). These results demonstrate that the IBL-Inertia model can generalize to new problems more accurately (in terms of average overall performance in both the

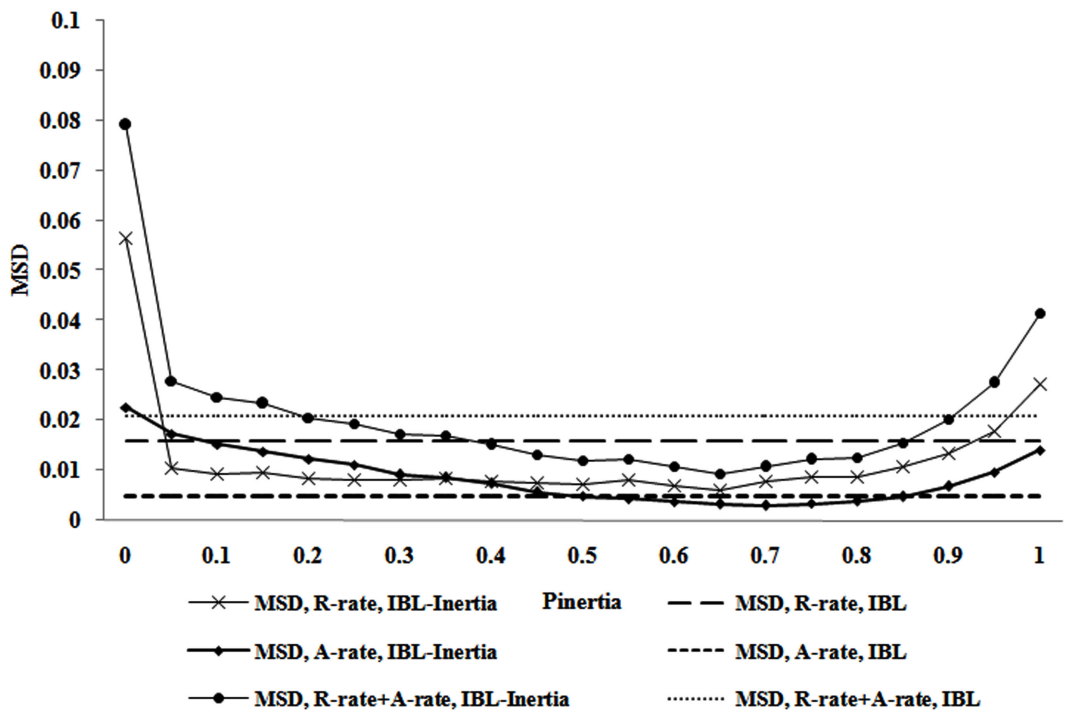


FIGURE 3 | The MSD for the R-rate, the MSD for the A-rate, and the MSD for the combined R-rate and A-rate for different values of *pInertia* parameter in IBL-Inertia model (the corresponding MSDs for the IBL model are also plotted as dotted lines for comparison). The IBL-Inertia model used the calibrated parameters for *d* and *s* parameters (i.e., $d = 6.41$ and $s = 1.40$).

Table 2 | The values of MSD, *r*, and MSE for IBL and IBL-Inertia models upon their generalization in TPT's competition set.

Model	MSD	*r*	MSE
IBL	0.011 (A-rate)	0.96 (A-rate)	0.010 (A-rate)
	0.022 (R-rate)	0.96 (R-rate)	0.010 (R-rate)
	0.033 (R-rate + A-rate)		
IBL-Inertia	0.003 (A-rate)	0.87 (A-rate)	0.003 (A-rate)
	0.007 (R-rate)	0.94 (R-rate)	0.001 (R-rate)
	0.010 (R-rate + A-rate)		

A-rate and R-rate measures across problems and across trials compared with the IBL model; but the IBL-Inertia model also cannot account for trends across trials in these measures compared with the IBL model without inertia).

Figure 4 shows the R-rate and the A-rate over trials for human data, and how the IBL and IBL-Inertia models generalized in the competition set. The IBL model, upon generalization, underestimates the observed R-rate and overestimates the observed A-rate in the competition set. These patterns of under- and over-estimations are similar to those observed in the model's predictions in the estimation set in **Figure 2**. The IBL-Inertia model's predictions about the human R-rate and A-rate in the competition set, however, were very good with very little under- and over-estimations of the observed R-rate and A-rate curves. Furthermore, because the *pInertia* parameter (=0.62) is fixed across trials at a high value in the IBL-Inertia model, the model does not alternate as much as humans in the first few trials. As seen in the lower right graph, the IBL-Inertia model's A-rate starts around 40%, rather than the 85% as observed in human data. Thus, the IBL-Inertia model is not able to account for the initially high A-rate and the rapid decrease in the A-rate in the first few trials compared with the IBL model in its predictions.

DISCUSSION

Some computational models of DFE do not include any inertia assumptions and are still able to account for the observed risk-taking behavior (Lejarraga et al., 2012). However, a number of recent computational models have included some form of inertia to account for observed DFE (Erev et al., 2010a; Gonzalez and Dutt, 2011; Gonzalez et al., 2011). Three different inertial forms have been proposed: random inertia (Gonzalez and Dutt, 2011); inertia as a function of surprise determined by the differences in expected values (Erev et al., 2010a); and inertia as a function of surprise determined by the differences in blended values (Gonzalez et al., 2011). This research uses the particular case of random inertia in an IBL model and determines the benefits of this mechanism by considering two IBL models with and without this mechanism. We selected the random inertia form for our evaluation because of its simplistic formulation, but also because an existing IBL model with this definition accounts for DFE better than other best known models of DFE (Gonzalez and Dutt, 2011).

Our results reveal that a simple instantiation of the inertia mechanism can be used to improve the ability of the IBL model to account for the average risk-taking (R-rate) and alternations (A-rate; based upon MSDs, MSEs, and AICs) observed in human data. However, we also find that the inclusion of random inertia does not help the model to account for the trends across trials in the R-rate and A-rate compared with the same model without inertia (based upon correlation coefficients, *r*). We draw our conclusions based upon model calibration and model generalization that is known to account for increased model complexity (number of parameters) in novel test environments (Busemeyer and Diederich, 2009).

Most current models of DFE have been successful at capturing the risk-taking behavior, but not the underlying alternations observed in repeated binary-choice; such as the tendency to repeat choices irrespective of the obtained outcome in the last trial (Biele et al., 2009). This observation is perhaps not a coincidence, because predicting risk-taking behavior and alternation effects simultaneously is a very challenging task (Rapoport et al., 1997; Erev and Barron, 2005; Estes and Maddox, 2005). In order to overcome some of the challenges, a number of computational models have considered the inclusion of some form of inertia with some initial success (Erev et al., 2010a; Gonzalez and Dutt, 2011; Gonzalez et al., 2011). As can be seen in our results, the random inertia's inclusion into the IBL model helps the model to account for both the average A-rate and R-rate in terms of MSDs, MSEs, and AICs, but not in terms of trends in these rates over trials. Because random inertia accounts for the average A-rate and R-rate in human data, it helps to reduce the observed under-estimation and over-estimation of the observed R-rate and A-rate, respectively, which is seen in the model without the inertia mechanism. This finding might suggest that the inclusion of some form of inertia into computational models might be ecologically plausible for capturing the average risk-taking and alternation behaviors more accurately, but not for the trend in these behaviors over time.

Although the introduction of inertia into the IBL model generally improves the fits to the average human data (based upon MSDs, MSEs, and AICs), it is likely that few modelers may be impressed by this particular result. It is well-known that a model with more parameters (i.e., greater model complexity) can fit a dataset better than a model with fewer parameters (Pitt and Myung, 2002). We dealt with this issue through model generalization (Stone, 1977). The generalization helped to test models with different parametric assumptions in a novel environment (Busemeyer and Diederich, 2009). We used these procedures and generalized the IBL and the IBL-Inertia models to the TPT's competition dataset to compare their performance.

Although the error across trials between IBL-Inertia model and human data was smaller compared to that between IBL model and human data; however, unlike the IBL model, the IBL-Inertia model did not capture the trends in the R-rate and A-rate across trials. The most likely reason is that the inertia parameter in its current formulation is a noisy selection of choices across trials, which disregards the choices derived based upon blended values. Gonzalez et al. (2011) had assumed an inertia formulation that was based upon surprise, where surprise was a function of the difference in blended values of the two alternatives. Perhaps, if the inertia mechanism in the model is formulated as described by Gonzalez et al. (2011), then the trends across trials might be better accounted for

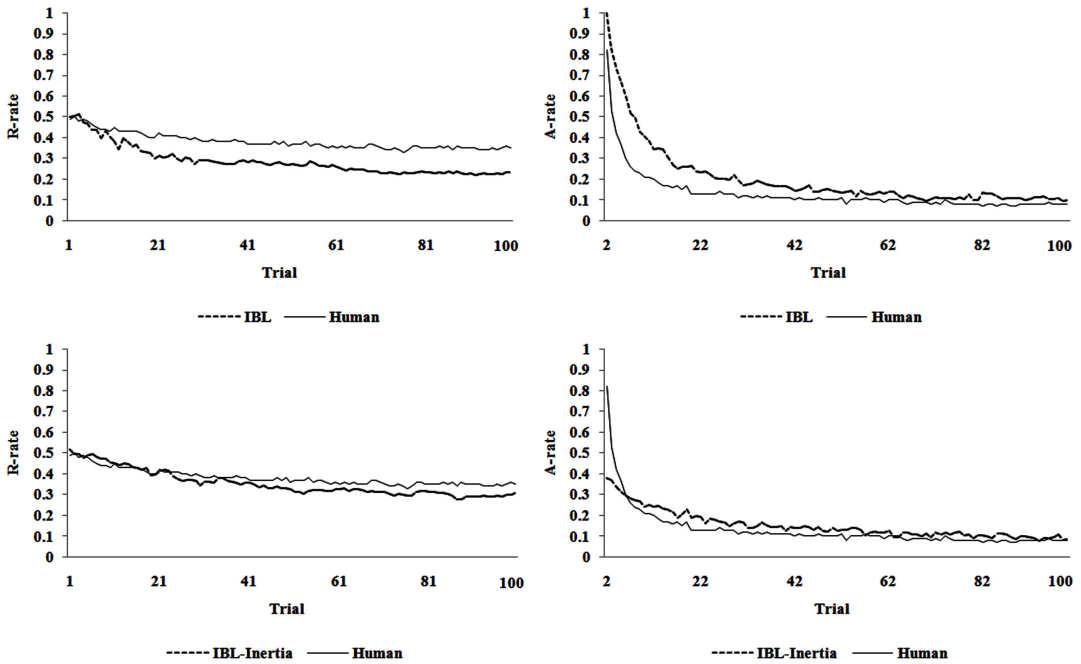


FIGURE 4 | The R-rate and A-rate over trials predicted by the IBL and IBL-Inertia models upon their generalization in the competition set. The R-rate and A-rate observed in human data in the competition set are also shown.

compared to that based upon inertia's current formulation. Overall, these observations indicate that there are many aspects still left in the literature to explore. For example, it is unclear whether people exhibit inertia after receiving both rewards and punishments. Although inertia has been defined as the tendency to repeat the last choice irrespective of the obtained outcomes (Biele et al., 2009), it is clear that inertia needs to be defined more precisely. Some researchers have argued inertia as an irrational behavior in which individuals hold onto choices that clearly do not provide the maximizing outcome for too long (Sandri et al., 2010). Inertia has also been portrayed as desirable, however, as it believed to be a key component of love, trust, and friendship in the real world (Cook et al., 2005). Even when we consider inertia as we defined it in this paper, it may be the result of strong preferences for the high outcomes or the result of an apparently irrational behavior of holding on too long to non-maximizing (low) outcomes. As part of our future research, we propose to define the reasons for inertia more precisely by investigating its relationship with the exploration of alternatives due to the nature of outcomes, high or low. One way we may do this analysis is by controlling for the nature of rewards or punishments received after a decision choice and by evaluating its effects on repeating the last choice as the current decision. Also, we would like to consider the alternation behaviors of individuals depending upon the nature of rewards or punishments received by them in the last trial.

Finally, as part of future research, we would also like to compare the different formulations of inertia in computational models of DFE. As detailed above, there have been at least three different inertia formulations proposed: A random variation across trials (Gonzalez and Dutt, 2011), a function of surprise determined by the difference in expected values (Erev et al., 2010a), and a function of surprise determined by the differences in blended values (Gonzalez et al., 2011). Which one of these formulations performs best in different DFE tasks? How well do these different formulations account for the over trial trends in the R-rates and A-rates? Still, how are these formulations impacted by task complexity: by the nature and number of outcomes on each alternative, and the nature of the probability distribution of outcomes on each alternative? These are also some important questions that we would like to attend to as part of future research.

ACKNOWLEDGMENTS

This research is supported by the Defense Threat Reduction Agency (DTRA) grant number: HDTRA1-09-1-0053 to Dr. Cleotilde Gonzalez. We thank Ms. Hau-Yu Wong and Dr. Noam Ben-Asher of the Dynamic Decision Making Laboratory for their help in proofreading this manuscript and providing insightful comments. We would also like to thank Dr. Ido Erev of the Technion-Israel Institute of Technology for making the data from the Technion Prediction Tournament available.

REFERENCES

Ahn, W. Y., Busemeyer, J. R., Wagenmakers, E. J., and Stout, J. C. (2008). Comparison of decision learning models using the generalization criterion method. *Cogn. Sci.* 32, 1376–1402.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., and Qin, Y. (2004). An integrated theory of the mind. *Psychol. Rev.* 111, 1036–1060.

Anderson, J. R., and Lebiere, C. (1998). *The Atomic Components of Thought.* Mahwah, NJ: Erlbaum.

Anderson, J. R., and Lebiere, C. (2003). The Newell test for a theory of mind. *Behav. Brain Sci.* 26, 587–639.

Barron, G., and Erev, I. (2003). Small feedback-based decisions and their limited correspondence to description-based decisions. *J. Behav. Decis. Mak.* 16, 215–233.

Barron, G., Leider, S., and Stack, J. (2008). The effect of safe experience on a warnings' impact: sex, drugs, and rock-n-roll. *Organ. Behav. Hum. Decis. Process.* 106, 125–142.

Biele, G., Erev, I., and Ert, E. (2009). Learning, risk attitude and hot stoves in restless bandit problems. *J. Math. Psychol.* 53, 155–167.

Börgers, T., and Sarin, R. (2000). Naive reinforcement learning with endogenous aspirations. *Int. Econ. Rev.* 41, 921–950.

Busemeyer, J. R., and Diederich, A. (2009). *Cognitive Modeling.* Thousand Oaks: Sage Publications, Inc.

Busemeyer, J. R., and Wang, Y. (2000). Model comparisons and model selections based on the generalization criterion methodology. *J. Math. Psychol.* 44, 171–189.

Chen, W., Liu, S.-Y., Chen, C.-H., and Lee, Y.-S. (2011). Bounded memory, inertia, sampling and weighting model for market entry games. *Games* 2, 187–199.

Cook, K. S., Hardin, R., and Levi, M. (2005). *Cooperation without Trust?* New York: Russell Sage Foundation Publications.

Dutt, V., Yamaguchi, M., Gonzalez, C., and Proctor, R. W. (2009). "An instance-based learning model of stimulus-response compatibility effects in mixed location-relevant and location-irrelevant tasks," in *Proceedings of the 9th International Conference on Cognitive Modeling – ICCM2009,* eds A. Howes, D. Peebles, and R. Cooper (Manchester: University of Huddersfield).

Erev, I., and Barron, G. (2005). On adaptation, maximization and reinforcement learning among cognitive strategies. *Psychol. Rev.* 112, 912–931.

Erev, I., Ert, E., and Roth, A. E. (2010a). A choice prediction competition for market entry games: an introduction. *Games* 1, 117–136.

Erev, I., Ert, E., Roth, A. E., Haruvy, E., Herzog, S. M., Hau, R., Hertwig, R., Stewart, T., West, R., and Lebiere, C. (2010b). A choice prediction competition: choices from experience and from description. *J. Behav. Decis. Mak.* 23, 15–47.

Estes, W. K., and Maddox, W. T. (2005). Risks of drawing inferences about cognitive processes from model fits to individual versus average performance. *Psychon. Bull. Rev.* 12, 403–408.

Gladwell, M. (2007). *Blink.* New York: Back Bay Books.

Gonzalez, C., Best, B. J., Healy, A. F., Bourne, L. E. Jr., and Kole, J. A. (2010). A cognitive modeling account of simultaneous learning and fatigue effects. *Cogn. Syst. Res.* 12, 19–32.

Gonzalez, C., and Dutt, V. (2011). Instance-based learning: integrating sampling and repeated decisions from experience. *Psychol. Rev.* 118, 523–551.

Gonzalez, C., Dutt, V., and Lejarraga, T. (2011). A loser can be a winner: comparison of two instance-based learning models in a market entry competition. *Games* 2, 136–162.

Gonzalez, C., and Lebiere, C. (2005). "Instance-based cognitive models of decision making," in *Transfer of Knowledge in Economic Decision-Making,* eds D. Zizzo and A. Courakis (New York: Palgrave Macmillan), 148–165.

Gonzalez, C., Lerch, F. J., and Lebiere, C. (2003). Instance-based learning in real-time dynamic decision making. *Cogn. Sci.* 27, 591–635.

Hertwig, R., and Erev, I. (2009). The description-experience gap in risky choice. *Trends Cogn. Sci. (Regul. Ed.)* 13, 517–523.

Hodgkinson, G. P. (1997). Cognitive inertia in a turbulent market: the case of UK residential estate agents. *J. Manage. Stud.* 34, 921–945.

Kline, R. B. (2004). *Principles and Practice of Structural Equation Modeling (Methodology in the Social Sciences),* 2nd Edn. New York: The Guilford Press.

Lebiere, C. (1999). "Blending," in *Proceedings of the Sixth ACT-R Workshop* (Fairfax, VA: George Mason University).

Lebiere, C., Gonzalez, C., and Martin, M. (2007). "Instance-based decision making model of repeated binary choice," in *Proceedings of the 8th International Conference on Cognitive Modeling,* Ann Arbor, MI.

Lejarraga, T., Dutt, V., and Gonzalez, C. (2012). Instance-based learning: a general model of repeated binary choice. *J. Behav. Decis. Mak.* 25, 143–153.

Loehlin, J. C. (2003). *Latent Variable Models: An Introduction to Factor, Path, and Structural Equation Analysis.* Mahwah, NJ: Lawrence Erlbaum.

Martin, M. K., Gonzalez, C., and Lebiere, C. (2004). "Learning to make decisions in dynamic environments: ACT-R plays the beer game," in *Proceedings of the Sixth International Conference on Cognitive Modeling* (Pittsburgh, PA: Carnegie Mellon University), 178–183.

Maruyama, G. M. (1997). *Basics of Structural Equation Modeling.* Thousand Oaks, CA: Sage Publications, Inc.

Nevo, I., and Erev, I. (2012). On surprise, change, and the effect of recent outcomes. *Front. Psychol.* 3:24. doi:10.3389/fpsyg.2012.00024

Pitt, M. A., and Myung, I. J. (2002). When a good fit can be bad. *Trends Cogn. Sci. (Regul. Ed.)* 6, 421–425.

Rapoport, A., Erev, I., Abraham, E. V., and Olson, D. E. (1997). Randomization and adaptive learning in a simplified poker game. *Organ. Behav. Hum. Decis. Process.* 69, 31–49.

Reger, R. K., and Palmer, T. B. (1996). Managerial categorization of competitors: using old maps to navigate new environments. *Organ. Sci.* 7, 22–39.

Samuelson, L. (1994). Stochastic stability in games with alternative best replies. *J. Econ. Theory* 64, 35–65.

Sandri, S., Schade, C., Mußhoff, O., and Odening, M. (2010). Holding on for too long? An experimental study on inertia in entrepreneurs' and non-entrepreneurs' disinvestment choices. *J. Econ. Behav. Organ.* 76, 30–44.

Stone, M. (1977). Asymptotics for and against cross-validation. *Biometrika* 64, 29–35.

Tripsas, M., and Gavetti, G. (2000). Capabilities, cognition, and inertia: evidence from digital imaging. *Strateg. Manage. J.* 21, 1147–1161.

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 February 2012; accepted: 16 May 2012; published online: 06 June 2012.

Citation: Dutt V and Gonzalez C (2012) The role of inertia in modeling decisions from experience with instance-based learning. Front. Psychology ***3****:177. doi: 10.3389/fpsyg.2012.00177*

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

APPENDIX

IBL MODEL EQUATIONS

Inertia mechanism

A choice is made in the model in trial $t+1$ as:

If

The draw of a random value in the uniform distribution U (0, 1) $<$ *pInertia*,

Then

Repeat the choice as made in the previous trial

Else

Select an alternative with the highest blended value as per Eq. A2 (below) (A1)

The *pInertia* parameter could vary between 0 and 1, and it does not change across trials or participants.

Blending and activation mechanisms

The blended value of alternative j is defined as

$$V_j = \sum_{i=1}^{n} p_i x_i \tag{A2}$$

Where x_i is the value of the observed outcome in the outcome slot of an instance i corresponding to the alternative j, and p_i is the probability of that instance's retrieval from memory (for the case of our binary-choice task in the experience condition, the value of j in Eq. A2 could be either risky or safe). The blended value of an alternative is the sum of all observed outcomes x_i in the outcome slot of corresponding instances, weighted by the instances' probability of retrieval.

Probability of retrieving instances

In any trial t, the probability of retrieving instance i from memory is a function of that instance's activation relative to the activation of all other instances corresponding to that alternative, given by

$$p_{i,t} = \frac{e^{A_{i,t}/\tau}}{\sum_j e^{A_{j,t}/\tau}} \tag{A3}$$

Where τ is random noise defined as $s \times \sqrt{2}$, and s is a free noise parameter. The noise parameter s captures the imprecision of retrieving instances from memory.

Activation of instances

The activation of each instance in memory depends upon the *activation* mechanism originally proposed in ACT-R (Anderson and Lebiere, 2003). According to this mechanism, for each trial t, *activation* $A_{i,t}$ of instance is:

$$A_{i,t} = \ln\left(\sum_{t_i \in \{1,\ldots,t-1\}} (t - t_i)^{-d}\right) + s \times 1n\left(\frac{1-\gamma_{i,t}}{\gamma_{i,t}}\right) \tag{A4}$$

Where d is a free decay parameter, and t_i is a previous trial when the instance i was created or its activation was reinforced due to an outcome observed in the task (the instance i is the one that has the observed outcome as the value in its outcome slot). The summation will include a number of terms that coincides with the number of times an outcome has been observed in previous trials and the corresponding instance i's activation that has been reinforced in memory (by encoding a timestamp of the trial t_i). Therefore, the activation of an instance corresponding to an observed outcome increases with the frequency of observation and with the recency of those observations. The decay parameter d affects the activation of an instance directly, as it captures the rate of forgetting or reliance on recency.

Noise in activation

The $\gamma_{i,t}$ term is a random draw from a uniform distribution U (0, 1), and the $s \times \ln\left(\frac{1-\gamma_{i,t}}{\gamma_{i,t}}\right)$ term represents Gaussian noise important for capturing the variability of human behavior.

Definition of Akaike information criterion

$$\text{AIC} = G^2 + 2*k \tag{A5}$$

$$G^2 = t * \ln\frac{\text{SSE}}{t} \tag{A6}$$

$$\text{SSE} = \sum_{i=1}^{t} \left(x_{\text{model},i} - x_{\text{human},i}\right)^2 \tag{A7}$$

Where, G^2 is defined as the lack of fit between model and human data (Gonzalez and Dutt, 2011). Furthermore, the $x_{\text{model},i}$ and $x_{\text{human},i}$ refer to the average dependent measure (e.g., average R-rate or A-rate) in the model and human data over t trials of a task ($t = 100$ for the R-rate and $t = 99$ for the A-rate). The average in the dependent measure (R-rate or A-rate) has been taken over all problems and participants. The SSE is the sum of squared errors between human and model datasets that is calculated for the average dependent measure (A-rate or R-rate). The mean-squared error (MSE) is defined as SSE/100 for the R-rate measure and SSE/99 for the A-rate measure. The t is the number of trials in the task, and k is the number of parameters in the model. The AIC in its formulation incorporates both the effect of an MSD (the G^2 term) as well as the number of free parameters in a model (the $2 * k$ term). The smaller the value of AIC, the better the respective model is.

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 29 June 2012
doi: 10.3389/fpsyg.2012.00215

Temporal dynamics of hypothesis generation: the influences of data serial order, data consistency, and elicitation timing

Nicholas D. Lange[1,2]*, Rick P. Thomas[2] and Eddy J. Davelaar[1]

[1] *Department of Psychological Sciences, Birkbeck College, University of London, London, UK*
[2] *Department of Psychology, University of Oklahoma, Norman, OK, USA*

Edited by:
David Albert Lagnado, University College London, UK

Reviewed by:
Richard P. Cooper, Birkbeck College, UK
Daniel Navarro, University of Adelaide, Australia
Amber M. Sprenger, University of Maryland College Park, USA

***Correspondence:**
Nicholas D. Lange, Department of Psychological Sciences, Birkbeck College, University of London, Malet Street, London WC1E 7HX, UK.
e-mail: ndlange@gmail.com

The pre-decisional process of hypothesis generation is a ubiquitous cognitive faculty that we continually employ in an effort to understand our environment and thereby support appropriate judgments and decisions. Although we are beginning to understand the fundamental processes underlying hypothesis generation, little is known about how various temporal dynamics, inherent in real world generation tasks, influence the retrieval of hypotheses from long-term memory. This paper presents two experiments investigating three data acquisition dynamics in a simulated medical diagnosis task. The results indicate that the mere serial order of data, data consistency (with previously generated hypotheses), and mode of responding influence the hypothesis generation process. An extension of the HyGene computational model endowed with dynamic data acquisition processes is forwarded and explored to provide an account of the present data.

Keywords: hypothesis generation, temporal dynamics, working memory, information acquisition, decision making

Hypothesis generation is a pre-decisional process by which we formulate explanations and beliefs regarding the occurrences we observe in our environment. The hypotheses we generate from long-term memory (LTM) bring structure to many of the ill-structured decision making tasks we commonly encounter. As such, hypothesis generation represents a fundamental and ubiquitous cognitive faculty on which we constantly rely in our day-to-day lives. Given the regularity with which we employ this process, it is no surprise that hypothesis generation forms a core component of several professions. Auditors, for instance, must generate hypotheses regarding abnormal financial patterns, mechanics must generate hypotheses concerning car failure, and intelligence analysts must interpret the information they receive. Perhaps the clearest example, however, is that of medical diagnosis. A physician observes a pattern of symptoms presented by a patient (i.e., data) and uses this information to generate likely diagnoses (i.e., hypotheses) in an effort to explain the patient's presenting symptoms. Given these examples, the importance of developing a full understanding of the processes underlying hypothesis generation is clear, as the consequences of impoverished or inaccurate hypothesis generation can be injurious.

Issues of temporality pervade hypothesis generation and its underlying information acquisition processes. Hypothesis generation is a task situated at the confluence of external environmental dynamics and internal cognitive dynamics. External dynamics in the environment dictate the manifestation of the information we acquire and use as cues to retrieve likely hypotheses from LTM. Internal cognitive dynamics then determine how this information is used in service of the generation process and how the resulting hypotheses are maintained over the further course of time as judgments and decisions are rendered. Additionally, these further internal processes are influenced by and interact with the ongoing environmental dynamics as new information is acquired. These complicated interactions govern the beliefs (i.e., hypotheses) we entertain over time. It is likely that these factors interact in such a manner that would cause the data acquisition process to deviate from normative prescriptions.

Important to the present work is the fact that data acquisition generally occurs serially over some span of time. This, in turn, dictates that individual pieces of data are acquired in some relative temporal relation to one another. These constraints, individual data acquisition over time and the relative ordering of data, are likely to have significant consequences for hypothesis generation processes. Given these basic constraints, it is intuitive that temporal dynamics must form an integral part of any comprehensive account of hypothesis generation processes. At present there exists only a scant amount of data concerning the temporal dynamics of hypothesis generation. Thus, the influences of the constraints operating over these processes are not yet well understood. Until such influences are addressed more deeply at an empirical and theoretical level, a full understanding of hypothesis generation processes will remain speculative.

The empirical paradigm used in the following experiments is a simulated diagnosis task comprised of two main phases. The first phase represents a form of category learning in which the participant learns the conditional probabilities of medical symptoms (i.e., data) and fictitious diseases (i.e., hypotheses), from experience over time by observing a large sample of hypothetical pre-diagnosed patients. The second phase of the task involves presenting symptoms to the participant whose task it is to generate

(i.e., retrieve) likely disease states from memory. At a broader level, such experiments involving a learning phase followed by a decision making phase have been utilized widely in previous experiments (e.g., McKenzie, 1998; Cooper et al., 2003; Nelson et al., 2010; Sprenger and Dougherty, 2012). In the to-be-presented experiments, we presented the symptoms sequentially and manipulated the symptom's sequence structures in the "decision making phase." As the data acquisition unfolds over time, the results of these experiments provide insight into the dynamic data acquisition and hypothesis generation processes operating over time that are important for computational models.

In this paper, we present a novel extension of an existing computational model of hypothesis generation. This extension is designed to capture the working memory dynamics operating during data acquisition and how these factors contribute to the process of hypothesis generation. Additionally, two experiments exploring three questions of interest to dynamic hypothesis generation are described whose results are captured by this model. Experiment 1 utilized an adapted generalized order effects paradigm to assess how the serial position of an informative piece of information (i.e., a diagnostic datum), amongst uninformative information (i.e., non-diagnostic data), influences its contribution to the generation process. Experiment 2 investigated (1) how the acquisition of data inconsistent with previously generated hypotheses influences further generation and maintenance processes and (2) if generation behavior differs when it is based on the acquisition of a set of data vs. when those same pieces of data are acquired in isolation and generation is carried out successively as each datum is acquired. This distinction underscores different scenarios in which it is advantageous to maintain previously acquired data vs. previously generated hypotheses over time.

HYGENE: A COMPUTATIONAL MODEL OF HYPOTHESIS GENERATION

HyGene (Thomas et al., 2008; Dougherty et al., 2010), short for hypothesis generation, is a computational architecture addressing hypothesis generation, evaluation, and testing. This framework has provided a useful account through which to understand the cognitive mechanisms underlying these processes. This process model is presented in **Figure 1**.

HyGene rests upon three core principles. First, as underscored by the above examples, it is assumed that hypothesis generation represents a generalized case of cued recall. That is, the data observed in the environment (D_{obs}), which one would like to explain, act as cues prompting the retrieval of hypotheses from LTM. For instance, when a physician examines a patient, he/she uses the symptoms expressed by the patient as cues to related experiences stored in LTM. These cues activate a subset of related memories from which hypotheses are retrieved. These retrieval processes are indicated in Steps 1, 2, and 3 shown in **Figure 1**. Step 1 represents the environmental data being matched against episodic memory. In step 2, the instances in episodic memory that are highly activated by the environmental data contribute to the extraction of an *unspecified probe* representing a prototype of these highly activated episodic instances. This probe is then matched against all known hypotheses in semantic memory as indicated in Step 3. Hypotheses are then sampled into working memory based on their activations resulting from this semantic memory match.

As viable hypotheses are retrieved from LTM, they are placed in the Set of Leading Contenders (SOC) as demonstrated in Step 4. The SOC represents HyGene's working memory construct to which HyGene's second principle applies. The second principle holds that the number of hypotheses that can be maintained at one time is constrained by cognitive limitations (e.g., working memory capacity) as well as task characteristics (e.g., divided attention, time pressure). Accordingly, the more working memory resources that one has available to devote to the generation and maintenance of hypotheses, the greater the number of additional hypotheses can be placed in the SOC. Working memory capacity places an upper bound on the amount of hypotheses and data that one will be able to maintain at any point in time. In many circumstances, however, attention will be divided by a secondary task. Under such conditions this upper bound is reduced as the alternative task siphons resource that would otherwise allow the population of the SOC to its unencumbered capacity (Dougherty and Hunter, 2003a,b; Sprenger and Dougherty, 2006; Sprenger et al., 2011).

The third principle states that the hypotheses maintained in the SOC form the basis from which probability judgments are derived and provide the basis from which hypothesis testing is implemented. This principle underscores the function of hypothesis generation as a pre-decisional process underlying higher-level decision making tasks. The tradition of much of the prior research on probability judgment and hypothesis testing has been to provide the participant with the options to be judged or tested. HyGene highlights this as somewhat limiting the scope of the conclusions drawn from such procedures, as decision makers in real world tasks must generally generate the to-be-evaluated hypotheses themselves. As these higher-level tasks are contingent upon the output of the hypothesis generation process, any conclusions drawn from such experimenter-provided tasks are likely limited to such conditions.

HYPOTHESIS GENERATION PROCESSES IN HYGENE

The representation used by HyGene was borrowed from the multiple-trace global matching memory model MINERVA II (Hintzman, 1986, 1988) and the decision making model MINERVA-DM (Dougherty et al., 1999)[1]. Memory traces are represented in the model as a series of concatenated minivectors arbitrarily consisting of 1, 0, and −1 s where each minivector represents either a hypothesis or a piece of data (i.e., a feature of the memory). Separate episodic and semantic memory stores are present in HyGene which are made up of separate instances of such concatenated feature minivectors. While semantic memory contains prototypes of each disease, episodic memory contains individual traces for every experience the model acquires.

Retrieval is initiated when D_{obs} are matched against each of data minivectors in episodic LTM. This returns an LTM activation value for each trace in episodic LTM whereby greater overlap of features present in the trace and present in the D_{obs} results in greater activation. A threshold is applied to these episodic

[1]For a more thorough treatment of HyGene's computational architecture please see Thomas et al. (2008) or Dougherty et al. (2010).

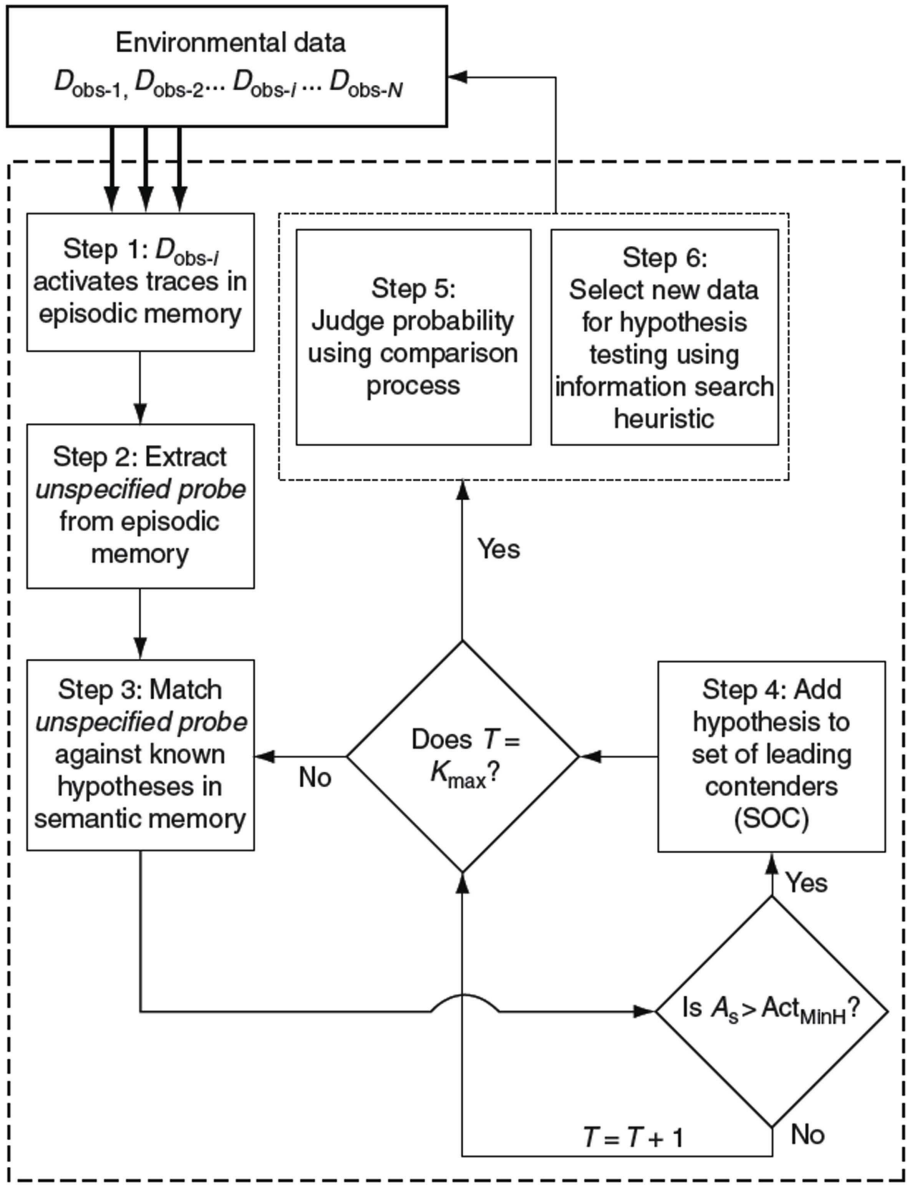


FIGURE 1 | Flow diagram of the HyGene model of hypothesis generation, judgment, and testing. A_s, semantic activation of retrieved hypothesis; Act_{MinH}, minimum semantic activation criterion for placement of hypothesis in SOC; *T*, total number of retrieval failures; and K_{max}, number of retrieval failures allowed before terminating hypothesis generation.

activation values such that only traces with long-term episodic activation values exceeding this threshold contribute to additional processing in the model. A prototype is extracted from this subset of traces which is then used as a cue to semantic memory for the retrieval of hypotheses. We refer to this cue as the *unspecified probe*. This unspecified probe is matched against all hypotheses in semantic memory which returns an activation value for each known hypothesis. The activation values for each hypothesis serve as input into retrieval through sampling via Luce's choice rule. Generation proceeds in this way until a stopping rule is reached based on the total number of resamples of previously generated hypotheses (i.e., retrieval failures).

In its current form, the HyGene model is *static* with regards to data acquisition and utilization. The model receives all available data from the environment simultaneously and engages in only a single iteration of hypothesis generation. Given the static nature of the model, each piece of data used to cue LTM contributes equally to the recall process. Based on effects observed in related domains, however, it seems reasonable to suspect that all available data do not contribute equally in hypothesis generation tasks. For example, Anderson (1965), for instance, observed primacy weightings in an impression formation task in which attributes describing a person were revealed sequentially. Moreover, recent work has demonstrated biases in the serial position of data used to support

hypothesis generation tasks (Sprenger and Dougherty, 2012). By ignoring differential use of available data in the generation process, HyGene, as previously implemented, ignores temporal dynamics influencing hypothesis generation tasks. In our view, what is needed is an understanding of working memory dynamics as data acquisition, hypothesis generation, and maintenance processes unfold and evolve over time in hypothesis generation tasks.

DYNAMIC WORKING MEMORY BUFFER OF THE CONTEXT-ACTIVATION MODEL

The context-activation model of memory (Davelaar et al., 2005) is one of the most comprehensive models of memory recall to date. It is a dual-trace model of list memory accounting for a large set of data from various recall paradigms. Integral to the model's behavior are the activation-based working memory dynamics of its buffer. The working memory buffer of the model dictates that the activations of the items in working memory systematically fluctuate over time as the result of competing processes described by Eq. 1.

$$x_i\,(t+1) = \lambda x_i\,(t) + (1-\lambda)\,\{\alpha F[x_i\,(t)] + I_i\,(t) - \beta \sum_{j \neq i} F[x_j\,(t)] + N(0,\sigma)\} \quad (1)$$

Equation 1: activation calculation of the context-activation model

The activation level of each item, x_i, is determined by the item's activation on the previous time step, self-recurrent excitation that each item recycles onto itself α, inhibition from the other active items β, and zero-mean Gaussian noise N with standard deviation σ. Lastly, λ is the Euler integration constant that discretizes the differential equation. Note, however, that as this equation is applied in the present model, noise was only applied to an item's activation value once it was presented to the model[2].

Figure 2 illustrates the interplay between the competitive buffer dynamics in a noiseless run of the buffer when four pieces of data have been presented to the model successively. The activation of each datum rises as it is presented to the model and its bottom-up sensory input contributes to the activation. These activations are then dampened in the absence of bottom-up input as inhibition from the other items drive activation down. Self-recurrency can keep an item in the buffer in the absence of bottom-up input, but this ability is in proportion to the amount of competition from other items in the buffer. The line at 0.2 represents the model's working memory threshold. In the combined dynamic HyGene model (utilizing the dynamics of the buffer to determine the weights of the data) this WM threshold separates data that are available to contribute to generation (>0.2) from those that will not (<0.2). That is, if a piece of data's activation is greater than this threshold at the time of generation then it contributes to the retrieval of hypotheses from LTM and is weighted by its amount of activation. However, if, on the other hand, a piece of data falls below the WM threshold then it is weighted zero and as a result does not contribute to the hypothesis retrieval.

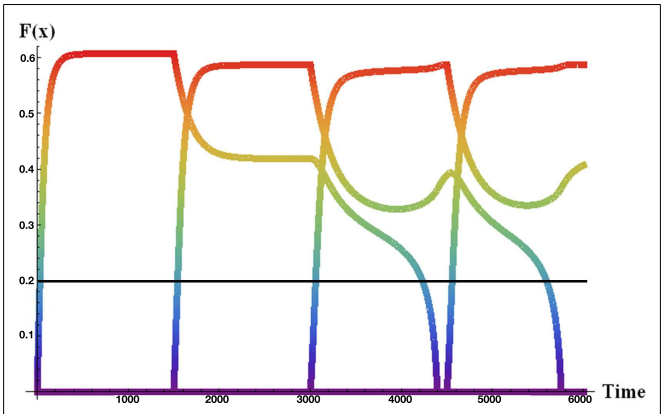


FIGURE 2 | Noiseless activation trajectories for four sequentially received data in the dynamic activation-based buffer. Each item presented to the buffer for 1500 iterations. $F(x)$ = memory activation.

The activations of individual items are sensitive the amount of recurrency (alpha) and inhibition (beta) operating in the buffer. **Figure 3** demonstrates differential sensitivity to values of alpha and beta by item presentation serial position (1 through 4 in this case). This plot was generated by running the working memory buffer across a range of alpha and beta values for 50 runs at each parameter combination. Each panel presents the activation of an item in a four-item sequence after the final item has been presented. The activation levels vary with serial position, as shown by the differences among the four panels and with the value of the alpha and beta parameters, as shown within each panel. It can be seen that items one and two are mainly sensitive to the value of alpha. As alpha is increased, these items are more likely to maintain high activation values at the end of the data presentation. Item three demonstrates a similar pattern under low values of beta, but under higher values of beta this item only achieves modest activation as it cannot overcome the strong competition exerted by item one and two. Item four demonstrates a pattern distinct from the others. Like the previous three items the value of alpha limits the influence of beta up to a certain point. At moderate to high values of alpha, however, beta has a large impact on the activation value of the fourth item. At very low values of beta (under high alpha) this item is able to attain high activation, but quickly moves to very low activation values with modest increases in beta. These modest increases in beta are enough to make the competition from the three preceding items severe enough that the fourth item cannot overcome it.

Taken as a whole, these plots describe differences in the activation gradients (profiles of activation across all four items) taken on by the buffer across various values of alpha and beta. For instance, the stars in the plot represent two different settings of alpha and beta which result in different activation gradients across the items. The settings of alpha = 2 and beta = 0.2 represented by the white stars, for instance, represent an instance of recency in the item activations. That is, the earlier items have only slight activation,

[2]This was done from the pragmatic view that the buffer cannot apply noise to an item representation that does not yet exist in the environment or in the system. A full and systematic analysis of how this assumption affects the behavior of the buffer has not been carried out as of yet, but in the context of the current simulations preliminary analysis suggests that this change affects the activation values produced by the buffer only slightly.

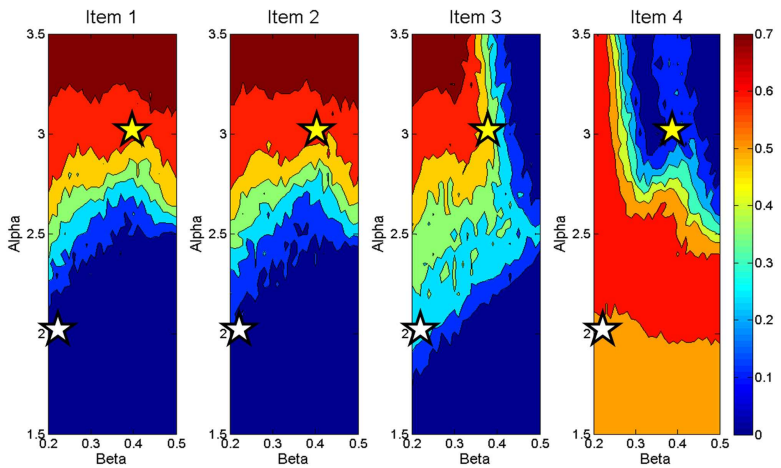


FIGURE 3 | Contour plot displaying activation values of four items at end of data presentation across a range of Beta (*X* axes) and Alpha (*Y* axes) demonstrating differences in activation weight gradients produced by the working memory buffer.

the third item modest activation, and the last item is highly active relative to the others. Tracing the activations across the settings of alpha = 3 and beta = 0.4 represented by the yellow stars, on the other hand, shows a primacy gradient in which the earlier items are highly active, item three is less so, and the last item's activation is very low. As will be seen, this pattern of activation values across different values of alpha and beta will become important for the computational account of Experiment 2. At a broader level, however, this plot shows possible activation gradients that can be obtained with the working memory buffer. In general, the activation gradients produce recency, but primacy gradients are also possible. Additionally, there are patterns of activation across items that the buffer cannot produce. For instance an inverted U shape of item activations would not result from the buffer's processes.

These dynamics are theoretically meaningful as they produce data patterns which item-based working memory buffers (e.g., SAM; Raaijmakers and Shiffrin, 1981) cannot account for. For example, the buffer dynamics of the context-activation model dictate that items presented early in a sequence will remain high in activation (i.e., remain in working memory) under fast presentation rates. That is, under fast presentation rates the model predicts a primacy effect. Such effects have been observed in cued recall (Davelaar et al., 2005), free recall (Usher et al., 2008), and in a hypothesis generation task (Lange et al., 2012). Given these findings and the unique ability of the activation-based buffer to account for these effects, we have selected the activation-based buffer as our starting point for endowing the HyGene model with dynamic data acquisition processes.

A DYNAMIC MODEL OF HYPOTHESIS GENERATION: ENDOWING HYGENE WITH DYNAMIC DATA ACQUISITION

The competitive working memory processes of the context-activation model's dynamic buffer provide a principled means for incorporating fine-grained temporal dynamics into currently static portions of HyGene. As a first step in incorporating the dynamic working memory processes of the working memory buffer, we use the buffer as a means to endow HyGene with dynamic data acquisition. In so doing, the HyGene architecture gains two main advantages. As pointed out by Sprenger and Dougherty (2012), any model of hypothesis generation seeking to account for situations in which data are presented sequentially needs a means of weighting the contribution of individual data. In using the buffer's output as weights on the generation process we provide such a weighting mechanism. Additionally, as a natural consequence of utilizing the buffer to provide weights on data observed in the environment, working memory capacity constraints are imposed on the amount of data that can contribute to the generation process. As data acquisition was not a focus of the original instantiation of HyGene, capacity limitations in this part of the generation process were not addressed. However, recent data suggest that capacity constraints operating over data acquisition influence hypothesis generation (Lange et al., 2012). Lastly, at a less pragmatic level, this integration provides insight into the working memory dynamics unfolding throughout the data acquisition period thereby providing a window into processing occurring over this previously unmodeled epoch of the hypothesis generation process.

In order to endow HyGene with dynamic data acquisition, each run of the model begins with the context-activation model being sequentially presented with a series of items. In the context of this model these items are the environmental data the model has observed. The activation values for each piece of data at the end of the data acquisition period are then used as the weights on the generation process. A working memory threshold is imposed on the data activations such that data with activations falling below 0.2 are weighted with a zero rather than their actual activation value[3]. Specifically, the global memory match performed between the current D_{obs} and episodic memory in HyGene is weighted by the individual item activations in the dynamic working memory buffer (with the application of the working memory threshold).

[3]This working memory threshold has been carried over from the context-activation model as it proved valuable for that model's account of data from a host of list recall paradigms (Davelaar et al., 2005).

As each trace in HyGene's episodic memory is made up of concatenated minivectors, each representing a particular data feature (e.g., fever vs. normal temperature), this weighting is applied in a feature by feature manner in the global matching process. From this point on in the model everything operates in accordance with the original instantiation of HyGene. That is, a subset of the highly activated traces in episodic memory is then used as the basis for the extraction of the *unspecified probe*. This probe is then matched against semantic memory from which hypotheses are serially retrieved into working memory for further processing.

In order to demonstrate how the integrated dynamic HyGene model responds to variation in the buffer dynamics a simulation was run in which alpha and beta were manipulated at the two levels highlighted above in **Figure 3**. In this simulation, the model was sequentially presented with four pieces of data. Only one of these pieces of data was diagnostic whereas the remaining three were completely non-diagnostic. An additional independent variable in this simulation was the serial position in which the diagnostic piece of data was placed. Displayed in **Figure 4** is the model's generation of the most likely hypothesis (i.e., the hypothesis suggested by the diagnostic piece of data) across that data's serial position plotted by the two levels of alpha (recurrent activation) and beta (global lateral inhibition). What this plot demonstrates, in effect, is how the contribution of each data's serial position to the model's generation process is influenced by alpha and beta. As displayed on the left side of the plot, at the lower value of alpha there are clear recency effects. This is due to the buffer dynamics which under these settings predict an "early in – early out" cycling of items through the buffer as shown in **Figure 2**. The recency effects emerge as earlier data are less likely to reside in the buffer at the time of generation than later data. It should be noted that these parameters (alpha = 2, beta = 0.2) have been used in previous work accounting for the data from multiple list recall paradigms (Davelaar et al., 2005). By means of preview, we utilize the model's prediction of recency under these standard parameter settings in guiding our expectations and the implementation of Experiment 1.

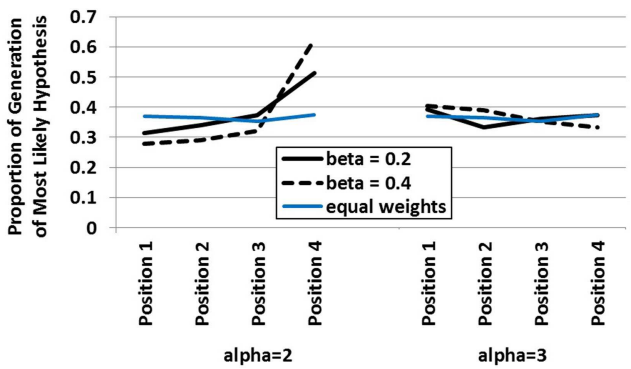


FIGURE 4 | Influence of data serial position on the hypothesis generation behavior of the dynamic HyGene model at two levels of alpha and beta (and the performance of an equal weighted model in blue). Data plotted represents the proportion of simulation runs on which the most likely hypothesis was generated.

Under the higher value of alpha however, recency does not obtain. In this case, the serial position function flattens substantially as the increased recurrency allows more items to be available to contribute to generation at the end of the sequential data presentation. That is, even when the diagnostic datum appears early, it is maintained long enough in the buffer to be incorporated into the cue to episodic memory. Under the higher value of beta, we see this flattening out transition to a mild primacy gradient. This results from the increased inhibition making it more difficult for the later items to gain enough activation in working memory to contribute to the retrieval process. The greater amount of inhibition essentially renders the later items uncompetitive as they face more competition than they are able, in general, to overcome. **Figure 4** additionally plots a line in blue demonstrating the generation level of the static HyGene model in which, rather than utilizing the weights produced by the buffer, each piece of data was weighted equally with a value of one. It can be seen that this line of performance is intermediate under low alpha, but somewhat consistent with the high alpha condition in which more data contribute to the generation process more regularly.

EXPERIMENT 1: DATA SERIAL POSITION

Order effects are pervasive in investigations of memory and decision making (Murdock, 1962; Weiss and Anderson, 1969; Hogarth and Einhorn, 1992; Page and Norris, 1998). Such effects have even been obtained in a hypothesis generation task specifically. Although observed under different conditions than addressed by the present experiment, Sprenger and Dougherty, 2012, Experiments 1 and 3) found that people sometimes tend to generate hypotheses suggested by more recent cues.

The generalized order effect paradigm was developed by Anderson (1965, 1973) and couched within the algebra of information integration theory to derive weight estimates for individual pieces of information presented in impression formation tasks (e.g., adjectives describing a person). This procedure involved embedding a fixed list of information with a critical piece of information at various serial positions. The differences in the serial position occupied by the piece of critical information thus defined the independent variable, and given that all other information was held constant between conditions, the differences in final judgment were attributable to this difference in serial position. The present experiment represents an adaptation of this paradigm to assess the impact of data serial position on hypothesis generation.

METHOD

Participants

Seventy-two participants from the University of Oklahoma participated in this experiment for course credit.

Design and procedure

The design of Experiment 1 was a one-way within-subjects design with symptom order as the independent variable. The statistical ecology for this experiment, as defined by the conditional probabilities between the various diseases and symptoms, is shown in **Table 1**. Each of the values appearing in this table represents the probability that the symptom will be positive (e.g., fever) given the disease [where the complementary probability represents the probability of the symptom being negative (e.g., normal

temperature) given the disease]. The only diagnostic (i.e., informative) symptom is S1 whereas the remaining symptoms, S2–S4, are non-diagnostic (uninformative).

Table 2 displays the four symptom orders. Each of these orders was identical (S2 → S3 → S4) except for the position of S1 within them. All participants received and judged all four symptom orders.

There were three main phases to the experiment, an exemplar training phase to learn the contingencies displayed in **Table 1**, a learning test to allow discrimination of participants that had learned in the training from those that had not, and an elicitation phase in which the symptom order manipulation was applied in a diagnosis task in which the patient's symptoms were presented sequentially. The procedure began with the exemplar training phase in which a series of hypothetical pre-diagnosed patients was presented to the participant in order for them to learn, through experience, the contingencies between the diseases and symptoms. Each of these patients was represented by a diagnosis at the top of the screen and a series of test results (i.e., symptoms) pertaining to the columns of S1, S2, S3, and S4 as can be seen in the example displayed by **Figure 5**.

Each participant saw 50 exemplars of each disease for a total of 150 exemplars, thus making the base rates of the diseases equal. The specific results of these tests respected the probabilities in **Table 1**. The exemplars were drawn in blocks of 10 in which the symptoms would be drawn from the fixed distribution of symptom states given that disease. These symptom states were sampled independently without replacement from exemplar to exemplar. Therefore over the 10 exemplars presented in each individual disease block, the symptoms observed by the participant perfectly represented the distribution of symptoms for that disease. The disease blocks were randomly sampled without replacement which was repeated after the third disease block was presented. Thus, over the course of training the participants were repeatedly presented with the exact probabilities displayed in **Table 1**. Each exemplar appeared on the screen for a minimum of 5000 ms at which point they could continue studying the current exemplar or advance to the next exemplar by entering (on the keyboard) the first letter of the current disease exemplar. This optional prolonged studying made the training pseudo-self-paced. Prior to beginning the exemplar training phase, the participants were informed that they had an opportunity to earn a $5.00 gift card to Wal-Mart if they performed well enough in the task.

Table 1 | Disease × Symptom ecology of Experiment 1.

		Symptoms			
		S1	**S2**	**S3**	**S4**
Diseases	D1: Metalytis	0.8	0.6	0.6	0.6
	D2: Zymosis	0.2	0.6	0.6	0.6
	D3: Gwaronia	0.2	0.6	0.6	0.6

Values represent the probability of the symptom being positive (i.e., present) given the disease state. S1 was the only diagnostic symptom as indicated by the values in gray.

Table 2 | Symptom presentation orders used in Experiment 1.

	→Presentation position→			
	1	**2**	**3**	**4**
Order 1	**S1**	S2	S3	S4
Order 2	S2	**S1**	S3	S4
Order 3	S2	S3	**S1**	S4
Order 4	S2	S3	S4	**S1**

The diagnostic symptom, S1, appeared in different serial positions within each cue order condition as indicated in bold.

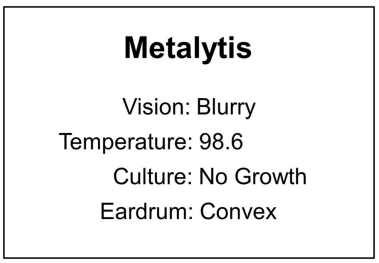


FIGURE 5 | Example exemplar used in Experiment 1.

The diagnosis test phase directly followed exemplar training. This test was included to allow discrimination of participants that learned the contingencies between the symptoms and the diseases in the training phase[4]. The participants were presented with the symptoms of a series of 12 patients (four of each disease) as defined principally by the presence or absence of S1. That is, four of the patients had S1 present (suffering from Metalytis) and the remaining eight had S1 absent (four suffering from Zymosis and four suffering from Gwaronia). The remaining symptoms for the four patients of each disease were the same across the three diseases. On one patient these symptoms were all positive. On the remaining three patients one of these symptoms (S2, S3, S4) was selected without replacement to be absent while the other two were present. Note that as S2, S3, and S4 were completely non-diagnostic as the presence or absence of their symptoms does not influence the likelihood of the disease state. The disease likelihood is completely dependent on the state of S1. The symptoms of each of the patients were presented simultaneously on a single screen. The participants' task was to correctly diagnose the patients with the disease of greatest posterior probability given their presenting symptoms. No feedback on this test performance was provided. As only S1 was diagnostic, the participants' scores on this test were tallied based on their correct discrimination of each patient as Metalytis vs. Gwaronia or Zymosis. There were 12 test patients in this diagnosis test. If the participant scored greater than 60% on

[4]Previous investigations in our lab utilizing exemplar training tasks have demonstrated variation in conclusions drawn from results conditionalized on such learning data against entire non-conditionalized data set. Therefore including this learning test allows us a check on the presence of such discrepancies in addition to obtaining data that may inform how greater or lesser learning influences the generation process.

a diagnosis test they were awarded the gift card at the end of the experiment[5]. Prior to the end of the experiment, the participants were not informed of their performance on the diagnosis test. The participant then completed a series of arithmetic distracters in order to clear working memory of information processed during the diagnosis test phase. The distracter task consisted of a series of 15 arithmetic equations for which the correctness or incorrectness was to be reported (e.g., $15/3 + 2 = 7$? Correct or Incorrect?). This distracter task was self-paced.

The elicitation phase then proceeded. First, the diagnosis task was described to the participants as follows: "You will now be presented with additional patients that need to be diagnosed. Each symptom of the patient will be presented one at a time. Following the last symptom you will be asked to diagnose the patient based on their symptoms. Keep in mind that sometimes the symptoms will help you narrow down the list of likely diagnoses to a single disease and other times the symptoms may not help you narrow down the list of likely diagnoses at all. It is up to you to determine if the patient is likely to be suffering from 1 disease, 2 diseases, or all 3 diseases. When you input your response make sure that you respond with the most likely disease first. You will then be asked if you think there is another likely disease. If you think so then you will enter the next most likely disease second. If you do not think there is another likely disease then just hit the Spacebar. You will then have the option to enter a third disease or hit the Spacebar in the same manner. To input the diseases you will use the first letter of the disease, just as you have been during the training and previous test."

The participant was then presented with the first patient and triggered the onset of the stream of symptoms themselves when they were ready. Each of the four symptoms was presented individually for 1.5 s with a 250 ms interstimulus interval following each symptom. The order in which the symptoms were presented was determined by the order condition as shown in **Table 2**. Additionally, all of the patient symptoms presented in this phase positive (i.e., present, as the values in **Table 2** represent the likelihood of the symptoms being present given the disease state). The Bayesian posterior probability of D1 was 0.67 whereas the posterior probability of either D2 or D3 was 0.17. Following the presentation of the last symptom the participant responded to two sets of prompts: the diagnosis prompts (as previously described in the instructions to the participants) and a single probability judgment of their highest ranked diagnosis. The probability judgment was elicited with the following prompt: "If you were presented 100 patients with the symptoms of the patient you just observed how many would have [INSERT HIGHEST RANKED DISEASE]?" The participant was then presented with the remaining symptom orders in the same manner with distracter tasks intervening between each trial. The first order received by each participant was randomized between participants and the sequence of the remaining three orders was randomized within participants. Eighteen participants received each symptom order first.

[5]Thirty-five participants (48%) exceeded this 60% criterion.

Hypotheses and predictions

A recency effect was predicted on the grounds that more recent cues would be more active in working memory and contribute to the hypothesis generation process to a greater degree than less recent cues. Given that the activation of the diagnostic symptom (S1) in working memory at the time of generation was predicted to increase in correspondence with its serial position, increases in the generation of Metalytis were predicted to be observed with greater recency of S1. As suggested by **Figure 2**, the context-activation model, under parameters based on previous work in list recall paradigms (Davelaar et al., 2005) predicts this generally recency effect as later items are more often more active in memory at the end of list presentation. Correspondingly, decreases in the generation of the alternatives to Metalytis were expected with increases in the serial position of S1. This prediction stems directly from the buffer activation dynamics of the context-activation model.

RESULTS

The main DV for the analyses was the discrete generation vs. non-generation of Metalytis as the most likely disease (i.e., first disease generated). All participants were included in the analyses regardless of performance in the diagnosis test phase and there were no differences in results based on learning. Carry-over effects were evident as demonstrated by a significant interaction between order condition and trial, $\chi^2(3) = 12.68$, $p < 0.016$[6]. In light of this, only the data from the first trial for each participant was subjected to further analysis as it was assumed that this was the only uncontaminated trial for each subject. Nominal logistic regression was used to examine the effect of data serial position on the generation of Metalytis (the disease with the greatest posterior probability given the data). A logistic regression contrast test demonstrated a trend for the generation of Metalytis as it was more often generated as the most likely hypothesis with increases in the serial position of the diagnostic data, $\chi^2(1) = 4.32$, $p < 0.05$. The number of hypotheses generated between order conditions did not differ, $F(3,68) = 0.567$, $p = 0.64$, $\eta_p^2 = 0.02$, ranging from an average of 1.67–1.89 hypotheses. There were no differences in the probability judgments of Metalytis as a function of data order when it was generated as the most likely hypothesis (with group means ranging from 56.00 to 67.13), $F(3,33) = 0.66$, $p = 0.58$, $\eta_p^2 = 0.06$.

SIMULATING EXPERIMENT 1

To simulate Experiment 1, the model's episodic memory was endowed with the Disease-Symptom contingencies described in **Table 1**. On each trial, each symptom was presented to the buffer for 1500 iterations (mapping onto the presentation duration of 1500 ms) and the order of the symptoms was manipulated to match the symptom orders used in the experiment. 1000 iterations of the entire simulation were run for each condition[7]. The

[6]This carry-over effect was not entirely surprising as the same symptom states were presented for every patient and our manipulation of serial order was likely transparent on later trials.

[7]The parameters used for this simulation were the following. Original HyGene parameters: $L = 0.85$, Ac = 0.1, Phi = 4, KMAX = 8. Context-activation model parameters: Alpha = 2.0, Beta = 0.2, Lambda = 0.98, Delta = 1. Note, these parameters were based on values utilized in previous work and were not chosen based on fitting the model to the current data.

primary model output of interest was the first hypothesis generated on each trial. As is demonstrated in **Figure 6**, the model is able to capture the qualitative trend in the empirical data quite well. Although the rate of generation is slightly less for the model, the model clearly captures the recency trend as observed in the empirical data. Increased generation of the most likely hypothesis corresponded to the recency of the diagnostic datum. This effect is directly attributable to the buffer activation weights being applied to the generation process. Although **Figure 10** will become more pertinent later, the left hand side of this figure demonstrates the recency gradient in the data activation weights produced by the model under these parameter settings. Inspection of the average weights for the first two data acquired show them to be below the working memory threshold of 0.2. Therefore, on a large proportion of trials the model relied on only the third and fourth piece of data (or just the last piece). This explains why the model performs around chance under the first two data orders and only deviates under orders three and four. Additionally, it should be noted that the model could provide a suitable quantitative fit to the empirical data by incorporating an assumption concerning the rate of guessing in the task or potentially by manipulating the working memory threshold. Although the aim of the current paper is to capture the qualitative effects evidenced in the data, future work may seek more precise quantitative fits.

DISCUSSION

The primary prediction of the experiment was confirmed. The generation of the most likely hypothesis increased in correspondence with increasing recency of the diagnostic data (i.e., symptom). This finding clearly demonstrates that not all available data contribute equally to the hypothesis generation process (i.e., some data are weighted more heavily than others) and that the serial position of a datum can be an important factor governing the weight allocated to it in the generation process. Furthermore, these results are consistent with the notion that the data weightings utilized in the generation process are governed by the amount of working memory activation possessed by each datum.

There are, however, two alternative explanations for the present finding to consider that do not necessarily implicate unequal weightings of data in working memory as governing generation.

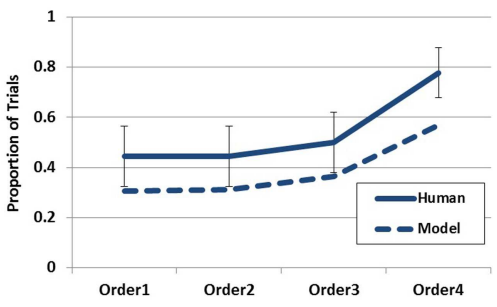


FIGURE 6 | Empirical data (solid line) and model data (dashed line) for Experiment 1 plotting the probability of reporting D1 (Metalytis) as most likely across order conditions. Error bars represent standard errors of the mean.

First, it could be the case that all data resident in working memory at the time of generation were equally weighted, but that the likelihood of S1 dropping out of working memory increased with its distance in time from the generation prompt. Such a discrete utilization (i.e., all that matters is that data are in or out of working memory regardless of the activation associated with individual data) would likely result in a more gradual recency effect than seen in the data. Future investigations measuring working memory capacity could provide illuminating tests of this account. If generation is sensitive to only the presence or absence of data in working memory (as opposed to graded activations of the data in working memory) it could be expected that participants with higher capacity would be less biased by serial order (as shown in Lange et al., 2012) or would demonstrate the bias at a different serial position relative to those with lower capacity.

A second alternative explanation could be that the participants engaged in spontaneous rounds of generation following each piece of data as it was presented. Because the hypothesis generation performance was only assessed after the final piece of data in the present experiment, such "step-by-step" generation would result in stronger generation of Metalytis as the diagnostic data is presented closer to the end of the list. For instance, if spontaneous generation occurs as each piece of data is being presented, then when the diagnostic datum is presented first, there remains three more rounds of generation (based on non-diagnostic data in this case) that could obscure the generation of the initial round. As the diagnostic data moves closer to the end of the data stream the likelihood that that particular round of generation will be obscured by forthcoming rounds diminishes. It is likely that the present data represents a mixture of participants that engaged in such spontaneous generation and those that did not engage in generation until prompted. This is likely the reason for the quantitative discrepancy between the model and empirical data. Future investigations could attempt to determine the likelihood that a participant will engage in such spontaneous generation and the conditions making it more or less likely.

The probability judgments observed in the present experiments did not differ across order conditions. Because the probability judgments were only elicited for the highest ranked hypothesis, the conditions under which the probability judgments were collected were highly constrained. It should be noted that the focus of the present experiment was to address generation behavior and the collection of the judgment data was ancillary. An independent experiment manipulating serial order in the manner done here and designed explicitly for the examination of judgment behavior would be useful for examining the influence of specific data serial positions on probability judgments. This would be interesting as HyGene predicts the judged probability of a hypothesis to be directly influenced by the relative support for the hypotheses currently in working memory. In so far as serial order influences the hypotheses generated into working memory, effects of serial position on probability judgment are likely to be observed as well.

The goal of Experiment 1 was to determine how relative data serial position affects the contribution of individual data to hypothesis generation processes. It was predicted that data presented later in the sequence would be more active in working memory and would thereby contribute more to the generation

process based on the dynamics of the context-activation buffer. Such an account predicts a recency profile for the generation of hypotheses from LTM. This effect was obtained and is well-captured by our model in which such differences in the working memory activation possessed by individual data govern the generation process. Despite these positive results, however, the specific processes underlying this data are not uniquely discernible in the present experiment as the aforementioned alternative explanations likely predict similar results. Converging evidence for the notion that data activation plays a governing role in the generation process should be sought.

EXPERIMENT 2: DATA MAINTENANCE AND DATA CONSISTENCY

When acquiring information from the world that we may use as cues for the generation of hypotheses we acquire these cues in variously sized sets. In some cases we might receive several pieces of environmental data over a brief period, such as when a patient rattles off a list of symptoms to a physician. At other times, however, we receive cues in isolation across time and generate hypotheses based on the first cue and update this set of hypotheses as further data are acquired, such as when an underlying cause of car failure reveals itself over a few weeks. Such circumstances are more complicated as additional processes come into play as further data are received and previously generated hypotheses are evaluated in light of the new data. Hogarth and Einhorn (1992) refer to this task characteristic as the response mode.

In the context of understanding dynamic hypothesis generation this distinction is of interest as it contrasts hypothesis generation following the acquisition of a set of data with a situation in which hypotheses are generated (and updated or discarded) while further data is acquired and additional hypotheses generated. An experiment manipulating this response mode variable in a hypothesis generation task was conducted by Sprenger and Dougherty, 2012, Experiment 3) in which people hypothesized about which psychology courses were being described by various keywords. The two response modes are step-by-step (SbS), in which a response is elicited following each piece of incoming data, and end-of-sequence (EoS), in which a response is made only after all the data has been acquired as a grouped set. Following the last piece of data, the SbS conditions exhibited clear recency effects whereas EoS conditions, on the other hand, did not demonstrate reliable order effects. A careful reader may notice a discrepancy between the lack of order effects in their EoS condition and the recency effect in the present Experiment 1 (which essentially represents an EoS mode condition). In the Sprenger and Dougherty experiment, the participants received nine cues from which to generate hypotheses as opposed to the four cues in our Experiment 1. As the amount of data in their experiment exceeded working memory capacity (more severely) it is likely that the cue usage strategies utilized by the participants differed between the two experiments. Indeed, it is important to gain a deeper understanding of such cue usage strategies in order to develop a better understanding of dynamic hypothesis generation.

The present experiment compared response modes to examine differences between data maintenance prior to generation (EoS mode) and generation that does not encourage the maintenance of multiple pieces of data (SbS mode). Considered in another light, SbS responding can be thought of as encouraging an anchoring and adjustment process where the set of hypotheses generated in response to the first piece of data supply the set of beliefs in which forthcoming data may be interpreted. The EoS condition, on the other hand, does not engender such belief anchoring as generation is not prompted until all data have been observed. As such, the SbS conditions provide investigation of a potential propensity to discard previously generated hypotheses and/or generate new hypotheses in the face of inconsistent data.

METHOD

Participants

One hundred fifty-seven participants from the University of Oklahoma participated in this experiment for course credit.

Design and procedure

As previously mentioned, the first independent variable was the timing of the generation and judgment promptings provided to the participant as dictated by the response mode condition. This factor was manipulated within-subject. The second independent variable, manipulated between-subjects, was the consistency of the second symptom (S2) with the hypotheses likely to be entertained by the participant following the first symptom. This consistency or inconsistency was manipulated within the ecologies learned by the participants as displayed in **Table 3**. In addition, this table demonstrates the temporal order in which the symptoms were presented in the elicitation phase of this experiment (i.e., S1 → S2 → S3 → S4). Note that only positive symptom (i.e., symptom present) states were presented in the elicitation phase. The only difference between the ecologies was the conditional probability of S2 being positive under D1. This probability was 0.9 in the "consistent ecology" and 0.1 in the "inconsistent ecology." Given that S1 should prompt the generation of D1 and D2, this manipulation of the ecology can be realized to govern the consistency of S2 with the hypothesis(es) currently under consideration following S1. This can be seen in **Table 4** displaying the Bayesian posterior probabilities for each disease following each symptom. Seventy-nine participants were in the consistent ecology condition and 78 participants were in the inconsistent ecology condition. Response mode was counter-balanced within ecology condition.

The procedure was much like that of Experiment 1: exemplar training to learn the probability distributions, a test to verify learning (for which a $5.00 gift card could be earned for performance greater than 60%)[8], and a distractor task prior to elicitation. The experiment was again cast in terms of medical diagnosis where D1, D2, and D3 represented fictitious disease states and S1–S4 represented various test results (i.e., symptoms).

There were slight differences in each phase of the procedure however. The exemplars presented in the exemplar training phase of were simplified and consisted of the disease name and a single test result (as opposed to all four). This change was made in an effort to enhance learning. Exemplars were blocked by disease

[8]Eighty-eight participants (56%) exceeded this 60% criterion.

Table 3 | Disease × Symptom ecologies of Experiment 2.

	S1	S2	S3	S4
CONSISTENT				
D1: Metalytis	0.9	**0.9**	0.5	0.5
D2: Zymosis	0.7	0.1	0.4	0.4
D3: Gwaronia	0.2	0.8	0.8	0.8
INCONSISTENT				
D1: Metalytis	0.9	**0.1**	0.5	0.5
D2: Zymosis	0.7	0.1	0.4	0.4
D3: Gwaronia	0.2	0.8	0.8	0.8

Values represent the probability of the symptom being positive (i.e., present) given the disease state. Only the value of S2|D1 differed between ecologies as indicated by underline and bold.

Table 4 | Bayesian posterior probabilities as further symptoms are acquired within each ecology of Experiment 2.

		Posterior probabilities across elicitations			
		Post S1	Post S2	Post S3	Post S4
Consistent ecology	D1	0.50	0.78	0.72	0.64
	D2	0.39	0.07	0.05	0.04
	D3	0.11	0.15	0.23	0.32
Inconsistent ecology	D1	0.50	0.28	0.22	0.17
	D2	0.39	0.22	0.14	0.08
	D3	0.11	0.50	0.64	0.75

such that a disease was selected at random without replacement. For each disease the participant would be presented with 40 exemplars selected at random without replacement. Therefore over the course of these 40 exemplars the entire (and exact) distribution of symptoms would be presented for that disease. This was then done for the remaining two diseases and the entire process was repeated two more times. Therefore the participant observed 120 exemplars per disease (inducing equal base rates for each disease) and observed the entire distribution three times. Each exemplar was again pseudo-self-paced and displayed on the screen for 1500 ms per exemplar prior to the participant being able to proceed to the next exemplar by pressing the first letter of the disease. Patient cases in the diagnosis test phase presented with only individual symptoms as well. Each of the eight possible symptom states were individually presented to the participants and they were asked to report the most likely disease given that particular symptom. Diseases with a posterior probability greater than or equal to 0.39 were tallied as correct responses.

In the elicitation phase, the prompts for hypothesis generation were the same as those used in Experiment 1, but the probability judgment prompt differed slightly. The judgment prompt used in the present experiment was as follows: "How likely is it that the patient has [INSERT HIGHEST RANKED DISEASE]? (Keep in mind that an answer of 0 means that there is NO CHANCE that the patient has [INSERT HIGHEST RANKED DISEASE] and that 100 means that you are ABSOLUTELY CERTAIN that the patient has [INSERT HIGHEST RANKED DISEASE].) Type in your answer from 1 to 100 and press Enter to continue." Probability judgments were taken following each generation sequence in the SbS condition (i.e., there were four probability judgments taken, one for the disease ranked highest on each round of generation).

Hypotheses and predictions

The general prediction for the end-of-sequence response mode was that recency would be demonstrated in both ecologies as the more recent symptoms should contribute more strongly to the generation process as seen in Experiment 1. Therefore, greater generation of D3 relative to the alternatives was expected in both ecologies. The focal predictions for the SbS conditions concerned the generation behavior following S2. It was predicted that participants in the consistent ecology would generate D1 to a greater extent than those in the inconsistent ecology who were expected to purge D1 from their hypothesis set in response to its inconsistency with S2. It was additionally predicted that those in the inconsistent ecology would generate D3 to a greater extent at this point than those in the consistent ecology as they would utilize S2 to repopulate working memory with a viable hypothesis.

RESULTS

As no interactions with trial order were detected, both trials from each subject were used in the present analyses and no differences in results were found with differences in learning. The main dependent variable analyzed for this experiment was the hypothesis generated as most likely on each round of elicitation. All participants were included in the analyses regardless of performance in the diagnosis test phase. In order to test if a recency effect obtained following the last symptom (S4), comparisons between the rates of generation of each disease were carried out within each of the four ecology-by-response mode conditions. Within the step-by-step conditions the three diseases were generated at different rates in the consistent ecology according to Cochran's Q Test, $\chi^2(2) = 9.14$, $p < 0.05$, but not in the inconsistent ecology $\chi^2(2) = 1$, $p = 0.61$. In the end-of-sequence conditions, significant differences in generation rates were revealed in both the consistent ecology, $\chi^2(2) = 17.04$, $p < 0.001$, and the inconsistent ecology, $\chi^2(2) = 7.69$, $p < 0.05$.

As D2 was very unlikely in both ecologies the comparison of interest in all cases is between D1 and D3. This pairwise comparison was carried out within each of the ecology-by-response mode conditions and reached significance only in the EoS mode in the consistent ecology, $\chi^2(1) = 6.79$, $p < 0.01$, as D1 was generated to a greater degree than D3 according to Cochran's *Q* Test. These results, displayed in **Figure 7**, demonstrate the absence of a recency effect in the present experiment. This difference between the EoS and SbS ecology is additionally observed by comparing rates of D1 generation across the entire design demonstrating a main effect of ecology, $\chi^2(1) = 8.87$, $p < 0.01$, but no effect of mode, $\chi^2(1) = 0.987$, $p = 0.32$, and no interaction, $\chi^2(1) = 0.554$, $p = 0.457$.

To test the influence of the inconsistent cue on the maintenance of D1 (the most likely disease in both ecologies following S1) in the SbS conditions, elicitation round (post S1 and post S2) was entered as an independent variable with ecology and tested in a 2 × 2 logistic regression. As plotted in **Figure 8**, this revealed a main effect

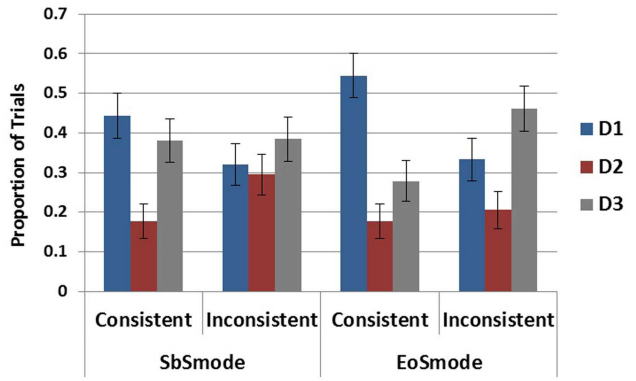


FIGURE 7 | Proportion of generation for each disease by response mode and ecology conditions. Error bars represent standard errors of the mean.

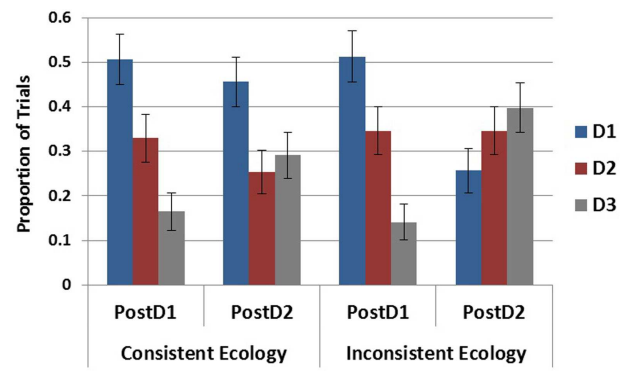


FIGURE 8 | Proportion of generation for each disease within the SbS condition following S1 and S2. Error bars represent standard errors of the mean.

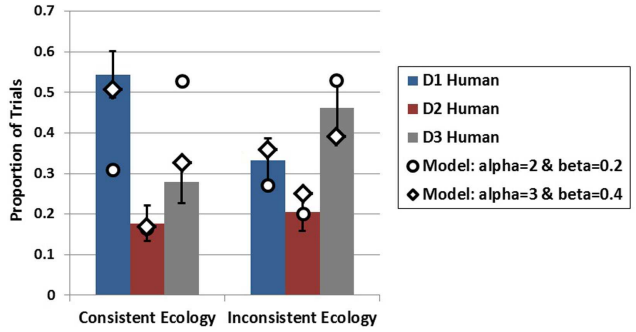


FIGURE 9 | Empirical data (bars) from Experiment 2 for the EoS conditions in both ecologies plotted with model data (diamonds and circles) at two levels of alpha and beta. Error bars represent standard errors of the mean.

of elicitation round, $\chi^2(1) = 10.51$, $p < 0.01$, an effect of ecology, $\chi^2(1) = 6.65$, $p < 0.05$, and a marginal interaction, $\chi^2(1) = 3.785$, $p = 0.052$. When broken down by ecology it is evident that the effect of round and the marginal interaction were due to the decreased generation of D1 following S2 in the inconsistent ecology, $\chi^2(1) = 10.51$, $p < 0.01$, as there was no difference between rounds in the consistent ecology, $\chi^2(1) = 0.41$, $p = 0.524$.

This same analysis was done with D3 to examine potential differences in its rate of generation over these two rounds of generation. This test revealed a main effect of elicitation round, $\chi^2(1) = 12.135$, $p < 0.001$, but no effect of ecology, $\chi^2(1) = 1.953$, $p = 0.162$, and no interaction, $\chi^2(1) = 1.375$, $p = 0.241$.

SIMULATING EXPERIMENT 2

To model the EoS conditions, the model was presented all four symptoms in sequence and run in conditions in which the model was endowed with either the consistent or inconsistent ecology. This simulation was run for 1000 iterations in each condition. As is intuitive from the computational results of Experiment 1, when the model is run with the same parameters utilized in the previous simulation it predicts greater generation for D3 in both ecologies (i.e., recency) which was not observed in the present experiment. However, the model is able to capture the data of the EoS mode quite well by increasing the amount of recurrent activation that each piece of data recycles onto itself (alpha parameter) and the amount of lateral inhibition applied to each piece of data (beta parameter) as it is acquired prior to generation. These results appear alongside the empirical results in **Figure 9**. Although the model is able the capture the qualitative pattern in the data in the inconsistent ecology reasonably well with either set of parameters, the model produces divergent results under the two alpha and beta levels in the consistent ecology. Only when recurrency and inhibition are increased does the model capture the data from both ecologies.

Examination of how the data activations are influenced by the increased alpha and beta levels reveals the underlying cause for this difference in generation. As displayed in **Figure 10**, there is a steep recency gradient for the data activations under alpha = 2 and beta = 0.2 (parameters from Experiment 1), but there is a markedly different pattern of activations under alpha = 3 and beta = 0.4[9]. Most notably, these higher alpha and beta levels cause the earlier pieces of data to reach high levels of activation which then suppress the activation levels of later data. This is due to the competitive dynamics of the buffer which restrict rise of activation for later items under high alpha and beta values resulting in a primacy gradient in the activation values as opposed to the recency gradient observed under the lower values.

To capture the SbS conditions for generation following S1 and generation following S2, the model was presented with different amounts of data on different trials. Specifically, the model was presented with S1 only, capturing the situation in which only the first piece of data had been received, or the model was presented with S1 and S2 successively in order to capture the SbS condition following the second piece of data. This was done for both ecologies in order to assess the effects of data inconsistency on the

[9]These parameter values were based on a grid search to examine the neighborhood of values capturing the qualitative patterns in the data and not based on a quantitative fit to the empirical data.

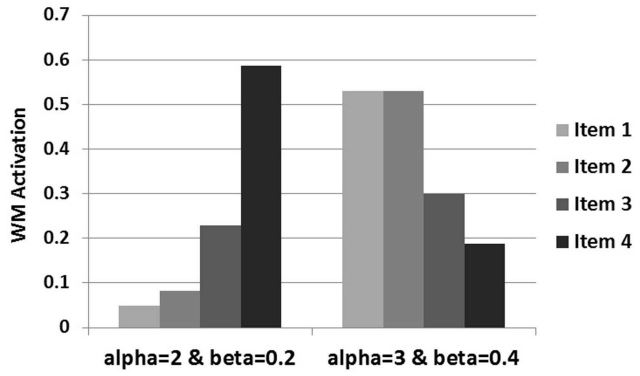


FIGURE 10 | Individual data activations under both levels of alpha and beta.

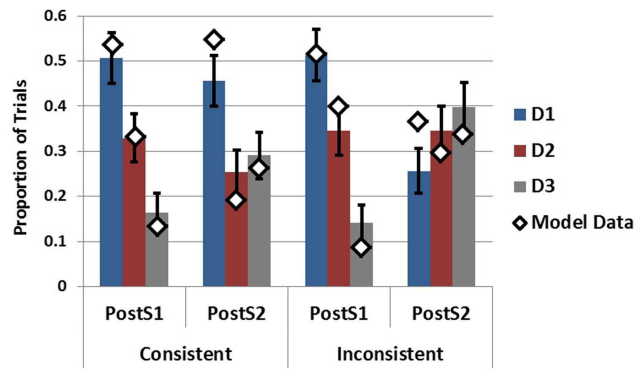


FIGURE 11 | Empirical data (bars) from Experiment 2 in the SbS conditions following S1 and S2 plotted with model data (diamonds). Error bars represent standard errors of the mean.

model's generation behavior[10]. As can be seen in **Figure 11** the model is able to capture the empirical data quite well following S1 while providing a decent, although imperfect, account of the post S2 data as well[11]. Focally, the model as implemented captures the influences of S2 on the hypothesis sets generated in response to S1. Following S2 in the inconsistent ecology D1 decreases substantially capturing its purging from working memory. Additionally, the increases in the generation of D3 are present in both ecologies.

DISCUSSION

The present experiment has provided a window into two distinct processing dynamics. The first dynamic under investigation was how generation differs when based on the acquisition of a set of data (EoS condition) vs. when each piece of data is acquired in isolation (SbS condition). The generation behavior between these conditions was somewhat similar overall, as neither D1 nor D3 dominated generation in three of the four conditions. The EoS consistent ecology condition, however, was clearly dominated by D1. This result obtained in contrast to the prediction of recency in the EoS conditions, which would have been evidenced by higher rankings for D3 (for both ecologies).

The divergence between the recency effect in Experiment 1 and the absence of recency effect in the EoS conditions of Experiment 2 is surprising. In order for the model to account for the amelioration of the recency effect an adjustment was made to the alpha and beta parameters governing how much activation each piece of data is able to recycle onto itself and the level of competition thereby eliminating the recency gradient in the activations. Moreover, the last piece of data did not contribute as often or as strongly to the cue to LTM under these settings. Therefore, rather than a recency effect, the model suggests a primacy effect whereby the earlier cues contributed more to generation than the later cues. As we have not manipulated serial order in the present experiment, it is difficult to assert a primacy effect based on the empirical data alone. The model's account of the current data, however, certainly suggests that a primacy gradient is needed to capture the results. Additionally, a recent experiment in a similar paradigm utilizing an EoS response mode demonstrated a primacy effect in a diagnostic reasoning task (Rebitschek et al., 2012) suggesting that primacy may be somewhat prevalent under EoS data acquisition situations.

As for why the earlier cues may have enjoyed greater activation in the present experiment relative to Experiment 1 we need to consider the main difference between these paradigms. The largest difference was that in the present experiment each piece of data present in the ecology carried a good amount of informational value whereas in Experiment 1 80% of the data in the ecology was entirely non-diagnostic. It is possible that this information rich vs. information scarce ecological difference unintentionally led to a change in how the participants allocated their attention over the course of the data streams between the two experiments. As all of the data in Experiment 2 was somewhat useful, the participants may have used this as a cue to utilize as much of the information as possible thereby rehearsing/reactivating the data as much as possible prior to generation. In contrast, being in the information scarce ecology of Experiment 1 would not have incentivized such maximization of the data activations for most of the data. Future experiments could address how the complexity of the ecology might influence dynamic attentional allocation during data acquisition.

The second dynamic explored was how inconsistent data influences the hypotheses currently under consideration. In the step-by-step conditions it was observed that a previously generated hypothesis was purged from working memory in response to the inconsistency of a newly received cue. This can be viewed as consistent with an extension of the consistency checking mechanism employed in the original HyGene framework. The present data suggests that hypotheses currently under consideration are checked against newly acquired data and are purged in accordance with their degree of (in)consistency. This is different from, although entirely compatible with, the operation of the original consistency checking mechanism operating over a single round of hypothesis generation. The consistency checking operation within the original version of HyGene checks each hypothesis retrieved into working memory for its consistency with the data used as a cue

[10]This is, of course, a simplification of the participant's task in the SbS condition. This is addressed in the general discussion.
[11]This simulation was run with alpha = 3 and beta = 0.4.

to its retrieval as the SOCs is populated. The consistency checking mechanism exposed in the present experiment, however, suggests that people check the consistency of *newly* acquired data against hypotheses generated from *previous* rounds of generation as well. If the previously generated hypotheses fall below some threshold of agreement with the newly acquired data they are purged from working memory. Recent work by Mehlhorn et al. (2011) also investigated the influence of consistent and inconsistent cues on the memory activation of hypotheses. They utilized a clever adaptation of the lexical decision task to assess the automatic memory activation of hypotheses as data were presented and found memory activation sensitivity to the consistency of the data. As the present experiment utilized overt report, these findings complement one another quite well as automatic memory activation can be understood as a precursor to the generation of hypotheses into working memory. The present experiment additionally revealed that S2 was used to re-cue LTM as evidenced by increased generation of D3 following S2. In contrast to the prediction that this would occur only in the inconsistent ecology, this recuing was observed in both ecologies. Lastly, although the model as currently implemented represents a simplification of the participant's task in the SbS conditions, it was able to capture these effects.

GENERAL DISCUSSION

This paper presented a model of dynamic data acquisition and hypothesis generation which was then used to account for data from two experiments investigating three consequences of hypothesis generation being extended over time. Experiment 1 varied the serial position of a diagnostic datum and demonstrated a recency effect whereby the hypothesis implied by this datum was generated more often when the datum appeared later in the data stream. Experiment 2 examined how generation might differ when it is based on isolated data acquired one at a time (step-by-step response mode) vs. when generation is based upon the acquisition of the entire set of data (end-of-sequence response mode). Secondly, the influence of an inconsistent cue (conflicting with hypotheses suggested by the first datum) was investigated by manipulating a single contingency of the data-hypothesis ecology in which the participants were trained. It was found that the different response modes did not influence hypothesis generation a great deal as the two most likely hypotheses were generated at roughly the same rates in most cases. The difference that was observed however was that the most likely hypothesis was favored in the EoS condition within the consistent ecology. This occurred in contrast to the prediction of recency for both EoS conditions, thereby suggesting that the participants weighted the data more equally than in Experiment 1 or perhaps may have weighted the earlier cues slightly more heavily. Data from the SbS conditions following the acquisition of the inconsistent cue revealed that this cue caused participants to purge a previously generated hypothesis from working memory that was incompatible with the newly acquired data. Moreover, this newly acquired data was utilized to re-cue LTM. Interestingly, this re-cueing was demonstrated in both ecologies and was therefore not contingent on the purging of hypotheses from working memory.

Given that the EoS conditions of Experiment 2 were procedurally very similar to the procedure used in Experiment 1 it becomes important to reconcile their contrasting results. As discussed above, the main factor distinguishing these conditions was the statistical ecology defining their respective data-hypothesis contingencies. The ecology of the first experiment contained mostly non-diagnostic data whereas each datum in the ecology utilized in Experiment 2 carried information as to the relative likelihood of each hypothesis. It is possible that this difference of relative information scarcity and information richness influenced the processing of the data streams between the two experiments. In order to capture the data from Experiment 2 with our model, the level of recurrent activation recycled by each piece of data was adjusted upwards and lateral inhibition increased thereby giving the early items a large processing advantage over the later pieces of data. Although *post hoc*, this suggests the presence of a primacy bias. It is then perhaps of additional interest to note that the EoS results resemble the SbS results following D2 and this is particularly so within the consistent ecology. This could be taken to suggest that those in the EoS condition were utilizing the initial cues more greatly than the later cues. Fisher (1987) suggested that people tend to use a subset of the pool of provided data and estimated that people generally use two cues when three are available and three cues when four are available. Interestingly the model forwarded in the present paper provides support for this estimate as it used three of the four available cues in accounting for the EoS data in Experiment 2. While the utilization of three as opposed to four data could be understood as resulting from working memory constraints, the determinants of why people would fail to utilize three pieces of data when only three data are available is less clear. Future investigation of the conditions under which people underutilize available data in three and four data-hypothesis generation problems could be illuminating for the working memory dynamics of these tasks.

It is also important to compare the primacy effect in the EoS conditions with the results of Sprenger and Dougherty (2012) in which the SbS conditions revealed recency (Experiments 1 and 3) and no order effects were revealed in the EoS conditions (only implemented in Experiment 3). As for why the SbS results of the present experiment do not demonstrate recency as in their Experiments 1 and 3 is unclear. The ecologies used in these experiments were quite different, however, and it could be the case that the ecology implemented in their experiment was better able to capture this effect. Moreover, they explicitly manipulated data serial order and it was through this manipulation that the recency effect was observed. As serial order was not manipulated in the present experiment we did not have the opportunity to observe recency in the same fashion and instead relied on relative rates of generation given one data ordering. Perhaps the manipulation of serial order within the present ecology would uncover recency as well.

In comparing the present experiment to the procedure of Sprenger and Dougherty's Experiment 3 a clearer reason for diverging results is available. In their experiment, the participants were presented with a greater pool of data from which to generate hypotheses, nine pieces in total. Participants in the present experiment, on the other hand, were only provided with four cues. It is quite possible that people's strategies for cue usage would differ between these conditions. Whereas the present experiment provided enough data to fill working memory to capacity (or barely

breach it), Sprenger and Dougherty's experiment provided an abundance of data thereby providing insight into a situation in which the data could not be held in working memory at once. It is possible that the larger pool of data engendered a larger pool of strategies to be employed than in the present study. Understanding the strategies that people employ and the retrieval plans developed under such conditions (Raaijmakers and Shiffrin, 1981; Gillund and Shiffrin, 1984; Fisher, 1987) as well as how these processes contrast with situations in which fewer cues are available is a crucial aspect of dynamic memory retrieval in need of better understanding.

The model presented in the present work represents a fusion of the HyGene model (Thomas et al., 2008) with the activation dynamics of the context-activation model of memory (Davelaar et al., 2005). As the context-activation model provides insight into the working memory dynamics underlying list memory tasks, it provides a suitable guidepost for understanding some of the likely working memory dynamics supporting data acquisition and hypothesis generation over time. The present model acquires data over time whose activations systematically ebb and flow in concert with the competitive buffer dynamics borrowed from the context-activation model. The resulting activation levels possessed by each piece of data are then used as weights in the retrieval of hypotheses from LTM. In addition to providing an account of the data from the present experiments this model has demonstrated further usefulness by suggesting potentially fruitful areas of future investigation.

The modeling presented here represents the first step of a work in progress. As we are working toward a fully dynamical model of data acquisition, hypothesis generation, maintenance, and use in decision making tasks, additional facets clearly still await inclusion. Within the current implementation of the model it is only the environmental data that are subject to the working memory activation dynamics of the working memory buffer. In future work, hypotheses generated into working memory (HyGene's SOCs) will additionally be sensitive to these dynamics. This will provide us with the means of fully capturing hypothesis maintenance dynamics (e.g., step-by-step generation) that the present model ignores. Moreover, by honoring such dynamic maintenance processes we may be able to address considerations of what information people utilize at different portions of a hypothesis generation task. For instance, when data is acquired over long lags (e.g., minutes), it is unclear what information people use to populate working memory with hypotheses at different points in the task. If someone is reminded of the diagnostic problem they are trying to solve, do they recall the hypotheses directly (e.g., via contextual retrieval) or do they sometimes recall previous data to be combined with new data and re-generate the current set of hypotheses? Presumably both strategies are prevalent, but the conditions under which they are more or less likely to manifest is unclear. It is hoped that this more fully specified model may provide insight into situations favoring one over the other.

As pointed out by Sprenger and Dougherty (2012) a fuller understanding of hypothesis generation dynamics will entail learning about how working memory resources are dynamically allocated between data and hypotheses over time. One-way that this could be achieved in the forthcoming model would be to have two sets of information available for use at any given time, one of which would be the set of relevant data (RED) and the other would be the SOC hypotheses. The competitive dynamics of the buffer could be brought to bear between these sets of items by allowing them to inhibit one another, thereby instantiating competition between the items in these sets for the same limited resource. Setting up the model in this or similar manners would be informative for addressing dynamic working memory tradeoffs that are struck between data and hypotheses over time.

In addition, this more fully elaborated model could inform maintenance dynamics as hypotheses are utilized to render judgments and decisions. The output of the judgment and decision processes could cohabitate the working memory buffer and its maintenance and potential influence on other items' activations could be gauged across time. Lastly, as the model progresses in future work it will be important and informative to examine the model's behavior more broadly. For the present paper we have focused on the first hypothesis generated in each round of generation. The generation behavior of people and the model of course furnishes more than one hypothesis into working memory. Further work with this model has the potential to provide a richer window into hypothesis generation behavior by taking a greater focus on the full hypothesis sets considered over time.

Developing an understanding of the temporal dynamics governing the rise and fall of beliefs over time is a complicated problem in need of further investigation and theoretical development. This paper has presented an initial model of how data acquisition dynamics influence the generation of hypotheses from LTM and two experiments considering three distinct processing dynamics. It was found that the recency of the data, sometimes but not always, biases the generation of hypotheses. Additionally, it was found that previously generated hypotheses are purged from working memory in light of new data with which they are inconsistent. Future work will develop a more fully specified model of dynamic hypothesis generation, maintenance, and use in decision making tasks.

REFERENCES

Anderson, N. H. (1965). Primacy effects in personality impression formation using a generalized order effect paradigm. *J. Pers. Soc. Psychol.* 2, 1–9.

Anderson, N. H. (1973). Serial position curves in impression formation. *J. Exp. Psychol.* 97, 8–12.

Cooper, R. P., Yule, P., and Fox, J. (2003). Cue selection and category learning: a systematic comparison of three theories. *Cogn. Sci. Q.* 3, 143–182.

Davelaar, E. J., Goshen-Gottstein,Y, Ashkenazi, A., Haarmann, H. J., and Usher, M. (2005). The demise of short term memory revisited: empirical and computational investigations of recency effects. *Psychol. Rev.* 112, 3–42.

Dougherty, M. R. P., Gettys, C. F., and Ogden, E. E. (1999). A memory processes model for judgments of likelihood. *Psychol. Rev.* 106, 180–209.

Dougherty, M. R. P., and Hunter, J. E. (2003a). Probability judgment and subadditivity: the role of WMC and constraining retrieval. *Mem. Cognit.* 31, 968–982.

Dougherty, M. R. P., and Hunter, J. E. (2003b). Hypothesis generation, probability judgment, and working memory capacity. *Acta Psychol. (Amst.)* 113, 263–282.

Dougherty, M. R. P., Thomas, R. P., and Lange, N. (2010). Toward an integrative theory of hypothesis generation,

probability judgment, and hypothesis testing. *Psychol. Learn. Motiv.* 52, 299–342.

Fisher, S. D. (1987). Cue selection in hypothesis generation: Reading habits, consistency checking, and diagnostic scanning. *Organ. Behav. Hum. Decis. Process.* 40, 170–192.

Gillund, G., and Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychol. Rev.* 91, 1–67.

Hintzman, D. L. (1986). "Schema Abstraction" in a multiple-trace memory model. *Psychol. Rev.* 93, 411–428.

Hintzman, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychol. Rev.* 95, 528–551.

Hogarth, R. M., and Einhorn, H. J. (1992). Order effects in belief updating: the belief-adjustment model. *Cogn. Psychol.* 24, 1–55.

Lange, N. D., Thomas, R. P., and Davelaar, E. J. (2012). "Data acquisition dynamics and hypothesis generation," in *Proceedings of the 11th International Conference on Cognitive Modelling*, eds N. Rußwinkel, U. Drewitz, J. Dzaack H. van Rijn, and F. Ritter (Berlin: Universitaetsverlag der TU), 31–36.

McKenzie, C. R. M. (1998). Taking into account the strength of an alternative hypothesis. *J. Exp. Psychol. Learn. Mem. Cogn.* 24, 771–792.

Mehlhorn, K., Taatgen, N. A., Lebiere, C., and Krems, J. F. (2011). Memory activation and the availability of explanations in sequential diagnostic reasoning. *J. Exp. Psychol. Learn. Mem. Cogn.* 37, 1391–1411.

Murdock, B. B. (1962). The serial position effect of free recall. *J. Exp. Psychol.* 64, 482–488.

Nelson, J. D., McKenzie, C. R. M., Cottrell, G. W., and Sejnowski, T. J. (2010). Experience matters: information acquisition optimizes probability gain. *Psychol. Sci.* 21, 960–969.

Page, M. P. A., and Norris, D. (1998). The primacy model: a new model of immediate serial recall. *Psychol. Rev.* 105, 761–781.

Raaijmakers, J. G. W., and Shiffrin, R. M. (1981). Search of associative memory. *Psychol. Rev.* 88, 93–134.

Rebitschek, F., Scholz, A., Bocklisch, F., Krems, J. F., and Jahn, G. (2012). "Order effects in diagnostic reasoning with four candidate hypotheses," in *Proceedings of the 34th Annual Conference of the Cognitive Science Society*, eds N. Miyake, D. Peebles, and R. P. Cooper (Austin, TX: Cognitive Science Society) (in press).

Sprenger, A., and Dougherty, M. P. (2012). Generating and evaluating options for decision making: the impact of sequentially presented evidence. *J. Exp. Psychol. Learn. Mem. Cogn.* 38, 550–575.

Sprenger, A., and Dougherty, M. R. P. (2006). Differences between probability and frequency judgments: the role of individual differences in working memory capacity. *Organ. Behav. Hum. Decis. Process.* 99, 202–211.

Sprenger, A. M., Dougherty, M. R., Atkins, S. M., Franco-Watkins, A. M., Thomas, R. P., Lange, N. D., and Abbs, B. (2011). Implications of cognitive load for hypothesis generation and probability judgment. *Front. Psychol.* 2:129. doi:10.3389/fpsyg.2011.00129

Thomas, R. P, Dougherty, M. R., Sprenger, A. M., and Harbison, J. I. (2008). Diagnostic hypothesis generation and human judgment. *Psychol. Rev.* 115, 155–185.

Usher, M., Davelaar, E. J., Haarmann, H., and Goshen-Gottstein, Y. (2008). Short term memory after all: comment on Sederberg, Howard, and Kahana (2008). *Psychol. Rev.* 115, 1108–1118.

Weiss, D. J., and Anderson, N. H. (1969). Subjective averaging of length with serial position. *J. Exp. Psychol.* 82, 52–63.

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: 24 January 2012; accepted: 09 June 2012; published online: 29 June 2012.

*Citation: Lange ND, Thomas RP and Davelaar EJ (2012) Temporal dynamics of hypothesis generation: the influences of data serial order, data consistency, and elicitation timing. Front. Psychology **3**:215. doi: 10.3389/fpsyg.2012.00215*

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

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 01 October 2012
doi: 10.3389/fpsyg.2012.00335

The dynamics of decision making in risky choice: an eye-tracking analysis

Susann Fiedler* and Andreas Glöckner

Max Planck Institute for Research on Collective Goods, Bonn, Germany

Edited by:
Marius Usher, Tel-Aviv University, Israel

Reviewed by:
Joseph G. Johnson, Miami University, USA
Neil Stewart, University of Warwick, UK
Ian Krajbich, University of Zurich, Switzerland

****Correspondence:***
Susann Fiedler, Max Planck Institute for Research on Collective Goods, Kurt-Schumacher-Street 10, D-53113 Bonn, Germany.
e-mail: fiedler@coll.mpg.de

In the last years, research on risky choice has moved beyond analyzing choices only. Models have been suggested that aim to describe the underlying cognitive processes and some studies have tested process predictions of these models. Prominent approaches are evidence accumulation models such as decision field theory (DFT), simple serial heuristic models such as the adaptive toolbox, and connectionist approaches such as the parallel constraint satisfaction (PCS) model. In two studies involving measures of attention and pupil dilation, we investigate hypotheses derived from these models in choices between two gambles with two outcomes each. We show that attention to an outcome of a gamble increases with its probability and its value and that attention shifts toward the subsequently favored gamble after about two thirds of the decision process, indicating a gaze-cascade effect. Information search occurs mostly within-gambles, and the direction of search does not change over the course of decision making. Pupil dilation, which reflects both cognitive effort and arousal, increases during the decision process and increases with mean expected value. Overall, the results support aspects of automatic integration models for risky choice such as DFT and PCS, but in their current specification none of them can account for the full pattern of results.

Keywords: risky choices, decision field theory, heuristics, parallel constraint satisfaction, eye tracking, arousal, gaze-cascade effect

INTRODUCTION

In many every day decisions individuals choose between options with different outcomes, each of which realizes with a certain probability. Conceptually, such risky choices can be reduced to decisions between gambles with monetary outcomes. Research in risky choice has a long tradition, and several models have been developed to account for the wealth of identified choice anomalies and data observed on risky choices in the lab and the field. Many of these models are extensions and refinements of the expected utility (EU) Model (von Neumann and Morgenstern, 1944; Edwards, 1954; Savage, 1954) assuming that some kind of integration of payoff and probability of outcomes drives individuals' choice. Cumulative prospect theory (CPT, Tversky and Kahneman, 1992), and the transfer-of-attention exchange model (Birnbaum, 2008a) are two of the most prominent models of this class. One important limitation of these models, however, is that they predict choices only and remain largely silent concerning the processes underlying choice (Luce and Raiffa, 1957; Luce, 2000).

PROCESS MODELS FOR RISKY CHOICE

The importance of investigating the underlying processes in more depth and of developing process models, however, has been highlighted repeatedly (e.g., Payne et al., 1988; Johnson et al., 2008; Franco-Watkins and Johnson, 2011; Schulte-Mecklenbeck et al., 2011). Two important classes of existing process models are simplifying heuristic models and automatic integration models. Simplifying heuristic models assume that individuals try to avoid integrating all pieces of information and apply short-cuts instead. According to a satisficing heuristics, for example, one could choose the first gamble that meets the criteria for adequacy on all outcomes (Simon, 1955, 1956). Alternatively, one might apply a minimax heuristic or a maximax heuristic, by choosing the gamble which is better then all alternatives concerning its worst outcome, or better in the best outcome, respectively. Persons might also apply a strategy that tests a sequence of such "reasons" as assumed by the priority heuristic (PH; Brandstätter et al., 2006). PH assumes that individuals compare reasons in the order: minimum outcomes (higher is better), probability of the minimum outcome (lower is better), and maximum outcome (higher is better). Participants consider these reasons in turn, make their decision as soon as a reason discriminates between the gambles, and decide in line with this reason.

Automatic integration models, in contrast, assume that individuals rely on powerful cognitive processes that allow integration of much information very quickly and efficiently (Schneider and Shiffrin, 1977; Shiffrin and Schneider, 1977). Examples for automatic integration models in the domain of risky choice are evidence accumulation models such as decision field theory (DFT, Busemeyer and Townsend, 1993; Johnson and Busemeyer, 2005), and the leaky competing accumulator model (Usher and McClelland, 2001, 2004); as well as interactive activation models such as the parallel constraint satisfaction (PCS) model for risky choice (Glöckner and Herbold, 2011; see also Glöckner and Betsch, 2008b; Betsch and Glöckner, 2010). Further automaticity-based evidence accumulation models that focus on decision making under certainty but could eventually also be applied to risky decisions

are multiattributive DFT (e.g., Diederich, 1997, 2003), and the attention drift-diffusion model (e.g., Krajbich et al., 2010, 2012; Milosavljevic et al., 2010; Krajbich and Rangel, 2011).

MODEL COMPARISONS AND TESTS OF PROCESS MODELS

A recent comprehensive model comparison indicates that none of the available simple heuristics (not even when assuming their adaptive usage) can predict risky choice in decisions between two-outcome gambles nearly as well as CPT (Glöckner and Pachur, 2012)[1]. Despite of some initial support for heuristics and particularly for the PH (Brandstätter et al., 2006), most later studies have challenged predictions derived from PH and similar semi-lexicographic heuristics with regard to choices (Birnbaum, 2008b; Birnbaum and LaCroix, 2008; Hilbig, 2008), decision times (Glöckner and Betsch, 2008a; Ayal and Hochman, 2009; but see Brandstätter et al., 2006), and information search (Glöckner and Betsch, 2008a; Johnson et al., 2008).

One study also used eye-tracking to directly compare predictions of PH and automatic integration models with regard to the process measures (Glöckner and Herbold, 2011). The results concerning information search, response time, and choice speak against the usage of PH as well as similar heuristics. The findings, however, also rule out that individuals use simple serial implementation of EU models assuming a stepwise calculation of weighted sums, which is one possible process implementation of EU suggested in previous research (Payne et al., 1988). Instead, results support automatic integration models such as DFT and PCS in that mainly short fixations were found, which indicate lower-level automatic processing instead of deliberate calculation (Horstmann et al., 2009). Furthermore, the decision time and number of fixations increased for decisions in which differences between gambles were small. In line with PCS predictions, there was an increase in the number of fixations to the favored gamble and particularly to the most attractive outcome of this gamble (defined by a high product of outcome and probability). Furthermore, and in line with previous studies using mouselab (Glöckner and Betsch, 2008a; Johnson et al., 2008), information search was conducted mainly within-gambles. Most simple heuristics assume attribute-wise comparisons between gambles and therefore cannot account for the information search behavior.

AIM OF THE CURRENT STUDIES

The current paper elaborates the eye-tracking approach by Glöckner and Herbold for detailed investigations of the dynamics in risky choice, that is, changes of process variables over the time course of a decision. Furthermore, we investigate changes in pupil size as further dependent measure which indexes both processing load (Beatty, 1982) and arousal (Partala and Surakka, 2003; Bradley et al., 2008) and can be informative for cognitive processes in decision research as well (cf. Franco-Watkins and Johnson, 2011). Finally, we aimed to go beyond Glöckner and Herbold (2011) by using more detailed analyses of the factors influencing attention and by additionally addressing the question whether individuals react rather homogeneously or heterogeneously on these influence factors. By putting the decision process under the microscope, we aim to improve our knowledge concerning the mechanisms underlying risky decision making.

PREVIOUS FINDINGS CONCERNING DYNAMICS AND AROUSAL IN DECISION MAKING

Some previous studies have investigated the dynamics of decision making using eye-tracking in situations under certainty (i.e., outcomes are certain in contrast to probabilistic). A *gaze-cascade effects*, that is the tendency that over the course of decision making attention shifts to the chosen option, has been repeatedly found in these studies. Gaze-cascade effects, have for example been demonstrated in attractiveness based decisions between faces (Shimojo et al., 2003) and other kinds of visual decision task (e.g., Glaholt and Reingold, 2009a,b). Similarly, a recent eye-tracking investigation of food choices testing predictions of the attention drift-diffusion model showed that the last fixated item was chosen more often than the alternative item (Krajbich et al., 2010). Additionally, it was shown in this study that decision time and the number of fixations decrease with the increase in the difference in valuations of the food items measured in a pre-test.

There is one recent eye-tracking paper comparing risky decisions from description and from experience (Glöckner et al., 2012). The former are decisions between gambles with stated probabilities and outcomes as the ones used in the current study. In decisions from experience, in contrast, no probabilities are provided and outcomes have to be sampled sequentially (e.g., Camilleri and Newell, 2011). In the description condition, arousal measured by pupil dilation and skin conductance response increased with the average expected value (EV) of the two gambles and with decreasing difference in EV between gambles. This effect was not found in the experience condition. Based on this study and other studies showing differences in choice behavior (e.g., Barron and Erev, 2003; Hertwig et al., 2004; Ungemach et al., 2009), we restrict the current investigation to decisions from description only. Finally a recent investigation by Franco-Watkins and Johnson (2011) shows that the pupil dilation increases over the course of the decision making and is influenced by the presentation format (basic eye-tracking vs. decision moving-window). Another eye-tracking study using a somewhat different card gambling task shows that in situations in which persons can form expectations pupil dilation signals surprise and not expected reward or uncertainty *per se* (Preuschoff et al., 2011).

PROCESS MEASURES AND THEORIES

In the following process analysis, we investigate decision time, number of fixations, distribution of attention, mean fixation duration, pupil dilation, and direction of information search as dependent measures. First, we investigate these measures in an overall perspective aggregated over time for entire decisions (i.e., similar to most analyses in Glöckner and Herbold, 2011). Second, we look at developments and changes in these variables over the time course of making a decision by splitting up the decision process in several parts (i.e., time bins).

DEPENDENT MEASURES

Decision time was measured as the time from the gamble onset to individuals choice response. The "number of fixations" refers to

[1]For important limitations of CPT particularly in predicting choices between gambles with more than two outcomes see Birnbaum (2006, 2008a,b).

the average fixation count in this period (per decision of each person). The dependent variable "distribution of attention" refers to the proportion of fixations to specific parts of the screen containing probability or outcome information, so-called areas of interest (AOI). AOIs can thereby also be combinations of smaller areas such as all areas that contain pieces of information belonging to the left or the right gamble. The variable "mean fixation duration" refers to the average duration of single fixations in a decision. Stated differently, it refers to how long each fixation was on average. It has been shown that mean fixation durations increase with level of processing in scene perception in driving (Velichkovsky et al., 2001). This finding generalizes to decision making in that deliberate processes of calculating weighted sums go along with long fixations, whereas more intuitive or superficial information processing is accompanied by shorter fixations (Horstmann et al., 2009). The dependent variable "pupil dilation" refers to the difference in pupil size between periods of task processing (i.e., decision trials) and periods of rest (i.e., intertrial intervals/baseline) as a measure of arousal and cognitive load. The measure "direction of search" is defined as proportion of fixation transitions within one gamble as compared to the sum of transitions within and between gambles (details see below).

MODELS

In the investigation we particularly consider the models: DFT, PCS, PH, minimax, maximax, LEX, and (for completeness) a deliberate application of EU which we will refer to as weighted additive strategy (WADD). These decision strategies are briefly described in **Table 1**.

Table 1 | Models for risky choice.

AUTOMATIC INTEGRATION MODELS
Decision field theory (DFT): Decision making is based on a dynamic, stochastic process. Each moment in the choice process is akin to mentally sampling one of these outcomes, producing affective reactions to the imagined result which are added up until a threshold for deciding for one or the other gamble is reached. Sampling is assumed to reflect the probabilities present in the stimuli, therefore outcome probabilities dictate where attention shifts, but only the outcome values are used in determining the momentary evaluation
Parallel constraint satisfaction (PCS): Decision making is based on a dynamic process of constructing coherence under parallel consideration of all constraints given by outcome-probability relations in a decision task between gambles. In a process akin to Gestalt-construction in perception, activation of information supporting the favored gamble is automatically increased, while activation of information speaking for the alternative is decreased to form a coherent interpretation. The option with the higher activation feels more attractive and is more likely to be chosen
SIMPLE HEURISTICS
Priority heuristic (PH): Decision making is based on simple set of rules: (1) Go through reasons in the order of: minimum gain, probability of minimum gain, and maximum gain. (2) Stop examination if the minimum gains differs by 1/10 (or more) of the maximum gain; otherwise, stop examination if probabilities differ by 1/10 (or more) of the probability scale. (3) Choose the gamble with the more attractive gain (probability)
Minimax: Decision making is based on a simple rule of choosing the gamble with highest minimum outcome
Maximax: Decision making is based on a simple rule of choosing the gamble with the highest outcome
Lexicographic (LEX): Determine the most likely outcome of each gamble and choose the gamble with the better outcome. If both outcomes are equal, determine the second most likely outcome of each gamble, and choose the gamble with the better (second most likely) outcome. Proceed until a decision is reached
SERIAL EXPECTATION MODEL
Weighted additive strategy (WADD). Multiple outcomes by their probability and add them up for each gamble. Choose the gamble with the higher sum.

Heuristics are from Brandstätter et al. (2006) and Payne et al. (1993).

MANIPULATION AND PREDICTIONS

Not all models allow for straightforwardly deriving predictions concerning all dependent variables. Nevertheless, to foster improvements and specifications of models based on our data, we tried to derive as many reasonable predictions as possible (for a detailed discussion of specification and model development see Glöckner and Betsch, 2011). In cases in which reasonable predictions could be derived on theoretical grounds, we did so even if the authors of the original models did not explicitly make these predictions (which is of course explicitly acknowledged). Furthermore, we report data relatively broadly, so that the reader has additionally the possibility to check his or her own hypotheses.

DECISION TIME AND INFORMATION SEARCH

There are several possibilities to compare models. In the current paper, we refrain from complex comparative model fitting but use the classic method of hypothesis testing. We thereby investigate predictions concerning general differences in dependent variables as well as predictions concerning the effect of manipulations. Specifically, we manipulate the mean EV of gambles (EVmean) and the difference in EVs between gambles (EVdiff). EVmean is basically a manipulation of the stakes involved in the task. A high EVmean should increase motivation since participants can win more money. EVdiff is a measure how similar the gambles are from a rational perspective, and according to some models it should be related to the difficulty of choice. Choices between gambles with very similar EVs are challenging in terms of finding the option with the higher EV whereas in gamble pairs with a high EVdiff the better alternative is easier recognizable. The process models considered make quite distinct predictions concerning whether or not individuals should be influenced by manipulations of these factors. According to DFT and PCS, decision time and number of fixations should increase with higher EVmean since this corresponds to an higher internal decision/coherence threshold within these models. An increase in decision time and number of fixations would also be predicted for decisions between-gamble pairs with low EVdiff, as compared to gamble pairs with high EVdiff, since according to DFT the drift rate towards the threshold is lower in these cases and according to PCS it is much harder to create a coherent interpretation of the task. According to heuristics and

WADD, measures should not be influenced by either manipulation, because strategies apply rather simple decision rules or a standardized weighting operation, which are independent of these factors as long as the number of elementary information processes (EIPs; see Newell and Simon, 1972; Payne et al., 1988) necessary to apply the strategy is not influenced by the manipulation.

MEAN FIXATION DURATION AND PUPIL DILATION

Increasing EVmean and decreasing (absolute) EVdiff could both potentially lead to qualitative changes in information processing in that more thorough, effortful, and deliberate information processing is used. Hence, mean fixation duration might increase. Similarly, individuals might be more aroused and diligent, which should be reflected in increasing pupil size. DFT and PCS can both predict the effects of EVmean on pupil size. According to DFT, a higher EVmean will induce persons to use higher thresholds that necessitate a higher accumulation of affective responses[2]. According to PCS, arousal is dependent on the general level of conflict (or dissonance) in the network (Betsch and Glöckner, 2010; Glöckner et al., 2012; see also Hochman et al., 2010; Glöckner and Hochman, 2011), which can be measured by the networks Hopfield energy (Hopfield, 1982; Read et al., 1997). The level of remaining conflict mainly depends on two factors: the general activation of the network and the structure of the network defined by its' constraints (i.e., the structure of inhibitory and excitatory links). Increasing EVmean influences the first factor and leads to higher activation and therefore also to higher arousal. In the model this would be captured by the fact that higher *a priori* valuations of outcomes result in stronger links between the general valuation node and the outcome nodes (see Glöckner and Herbold, 2011, **Figure 1**). Furthermore, PCS predicts increasing pupil size for decreasing EVdiff, which influences the second factors namely the structure of the network and makes it harder to construct a coherent interpretation (Hochman et al., 2010; Glöckner and Hochman, 2011). Heuristics predictions are less influenced by either manipulation. Individuals applying heuristics should not be aware of EV as they ignore parts of the information which would be needed to calculate it. They could, however, anyway realize that gambles are concerned with higher stakes by seeing higher outcomes and therefore at least react to the manipulation of EVmean (see also text footnote 2, above).

DISTRIBUTION OF ATTENTION

Most process models allow predictions concerning the distribution of attention over outcomes within a decision. For example, LEX, minimax, and maximax predict that certain outcomes (i.e., low vs. high outcomes) should receive particularly large amounts of attention. For individuals applying WADD, in contrast, all outcomes and probabilities should receive about equal attention. From DFT, even more refined predictions can be derived in that attention can be assumed to be proportional to the probability of an outcome (Busemeyer and Townsend, 1993; Johnson and Busemeyer, 2005)[3].

[2]The threshold is assumed to model the importance of a decision (e.g., Johnson and Busemeyer, 2005, p. 847) and arguably subjective importance increases with the stakes that are in place in a decision. Johnson and Busemeyer (2005, p. 843) furthermore state that: "Each moment in the DFT choice process is akin to mentally sampling one of these numbers [i.e., monetary outcomes of gambles], producing an affective reaction to the imagined result" (explanation in brackets are added). Everything else being equal (e.g., drift rate), scaling up stakes of a decision should therefore increase the amount of affective reactions that has to be sampled and accumulated. Besides this eventual effect of increased accumulation, the increased size of outcomes *per se* should lead to increased affective reaction for each sampled outcome as well. Given the well established link between the (absolute) strength of affective responses and pupil dilation (e.g., Janisse, 1974), both effects could cause larger pupil dilations. It, however, has to be acknowledged that the researchers that suggested DFT have (to the best of our knowledge) not explicitly derived these predictions concerning arousal from their theory.

[3]According to DFT, attention weights (often equated with probabilities) are assumed to determine the stochastic sampling process of outcomes. Therefore, "the outcome probabilities dictate where attention shifts, but only the outcome values are used in determining the momentary evaluation." (Johnson and Busemeyer, 2005, p. 843). Note, however, that DFT in its original formulation is mainly concerned with mentally sampling outcomes. Explicit predictions concerning overt attention are only found in some implementation of simplified evidence accumulation models (e.g., Raab and Johnson, 2007), while other authors have argued that the process is not directly observable (e.g., Diederich, 2003).

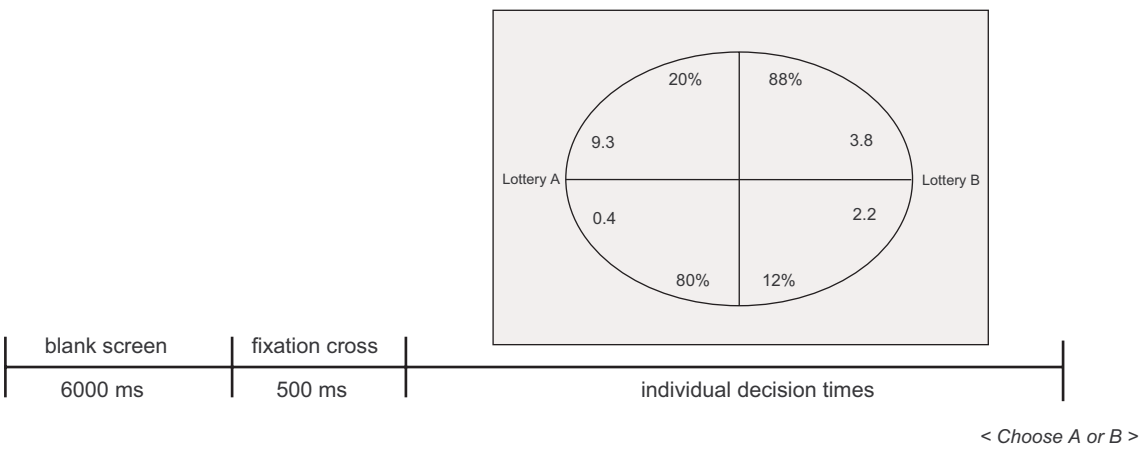


FIGURE 1 | Decision tasks between gambles with two outcomes each used in studies 1 and 2. Left outcomes belong to gamble A, right outcomes belong to gamble B.

PCS predicts that attention to outcomes should increase with both their probability and their value. PCS postulates that activation of the aspects supporting the later on favored option increases over time. Therefore, in a dynamic perspective, there should be a shift of attention toward the chosen options (and particularly to the most attractive outcomes of it) in the course of decision making. If one accepts the assumption that internal attention is at least correlated with overt attention in decisions from descriptions (Just and Carpenter, 1976), these predictions can be directly investigated in our fixation data.

DIRECTION OF INFORMATION SEARCH

WADD assumes mainly within-gambles information search. Heuristics, in contrast, mainly assume information search to be between gambles. DFT assumes a stochastic process of information sampling but does not explicitly specify the direction of information search. PCS also does not make clear predictions concerning the direction of information search.

MATERIALS AND METHODS

We conducted two eye-tracking studies to investigate the dynamics of risky choice. The first study used variants of the gambles taken from a previous study (Glöckner and Betsch, 2008a)[4]. To rule out the possibility that results might be dependent on the selection of these specific decisions, we conducted a second study using randomly generated gambles. To enhance cross-study comparisons, we will jointly report methods and results for both studies, although the studies were originally conducted in a logically motivated sequence and with a time lag of more than 1 year.

PARTICIPANTS AND DESIGN

Twenty-four residents of Bonn, most of them students (mean age: 24.6 years; 47% female), took part in the first study. Three of them had to be excluded from the analysis because their eye-tracking data was not recorded. In the second study, 37 participants took part (mean age: 22.53 years; 65% female). One had to be excluded because of incomplete eye-tracking data. Participants were recruited from the MPI Decision Lab subject pool using the database-system ORSEE (Greiner, 2004). In both studies participants repeatedly made decisions between two gambles with two outcomes each. In the first study, for many of the decisions the compared gambles were almost equal concerning their EV, that is, their EVdifference was close to zero, but the mean EV was manipulated to be high vs. low (details below). In the second study, mean EV and EV differences varied randomly between gambles within a certain pre-specified range.

Both studies lasted about 40 min each including a calibration phase and a short questionnaire about the decision behavior. Choices were incentivized. In addition to a fixed show-up fee of 6 € (approx. USD 7.90), participants earned money by playing one of the chosen lotteries (payoffs ranged from 0 to 49.5 €, $M = 6.2$€ (approx. USD 8.20) in Study 1 and from 0.1 to 49.8 €, $M = 9.6$ € (approx. USD 12.70, in Study 2). No deception was involved, and participants had not experienced deception in previous studies in our lab.

MATERIAL

Study 1

Participants completed 50 choices between pairs of gambles. The decision problems used consisted of 40 decision tasks adapted from Glöckner and Betsch (2008a), and 10 decision tasks taken from the standard economic test for risk aversion from Holt and Laury (2002). For the Glöckner-and-Betsch gambles, the mean EV of both gambles in the decision task was manipulated. For half of the decisions, the mean EV for both gambles was high (7.50 € ≤ EVmean ≤ 9.90 €) for the remaining decisions it was low (2.00€ ≤ EVmean ≤ 3.75€). The compared gambles were thereby always similar in EV to each other. In the Holt-and-Laury gambles, in contrast, the absolute difference in EVs between the gambles was varied gradually [0.32 € ≤ EVdiff (abs) ≤ 3.7 €]. Specifically, a safe gamble was compared to a risky gamble with increasing superiority in EV. Mean EV and differences in EVs between the gambles show a slight negative correlation, $r = -0.17$, $p = 0.23$. Decision tasks are listed in Section "Gambles Study 1" in Appendix.

Study 2

Decision tasks consisted of the 10 Holt-and-Laury gambles, and 40 decision tasks that were randomly generated. They contained pairs of gambles with varying mean EV (2.04 € ≤ EVmean ≤ 17.6 €) and varying the absolute difference in EV between gambles [0.006 € ≤ EVdiff (abs) ≤ 1.844 €]. Both factors were thereby almost uncorrelated ($r = -0.04$). The full list of decision problems can be found in Section "Gambles Study 2" in Appendix.

PROCEDURE

Both studies essentially used the same procedure. First, participants were informed about the experimental task, the incentive scheme, and the presentation format of the gambles. They were instructed to make good decisions, but to proceed as fast as possible. Each decision started with a blank screen (6 s), followed by the fixation cross (0.5 s) to direct attention to the center of the screen. Then the two gambles were presented simultaneously on the right and on the left side of the screen. An ellipsoid display format was used, in which all pieces of information (i.e., outcomes and probabilities) are present at equal distance from the initial fixation point (**Figure 1**). The ellipsoid format has been introduced in previous research, and it has been shown that choices are not systematically different from other more classic formats (see Glöckner and Betsch, 2008a; Glöckner and Herbold, 2011). The left (right) gamble was selected by pressing a key on the left (right) side of the keyboard (i.e., "y" and "m," which are the lower left and right letters, respectively, on German keyboards). Decision tasks were shown in randomized order. The presentation order of the gambles (gamble is presented on the left or right side of the screen) and the order of outcomes within each gamble (i.e., low outcome first vs. high outcome first) were also chosen according to a fixed random assignment.

Eye movements and pupil dilations were recorded by using the eye gaze binocular system (LC Technologies) with remote binocular sampling rate of 120 Hz and an accuracy of about 0.45°. Gambles were presented on a monitor (Samsung Synchmaster 740B,

[4]These gambles were also used in further eye-tracking studies by Glöckner and Herbold (2011) and Franco-Watkins and Johnson (2011) and comparative model fitting (together with additional gambles) in Glöckner and Pachur (2012).

refresh rate 60 Hz, reaction time 5 ms) with a native resolution of 1280 × 1024. Fixations were identified using a 30 pixel tolerance (i.e., added max-min deviation for *x*- and *y*-coordinates) and a minimum fixation time threshold of 50 ms. For analyses of mean fixation duration, first and last fixations in each decision trial were dropped since their length might be contaminated (see Krajbich et al., 2010; Glöckner and Herbold, 2011), although conclusions remained the same when including them in the analysis.

Non-overlapping AOI around the presented information on the screen were defined with the size of 100 × 100 pixels. These AOIs were used to determine which information was searched in a specific moment.

RESULTS

We first analyze dependent measures aggregated over the entire time course of decision making. In the second step, we look more closely at dynamics and systematic changes over time.

AGGREGATED PERSPECTIVE

Choices

First, we were interested in the influence of EVdiff on choices; therefore we conducted logistic regressions predicting choice for Gamble A by the difference in EV between Gambles A and B (positive numbers indicating an advantage of A over B and vice versa for negative numbers). This and all following regressions use random effects models to account for the repeated measurement design. The estimated models include random intercepts and slopes for all relevant predictors (excluding control factors). Coefficients *c* are assumed to vary randomly and independently from each other on the level of participants around their population mean μ_C with a SD of σ_C following a normal distribution $N[\mu_C, \sigma_C]$ (Wooldridge, 2001)[5]. As a further control factor, we included trial number to account for order effects. We find coefficients for choosing Gamble A higher than 0, indicating that the probability for choosing this gamble increases with increasing EV difference in favor of the gamble (**Table 2**).

Table 2 | Logistic regression predicting choices for gamble A ($p_{choiceA}$).

	$p_{choiceA}$ Study 1	$p_{choiceA}$ Study 2
EVdiff$_{(Gamble\ A-Gamble\ B)}$[a]	0.332*** (4.49)	0.801*** (12.63)
Constant	−0.124 (−0.98)	0.227* (2.13)
Observations	1026	1800

*Random effects model with varying intercepts and slopes for EVdiff. Order effects have been included as control factors in the regression (not reported). Reported are raw coefficients (z statistics in parentheses); number of participants were N = 21 and 36 in Studies 1 and 2, respectively. *p < 0.05, ***p < 0.001. [a]Values are in Euro difference.*

The influence of mean EV and EV difference on information search and processing

We were first interested in the influence of our manipulations of EVmean and absolute EVdiff on information search and information integration. As dependent measures, we analyzed decision time, number of fixations, mean fixation duration (excluding first and last fixations), and pupil dilation. Time was measured from gamble onset until key-press, so was the number of fixations (i.e., count), and the mean fixation duration (i.e., averaged duration of single fixations over the time-frame). Pupil dilation was calculated as peak pupil dilation scores, that is, the maximum increase of pupil size from baseline (measured at blank screen and fixation cross before each decision) in the same time period. Pupil dilation is measured as radius in mm. Descriptive statistics for the core dependent measures are provided in Section "Summary Statistics" in Appendix.

All dependent measures were regressed on EVmean and absolute EVdiff using random effects models with random intercepts and random slopes for both predictors and trial as control factor (**Table 3**). In the second study, we find that decision time, number of fixations, mean fixation duration, and pupil dilation increase with EVmean. Regression coefficients provide quantitative estimates for the influence. An increase in mean EV of 1.00€, for example, increases decision time by 0.197 s, the number of fixations by 0.701, etc. Except for mean fixation duration, we find trends in the same directions in Study 1, which do not reach conventional levels of significance, however. It remains unclear here whether these non-significant results are caused by the lower power due to reduced sample size or other factors such as the more systematic gamble construction and the lower variation of EVmean in Study 1.

Furthermore, we find that decision time, and number of fixations increase with decreasing EVdiff (abs) between gambles in both studies. Mean fixation duration also increases with decreasing EVdiff (abs) in both studies but only in the first study the effect reaches conventional significance levels. An increase in EVdiff of 1 € would here result in an decrease in mean fixation duration of around 2%[6].

To further investigate the effect of the EVmean manipulation on mean fixation durations, we categorized single fixations as short (<150 ms), medium (≥150 and <500 ms), and long (≥500 ms) according to their duration (Velichkovsky, 1999; Horstmann et al., 2009). In both studies (and in both high and low EVmean gambles) we find mainly medium and short fixations (Study 1: $M_{short} = 42.09\%$, $M_{medium} = 54.81\%$; Study 2: $M_{short} = 41.91\%$, $M_{medium} = 55.89\%$) and only a few long fixations that indicate more deliberate processing. Mixed effect regressions with proportion of short, medium, and long fixations as dependent variables and EV mean as predictor indicate that in both studies neither of the proportions is significantly influenced by EV mean, all *p*'s > 0.13.

Influence of probability and outcome on attention

To investigate driving factors for the distribution of attention, we regressed the amount of fixations toward each outcome on its probability, its value, and their interaction. As in the regression reported above, we corrected for the repeated measurement

[5]Alternative analyses using fixed effect correction and cluster corrected SE lead to the same conclusions except where explicitly stated otherwise.

[6]We also run all analyses reported in **Table 4** including additional interaction effects of EVmean and EVdiff (abs). As expected, main effects did not change but interactions turned out to be significant and positive for mean fixation duration in Study 1 and for decision time and number of fixations in Study 2.

Table 3 | Regression models predicting decision time, number of fixations, fixation durations, and pupil dilation by EVmean and EVdiff (abs).

	Decision time[b]		Number of fixations		Mean fixation duration[c]		Pupil dilation	
	Study 1	Study 2	Study 1	Study 2	Study 1	Study 2	Study 1	Study 2
EVmean[a]	0.0196	0.197***	0.0184	0.701***	−0.0003	0.0004***	0.0011	0.0016**
	(0.29)	(5.64)	(0.08)	(5.62)	(−0.93)	(4.56)	(0.88)	(2.81)
EVdiff (abs)[a]	−0.715**	−0.93***	−2.363**	−3.624***	−0.0042***	−0.0005	0.0029	0.0015
	(−2.95)	(−7.09)	(−2.95)	(−7.26)	(−4.21)	(−0.91)	(0.56)	(0.53)
Constant	8.891***	8.871***	33.24***	33.31***	0.199***	0.197***	0.107***	0.0743***
	(−3.88)	(14.92)	(9.85)	(−7.26)	(26.36)	(43.69)	(5.20)	(7.40)
Observations	1005	1728	1004	1728	1002	1728	995	1682

*Random effects model with varying intercepts and slopes for EVmean and EVdiff (abs). Order effects have been included as control factors in the regression (not reported). Reported are raw coefficients (z statistics in parentheses); number of participants were N = 21 and 36 in Studies 1 and 2, respectively. EVdiff (abs) denotes the absolute (i.e., none-negative) value of the difference between EVs of the gambles. **p < 0.01, ***p < 0.001. [a]Values are in Euro (difference). [b]Regression with log transformed decision times yield the same conclusions. [c]The random effect model failed to calculate the SE and as a result did not provide a stable estimate of the coefficients. Therefore we report for mean fixation duration a fixed effect model with cluster corrected SE (results and interpretation stay the same for both analysis).*

Table 4 | Regression models predicting fixations to outcomes by value and probability.

	Number of fixations	
	Study 1	Study 2
Value[a]	0.0741*** (4.97)	0.0913*** (9.78)
Probability[b]	1.533*** (8.21)	1.707*** (8.30)
Value × probability	0.0685+ (2.03)	0.125*** (7.04)
Constant	5.147*** (10.63)	5.428*** (15.37)
Observations	3892	6793

*Random effects model with varying intercepts and slopes for value and probability. Order and position effects have been included as control factors in the regression (not reported). Reported are raw coefficients (z statistics in parentheses); number of participants were N = 21 and 36 in Studies 1 and 2, respectively. +p < 0.10, ***p < 0.001. [a]Values are in Euro (centered). [b]Probabilities vary between 0 and 1(centered).*

design by using a random effects model with random intercept and random slopes for all three predictors. Furthermore we corrected for display position (i.e., all four combinations of right/left × up/down) using three display position dummies as well as for learning effects over trials by including trial number. In both studies the amount of fixations spent on an outcome increases with both its probability and the value of the outcome (**Table 4**). The interaction of probability and value is significant in Study 2 and marginally significant in Study 1. Significant main effects (but no interactions) are also found in regressions using fixation time as dependent variable (not reported).

Some heuristics models assume interindividual differences in that persons use qualitatively different strategies to make risky choices (Payne et al., 1988, 1992; Gigerenzer and Todd, 1999). These should be reflected in heterogeneity of attention patterns between individuals. We conducted regressions per individual to investigate eventual heterogeneity in the effects of probabilities and outcomes on fixations. **Figures 2** and **3** plot the resulting intercepts against the slopes from these regressions, each dot indicating one participant. We were thereby mainly interested in slopes. Interestingly, for both effects of probability and value on the number of fixations almost all participants showed behavior that was qualitatively in line with the overall regression results in that slopes for probabilities and values were positive. Hence, although slopes and intercepts differ between persons behavior seems to be relatively homogeneous.

DYNAMIC PERSPECTIVE

Furthermore, we investigated dynamics over the course of decision making. Some models predict changes in attention. Heuristics, such as PH and LEX, predict changes of attention from more important to less important comparisons over time. Most strategies considered here except for PCS (and also the above mentioned attention drift-diffusion models), however, predict that attention should be about equally distributed over the two gambles that are compared. As introduced above, Glöckner and Herbold (2011) showed an attention bias toward the option chosen later on in risky choices and gaze-cascade effects, that is a tendency to increasingly look at options that are later on chosen, have been demonstrated repeatedly in decisions under certainty (e.g., Shimojo et al., 2003). We investigate whether we can replicate the attentional bias and aim to explore in more detail when the bias starts to occur in the decision process. We expect to find a gaze-cascade effect in risky choice, in that the attentional bias is particularly driven by late fixations.

Attentional bias

To investigate the occurrence of an attentional bias over the course of decision making, the proportion of fixations to the left gamble is used as the dependent measure. **Figure 4** plots the attention proportion against proportional time bins with each bin containing 10% of the absolute decision time per person and decision.

Descriptively, the attentional bias toward the favored gamble is replicated. In both studies, a strong separation can be seen in the last third of the decision process (i.e., starting in the seventh or eighth bin). In both studies, we also consistently see an initial attention bias to the left which is driven by reading direction.

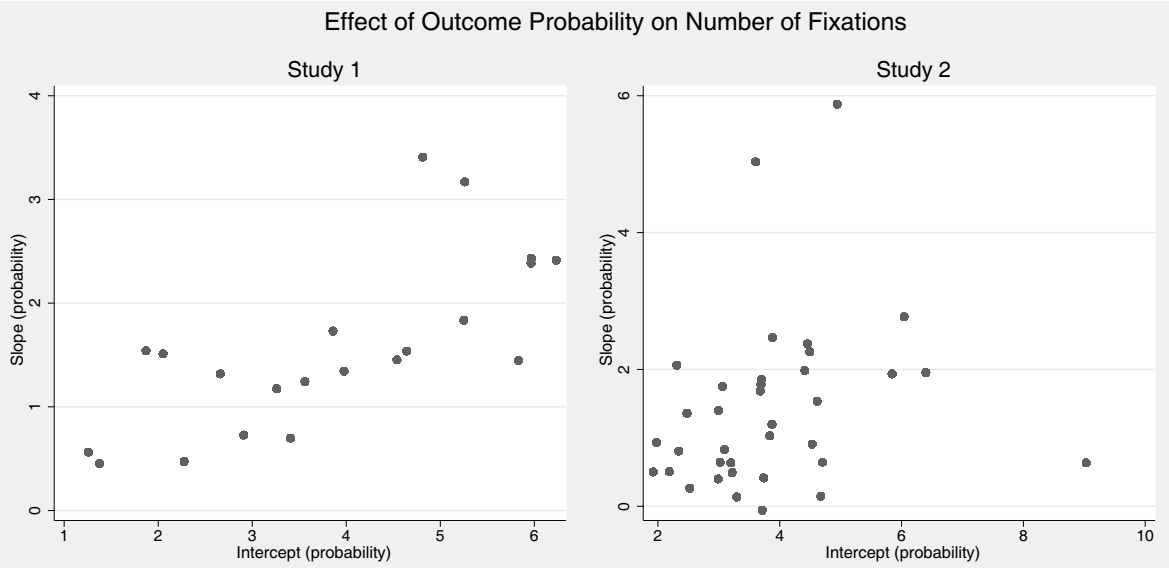


FIGURE 2 | Individual regression coefficients for predicting number of fixations by probability of outcomes (controlling for value). Graphs show intercepts plotted against slopes with positive slopes indicating that the number of fixation increases with probability.

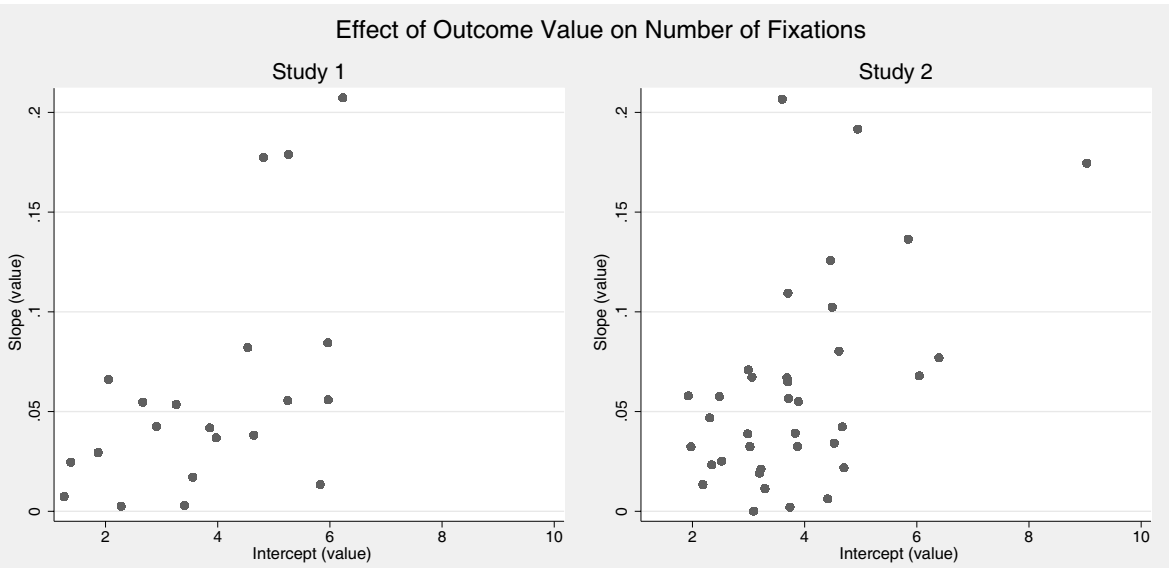


FIGURE 3 | Individual regression coefficients for predicting number of fixations by value of outcomes (controlling for probability). Graphs show intercepts plotted against slopes with positive slopes indicating that the number of fixation increases with value.

To test the results statistically, we regressed the proportion of fixation to the left gamble on choice (left/right), proportion of time (i.e., time bins), and their interaction (**Table 5**). We ran a random effects regression with random coefficients for intercept and slopes for all three predictors and included dummies for decision tasks as control factors. The significant main effect for choice indicates an overall attention bias in the direction of the chosen option. That is, if the left gamble was chosen, it was fixated 5–6% more often as compared to decisions in which the right gamble was chosen. The significant main effect for proportion of time on fixations to the left gamble indicates a general shift of fixations from left to right by 22–23% over the time course of decision making, probably driven by initial left bias due to natural reading order. More importantly, the significant interaction between choice and proportion of time indicates a strong gaze-cascade effect and captures the fact that the attentional bias mainly appears in the last part of the decision

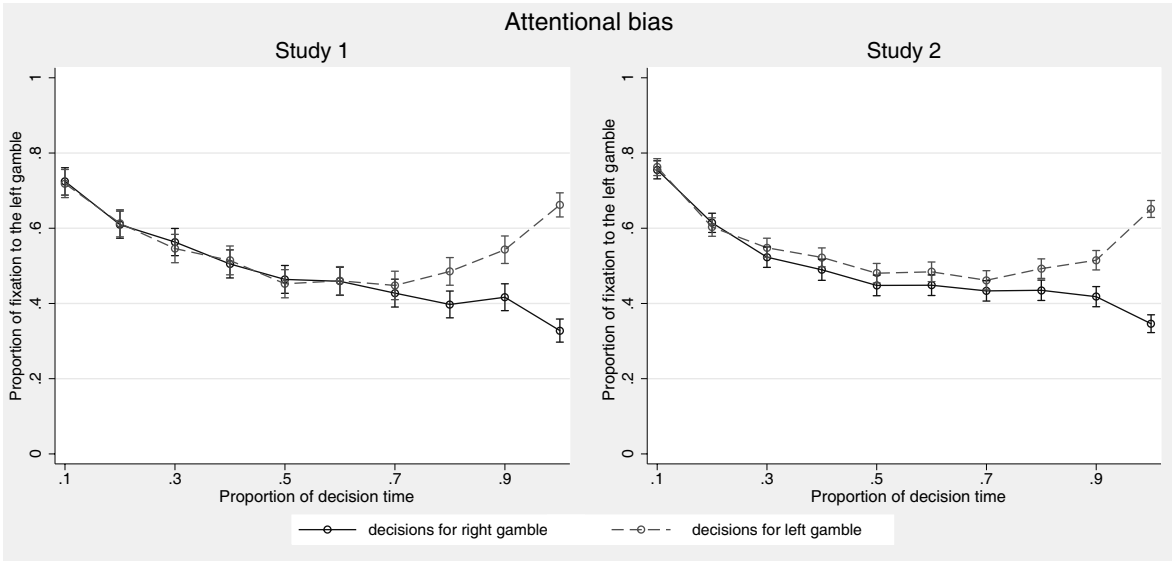


FIGURE 4 | Proportion of fixations to the left gamble (probabilities and outcomes) over the time course of the decision. Decisions for right vs. left gamble are depicted by separate lines. Error bars indicate 95% confidence intervals.

Table 5 | Regression model predicting proportion of fixations to the left gamble.

	Proportion of fixations to the left gamble	
	Study 1	**Study 2**
Choice (0 = right 1 = left)	5.088*** (3.71)	5.950*** (8.10)
Proportion of time[a]	−22.97*** (−4.79)	−22.16*** (−6.11)
Choice × proportion of time	27.24*** (7.56)	22.70*** (9.08)
Constant	47.68*** (15.46)	53.32*** (25.06)
Observations	9564	16944

*Random effects model with varying intercepts and slopes for choice, time, and the interaction. Task effects have been included as control factors in the regression (not reported). Reported are raw coefficients (z statistics in parentheses); number of participants were N = 21 and 36 in Studies 1 and 2, respectively. ***p < 0.001. The range of the variable is between 0 and 1 and we used bins of 0.10 for the analysis (both variables were centered for the analysis). Results stay essentially the same when excluding extreme values (i.e., p90, p10) or using fixation time as dependent measure.*

process in which the gamble receives attention which is chosen later on (independent of whether it is presented left or right)[7].

[7]For food choices, Krajbich et al. (2012) additionally report the effect that participants tend to choose the unlooked-at item if the last fixation was to a much disliked item. To investigate whether this effect generalizes to gambles, we rerun the regressions reported in **Table 5** additionally including a three-way interaction of EVdiff, choice, and time (also controlling for the respective main effect for EVdiff and the remaining two-way interaction; all variables centered). The results indicate that there is a positive interaction effect in study 2 but not study 1, whereby main effects and other interaction effects are not significantly changed. This lack of support, however, might be due to the fact that there were no losses involved in our material.

Pupil dilation

As reported above, on the aggregate level, we found that pupil dilation increased with EVmean, that is, with the stakes involved in the decision (note, however, that the effect was significant in Study 2 only). We were now interested in the development of pupil dilation over time. The measure of dilation peaks used in the aggregated analysis above is inappropriate for such an analysis of small time-blocks (bins), since no single peaks per time-bin can be expected. We therefore calculated for each time-bin average deviation scores in pupil size from the first time-bin. **Figure 5** shows that in both studies we observed an increase in pupil dilation over time. Additionally, there was an unexpected early drop in the second time-bin, which was observed in both studies and might eventually be caused by brightness changes due to stimulus onset.

To analyze pupil size development over time statistically, we regressed pupil dilation scores (i.e., pupil size minus pupil size in the first bin) on time, EVmean, absolute EVdiff, and their interactions with time (**Table 6**). We thereby again used a random effects model with intercept and slopes of all five predictors as random coefficients and controlling for order effects and the absolute time that passed within the trial. The analysis confirms that pupil dilation increases over the decision making process. The main effect for EVmean in Study 2 reproduces the effects from the overall analysis which used peak dilation scores (see **Table 3**, above).

Both in the aggregated and the dynamic analysis, we do not find a negative effect of EVdiff (abs) on pupil dilation, which stands in conflict with the findings reported in the introduction of this paper and fails to support predictions derived from PCS. Note, however, that in Study 1, EVdiff (abs) is only manipulated for a small subset of the gambles (i.e., the Holt–Laury gambles), so that eventual effects might have been hard to detect.

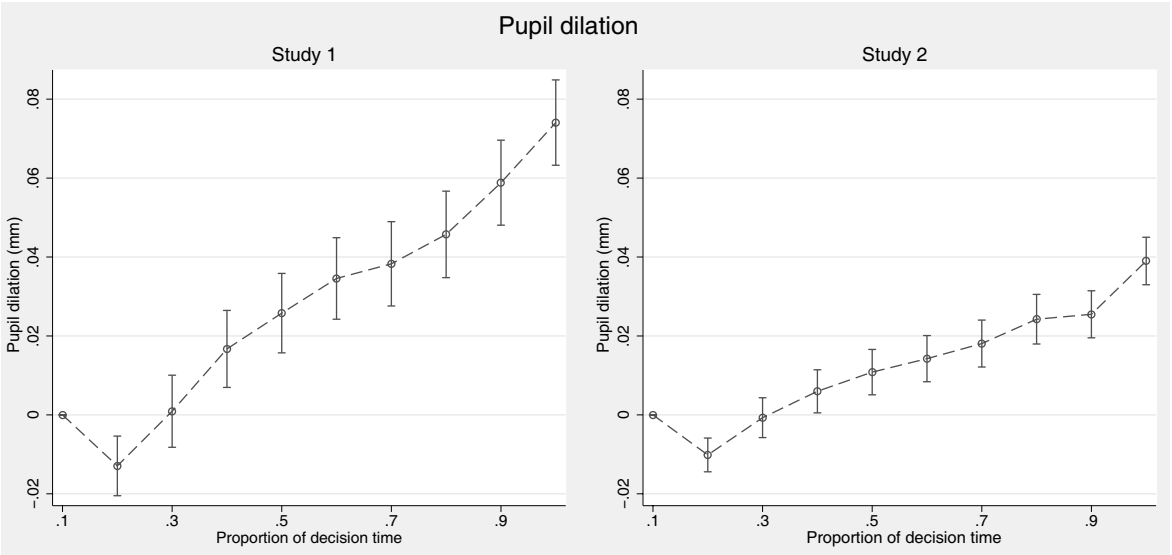


FIGURE 5 | Pupil dilation over time course of the decision. Error bars indicate 95% confidence intervals (Note that SEs are estimated assuming independence within participants since deviation scores are normalized per person and trial).

Table 6 | Regression predicting changes in pupil dilation over the decision making process.

	Pupil dilation	
	Study 1	**Study 2**
Proportion of time[a]	0.0434* (2.00)	0.0179^{+} (1.82)
EVmean[b]	0.000873 (0.80)	0.00105** (3.02)
EVdiff (abs)[b]	0.00369^{+} (1.93)	0.0027 (1.27)
Proportion of time × EVmean	0.00674 (1.39)	0.00703 (1.26)
Proportion of time × EVdiff (abs)	0.0140* (2.13)	0.00466(1.41)
Absolute time (in s)	0.00475*** (11.04)	0.0029*** (11.29)
Constant	0.0157 (11.04)	−0.00228 (−0.38)
Observations	9953	17427

Random effects model with varying intercepts and slopes for EVmean, EVdif(abs), time, and the interactions. Order effects have been included as control factors in the regression (not reported). Reported are raw coefficients (z statistics in parentheses); number of participants were N = 21 and 36 in Studies 1 and 2, respectively, $^{+}p < 0.10$, $*p < 0.05$, $**p < 0.01$, $***p < 0.001$. [a]The range of the variable is between 0 and 1 and we used bins of 0.10 for the analysis (and centered for the analysis). [b]Values are in Euro (and centered for the analysis).

Attention toward outcomes vs. probabilities

Changes in preferences for value and probability information during the decision can be informative for details of the decision process and were therefore analyzed as well. As mentioned above, some heuristics such as PH predict such shifts.

We analyzed the overall attention toward probabilities and outcomes of the gambles, as well as their development over time using the probability of fixations to probabilities as dependent measure (i.e., fixations to probability AOIs divided by fixations to all AOIs). As **Figure 6** shows, participants show a preference for probabilities very early in the decision process and a preference for outcomes later on. This is in line with one possible interpretation of DFT, which puts forward that individuals first have to learn probabilities, which are used later on to guide fixations to outcomes (see text footnote 3, above).

To investigate these effects statistically, we regressed the proportion of fixations dedicated to probabilities on time-bin, again using a random effects model and controlling for task and order (**Table 7**). The constant indicates that there is a general preference for value information but the significant main effect points to an even stronger reversal from preferring probabilities to preferring value information over the course of decision making.

Direction of information search

Finally, we investigated the direction of information search, which is one classic approach to investigate decision processes (e.g., Payne et al., 1988). As discussed above, WADD mainly predicts information search within-gambles, whereas heuristics mainly assume comparisons between gambles. We calculated the number of transitions within and between gambles as the number of times in which two subsequent fixations focused on different AOIs within the same gamble (within-gamble transition), as opposed to any AOI of the other gamble (between-gamble transition).

In line with previous findings (Franco-Watkins and Johnson, 2011; Glöckner and Herbold, 2011), we observe only a small proportion of transitions between gambles. Information search is mainly conducted within-gambles. We observe on average 5.12 (Study 1)/6.35 (Study 2) transitions between gambles and 14.44 (Study 1)/14.11 (Study 2) transitions within one of the gambles during each decision trial. We analyze changes in the proportion of transitions over the time course of decision making using a random effects model with randomly varying intercept and slope for

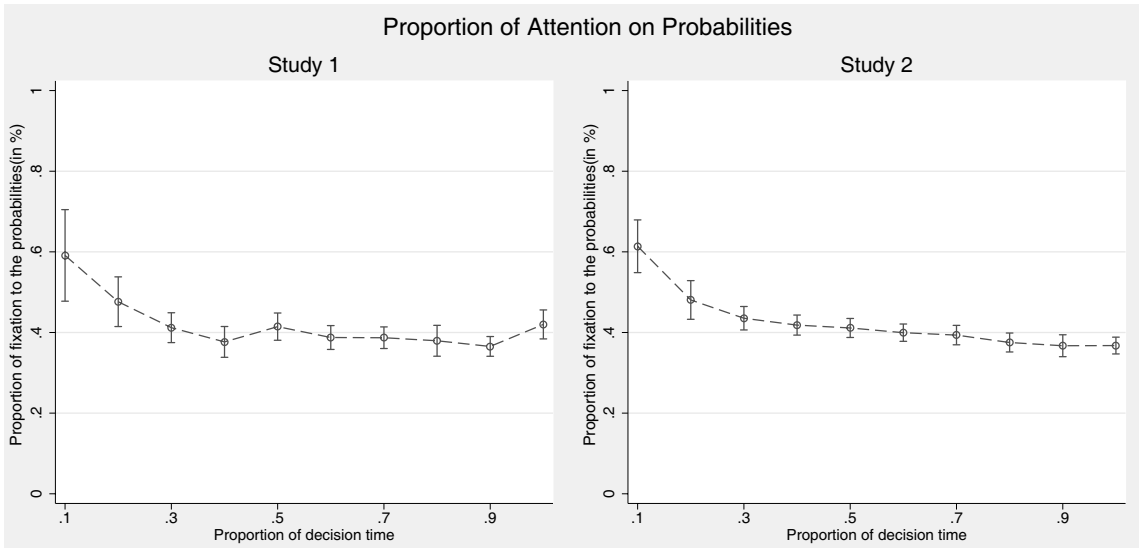


FIGURE 6 | Proportion of fixations to probabilities by time. Error bars indicate 95% confidence intervals.

Table 7 | Regression model predicting proportion of fixation to probability information over time.

	Proportion of fixations to probability information	
	Study 1	**Study 2**
Proportion of time (10%-steps)[a]	−14.69*** (−3.74)	−20.26*** (−7.61)
Constant	47.93*** (17.35)	49.62*** (23.90)
Observations	9428	16685

*Random effects model with varying intercepts and slopes for time. Order effects have been included as control factors in the regression (not reported). Reported are raw coefficients (z statistics in parentheses); number of participants were N = 21 and 36 in Studies 1 and 2, respectively. ***p < 0.001. [a]The range of the variable is between 0 and 1 and we used bins of 0.10 for the analysis.*

the time trend. Findings for both studies indicate that there is no change in the proportion of fixations by time (both z's < 1.16)[8].

SUMMARY OF RESULTS AND THEORY PREDICTIONS

A summary of the core results from the two reported eye-tracking studies with respect to the predictions of the theories that were considered in the current paper is provided in **Table 8**.

DISCUSSION

The current paper presents two studies which extend research on the cognitive mechanisms involved in risky choices with a special focus on dynamics, that is, changes in information search and processing over the time course of decision making. Eye-tracking is used to measure different indicators for information search and integration over time with the aim to test assumptions and predictions of existing process models for decision making to inspire further improvements and development within or outside the frameworks previously suggested.

In line with previous research (e.g., Glöckner and Betsch, 2008a; Ayal and Hochman, 2009; Glöckner and Herbold, 2011; Hilbig and Glöckner, 2011), we observed relatively short decision times, which speaks against the hypothesis that individuals deliberately integrate probabilities and outcomes according to a calculation of weighted sums. Note, that particularly in the second study we used randomly generated gambles that made calculations quite hard. Take for example the decision between Gamble A, which pays 9.3€ with 20% and 0.4€ otherwise vs. Gamble B which pays 3.8€ with 88% and 2.2€ otherwise. Even for mathematically skilled persons it seems hard to assume that they can deliberately calculate and compare the EVs of these Gambles within less than 10 s[9]. Nevertheless, participants' choices are still significantly predicted by the difference in EV between gambles and both outcomes and probabilities influence information search for almost all participants, which indicates that persons take into account value and probabilities and seem to integrate them to some degree. Additionally, we find mainly short and medium fixations over the entire time course of decision making, which supports the notion of individuals at least partially utilizing automatic processes in risky choice. DFT and PCS models describe possible process implementation that

[8]In this analysis, the application of a fixed effects model leads to different conclusions in that within gamble transitions increase over time in Study 2 ($b = 2.8$, $z = 3.01$, $p < 0.001$). Due to the inconsistency in analyses this result has to be interpreted cautiously. Note also that the finding conflicts with an opposite results by Glöckner and Herbold (2011).

[9]Note, that our manipulation of EVmean was partially confounded with mathematical complexity when assuming deliberate serial calculation of EVs. Specifically, gamble comparisons with high as compared to low mean EV also involve higher numbers that could make calculation harder and therefore increase decision time, number of fixations and arousal. Although we cannot completely rule out this possibility, we argue that given our findings concerning choices, decision time and process measures deliberate stepwise calculations of EV are not very likely in the first place.

Table 8 | Summary of the results.

Findings		Predicted by						
		DFT	**PCS**	**PH**	**Minimax**	**Maximax**	**LEX**	**WADD**
Decision time increases with	Higher EVmean	Yes	Yes	No	No	No	No	No
	Lower EVdiff	Yes	Yes	No	No	No	No	No
Average number of fixations increases with	Higher EVmean	Yes	Yes	No	No	No	No	No
	Lower EVdiff	Yes	Yes	No	No	No	No	No
Mean fixation duration increases with	Higher EVmean	Results are not conclusive						
	Lower EVdiff	Yes	Yes	No	No	No	No	No
Pupil dilation increases with	Higher EVmean	Yes	Yes	No	No	No	No	No
	Lower EVdiff	Results are not conclusive						
Attention increases with probability		Yes	Yes	No	No	No	Yes	No
Attention increases with value		No	Yes	No	No	Yes	No	No
Attention shift toward the chosen option		No	Yes	No	No	No	No	No
Mainly within-gambles information search		Unsp.	Unsp.	No	No	No	No	Yes
Pupil dilation increases over the time course of the decision		Yes	No	Unsp.	Unsp.	Unsp.	Unsp.	Unsp.

Unsp. = unspecified.

approximate weighted integration without relying on the unreasonable assumption that individuals indeed calculate them in a serial manner.

Quick decision responses could, of course, also be explained by simple heuristics. Considering that Gigerenzer and Gaissmaier (2011) start their definition of heuristics as follows: "A heuristic is a strategy that ignores part of the information. . .," the high average proportion of inspected information (i.e., 93%) speaks against this explanation. Also the interpretation that persons scan all pieces of information but later on ignore parts of it in their decision process seems to conflicts with the data since probabilities and values of outcomes influence attention for almost all participants. Furthermore, individuals show mainly within-gamble information search and are sensitive to manipulations of EV which also speaks against the application of simple heuristics.

We manipulated mean EV and EV difference between gambles of the decision problems to investigate whether individuals' reactions to these manipulations can be better explained by one of the models considered. Indicators for information search and processing are influenced by both manipulations. Deciding about gambles with somewhat higher stakes (i.e., higher mean EV) results in longer decision times, a more comprehensive information search (amount of fixations), deeper processing (mean length of a single fixation), and signs of increased arousal and/or cognitive load (greater pupil dilation). However, not only the stakes of a gamble influence the processing, but also the overall similarity of the two gambles in the choice pair, as indicated by the absolute EV difference. Deciding between two gambles that are very close to each other with respect to EV was associated with an increase in decision times and number of fixations, as well as a larger average fixation duration. Summing up the evidence, we observe that individuals make choices in less then 10 s and thereby take into account almost all pieces of information. Choice behavior is thereby highly sensitive to manipulations in EVmean and EVdiff, which arguably requires individuals quickly to develop rough impressions of the EVs of the gambles considered.

Since individuals seem to integrate many available pieces of information in a fast automatic response, these results provide evidence in favor of automatic integration models such as PCS and DFT. Results do not fit with the assumptions of simple heuristics ignoring parts of the information presented, or the much slower serial integration of information proposed by WADD.

Since most process models provide predictions about the distribution of attention, we used the amount of fixations on each outcome as a dependent measure. We were able to show that the look-up rate of an outcome depends on the value and the probability of an outcome and that most individuals show these effects. This result is challenging not only for heuristics, which often predict that attention is mainly driven by the value of an outcome (i.e., minimax and maximax), but also for models like DFT which assume that the attention to an outcome is only driven by its probability. Only PCS among the models considered here, which predicts holistic integration of all presented information, could account for the influence of both value and probability of an outcome on the attention devoted to it.

The results from both of our studies as well as evidence from previous research (Glöckner and Betsch, 2008a; Johnson et al., 2008; Franco-Watkins and Johnson, 2011; Glöckner and Herbold, 2011), show that within-gamble comparisons were much more frequent than between-gamble comparisons in these two-outcome risky choice tasks. This result supports WADD, but speaks against simple heuristics, which assume mainly attribute-wise comparisons between the gambles. DFT and PCS do not predict a specific direction of information search, and we suggest extending and specifying them by adding testable models of information search to increase their empirical content (Glöckner and Betsch, 2011).

DYNAMIC PERSPECTIVE

A more in-depth analysis of attention processes across the time course of decision making is unique to this study. To pin-point dynamics in the process, we analyzed changes in the distribution of attention between outcomes and probabilities, changes in pupil

dilation, as well as the attention biases toward one of the options as the decision process unfolded.

Individuals preferred looking at probability information in the very beginning of a decision. During the middle and later parts of the decision, however, individuals focused more strongly on outcome information. This shift in attention toward outcome information is in line with the assumption of DFT that probability information informs later on sampling of outcomes. It is hard to explain with most simple heuristics. PCS does not make clear predictions concerning this shift.

We also observed an attention bias toward the preferred option occurring in the later part of decision making showing the gaze-cascade effect for risky choices. This could be due to fixations toward the later chosen gamble within the decision making process or fixations in order to confirm the choice. Furthermore, pupil dilations increased over time in both studies, which can be interpreted as an accumulation of arousal over the information search process, which could be explained by DFT.

MODELS FOR RISKY CHOICES

The evidence presented in this paper advances our understanding of the time course of risky choice as well as the underlying processes. Results generally seem to support automatic integration models. Nevertheless, neither DFT nor PCS can in their current specifications fully account for all findings. In particular, with regard to DFT, the assumption of a stochastic sampling and evidence accumulation process directed by probabilities of outcomes can only be partially supported. The results at hand show that the probability of an outcome occurring is an important factor for the distribution of attention as predicted by DFT. However, it is not the only factor influencing information sampling. The value of an outcome influences the allocation of attention as well. Some implementations of DFT also assume that the allocation of attention between gambles should be constant over the entire decision process, which is not in line with our finding of an emerging bias toward the actually chosen option over time (i.e., the gaze-cascade effect). Other implementations of evidence accumulation models such as the attention drift-diffusion model (e.g., Krajbich et al., 2012), however, could account for such effects, at least when they concern last fixations only.

Concerning PCS the results are somewhat equivocal as well. The results concerning the attention bias and the gaze-cascade effect in risky choice strongly support the idea of systematic information distortions, particularly the accentuation of initial advantages of one option over time (Thagard, 1989; Holyoak and Simon, 1999; Simon et al., 2004; Glöckner and Herbold, 2011). Also the findings concerning the effects of EVmean on information search and arousal are in line with the models predictions and replicate and extend previous findings. However, it has to be qualified that the by PCS predicted and previously observed (Glöckner et al., 2012) effect of EVdiff on arousal was not found in the current studies. Also the observed increase in arousal over the time course of decision making is hard to explain with the PCS model. Furthermore, it has to be criticized that the PCS framework is currently not sufficiently specified to predict all information search parameters. Therefore, it seems necessary to extend the model by modeling information search more explicitly. The current findings, which provide a closer view of the choice process, can inform such model developments.

CAVEATS AND FURTHER RESEARCH

As already addressed in the introduction, one concern with regards to the method used in our experiment and the interpretations derived from our results could be that data from eye-tracking potentially provides only a vague proxy for the information search and processing, because it might neglect internal processes of attention and information retrieval. This would, for example, be the case if individuals look-up information only once and retrieve it from memory at a later stage of the decision making process. However, two arguments can be made against this objection: First, the current design (as opposed, for example, to mouselab) enabled effortless visual retrieval of information, making retrieval from memory unnecessary. Second, participants did actually constantly and systematically sample visual information across the entire decision process. These considerations allow for more faith in the methods, findings, and conclusions of the current study.

We could not replicate the previously observed effect of EV differences on arousal (Glöckner et al., 2012). For deeper understanding, we therefore strongly encourage further studies of the link between the decision task difficulty and observed arousal or cognitive load. Investigating the connection with additional measures, like skin conductance, could help to clarify the role of arousal and the driving factors.

Besides investigations of a methodological nature, a next step in analyzing information search and arousal should be replications of these findings in different decision making contexts by, on the one hand, changing the structure of the gamble tasks into more than two-outcome gambles and, on the other hand, using less abstract risky choices in order to test whether the result also holds for affect richer decisions and can therefore be generalized to risky decisions in the "real world" (see i.e., Goldstein and Weber, 1995, for a discussion of this caveat).

CONCLUSION

Given our results and evidence from previous studies (see Glöckner and Herbold, 2011), simple non-compensatory models such as the PH, LEX, minimax, or maximax heuristic do not seem to be appropriate to predict search behavior and processes involved in risky decisions in general. The same holds for the serial implementations of EU models in the form of WADD. Instead, the present results suggest that risky decision making seems to rely mainly on automatic-intuitive processes and can be partially described by models such as DFT and PCS. Nevertheless, also these models cannot account for all our findings in that they are underspecified in some respects and make no predictions or even make predictions which are clearly not in line with the findings. Since none of our findings, however, directly challenges core properties of these models, and also due to a lack of better alternatives, we think that both kinds of models are promising starting points for further theory developments concerning process models of risky choice.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at http://www.frontiersin.org/Cognitive_Science/10.3389/fpsyg.2012.00335/abstract

REFERENCES

Ayal, S., and Hochman, G. (2009). Ignorance or integration: the cognitive processes underlying choice behavior. *J. Behav. Decis. Mak.* 22, 455–474.

Barron, G., and Erev, I. (2003). Small feedback-based decisions and their limited correspondence to description-based decisions. *J. Behav. Decis. Mak.* 16, 215–233.

Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychol. Bull.* 91, 276–292.

Betsch, T., and Glöckner, A. (2010). Intuition in judgment and decision making: extensive thinking without effort. *Psychol. Inq.* 21, 279–294.

Birnbaum M. H. (2006). Evidence against prospect theories in gambles with positive, negative, and mixed consequences. *J. Econ. Psychol.* 27, 737–761.

Birnbaum, M. H. (2008a). New paradoxes of risky decision making. *Psychol. Rev.* 115, 463–501.

Birnbaum, M. H. (2008b). New tests of cumulative prospect theory and the priority heuristic: probability-outcome tradeoff with branch splitting. *Judgm. Decis. Mak.* 3, 304–316.

Birnbaum, M. H., and LaCroix, A. R. (2008). Dimension integration: testing models without trade-offs. *Organ. Behav. Hum. Decis. Process* 105, 122–133.

Bradley, M. M., Miccoli, L., Escrig, M. A., and Lang, P. J. (2008). The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology* 45, 602–607.

Brandstätter, E., Gigerenzer, G., and Hertwig, R. (2006). The priority heuristic: making choices without trade-offs. *Psychol. Rev.* 113, 409–432.

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.

Camilleri, A. R., and Newell, B. R. (2011). When and why rare events are underweighted: a direct comparison of the sampling, partial feedback, full feedback and description choice paradigms. *Psychon. Bull. Rev.* 18, 377–384.

Diederich, A. (1997). Dynamic stochastic models for decision making under time constraints. *J. Math. Psychol.* 41, 260–274.

Diederich, A. (2003). MDFT account of decision making under time pressure. *Psychon. Bull. Rev.* 10, 157–166.

Edwards, W. (1954). The theory of decision making. *Psychol. Bull.* 51, 380–417.

Franco-Watkins, A. M., and Johnson, J. G. (2011). Applying the decision moving window to risky choice: comparison of eye-tracking and mousetracing methods. *Judgm. Decis. Mak.* 6, 740–749.

Gigerenzer, G., and Gaissmaier, W. (2011). Heuristic decision making. *Annu. Rev. Psychol.* 62, 451–482.

Gigerenzer, G., and Todd, P. M. (1999). "Fast and frugal heuristics: the adaptive toolbox," in *Simple Heuristics that Make us Smart*, eds G. Gigerenzer and P. M. Todd (New York, NY: Oxford University Press), 3–34.

Glaholt, M. G., and Reingold, E. M. (2009a). Stimulus exposure and gaze bias: a further test of the gaze cascade model. *Atten. Percept. Psychophys.* 71, 445–450.

Glaholt, M. G., and Reingold, E. M. (2009b). The time course of gaze bias in visual decision tasks. *Vis. Cogn.* 17, 1228–1243.

Glöckner, A., and Betsch, T. (2008a). Do people make decisions under risk based on ignorance? An empirical test of the priority heuristic against cumulative prospect theory. *Organ. Behav. Hum. Decis. Process* 107, 75–95.

Glöckner, A., and Betsch, T. (2008b). Modeling option and strategy choices with connectionist networks: towards an integrative model of automatic and deliberate decision making. *Judgm. Decis. Mak.* 3, 215–228.

Glöckner, A., and Betsch, T. (2011). The empirical content of theories in judgment and decision making: shortcomings and remedies. *Judgm. Decis. Mak.* 6, 711–721.

Glöckner, A., Fiedler, S., Hochman, G., Ayal, S., and Hilbig, B. (2012). Processing differences between descriptions and experience: a comparative analysis using eye-tracking and physiological measures. *Front. Psychol.* 3:173. doi:10.3389/fpsyg.2012.00173

Glöckner, A., and Herbold, A.-K. (2011). An eye-tracking study on information processing in risky decisions: evidence for compensatory strategies based on automatic processes. *J. Behav. Decis. Mak.* 24, 71–98.

Glöckner, A., and Hochman, G. (2011). The interplay of experience-based affective and probabilistic cues in decision making: arousal increases when experience and additional cues conflict. *Exp. Psychol.* 58, 132–141.

Glöckner, A., and Pachur, T. (2012). Cognitive models of risky choice: parameter stability and predictive accuracy of prospect theory. *Cognition* 123, 21–32.

Goldstein, W. M., and Weber, E. U. (1995). Content and discontent: indications and implications of domain specificity in preferential decision making. *Psychol. Learn. Motiv.* 32, 83–136.

Greiner, B. (2004). The online recruitment system ORSEE 2.0 – a guide for the organization of experiments in economics. in *Working Paper series in Economics*, University of Cologne, Cologne.

Hertwig, R., Barron, G., Weber, E. U., and Erev, I. (2004). Decisions from experience and the effect of rare events in risky choice. *Psychol. Sci.* 15, 534–539.

Hilbig, B. E. (2008). One-reason decision making in risky choice? A closer look at the priority heuristic. *Judgm. Decis. Mak.* 3, 457–462.

Hilbig, B. E., and Glöckner, A. (2011). Yes, they can! Appropriate weighting of small probabilities as a function of information acquisition. *Acta Psychol. (Amst)* 138, 390–396.

Hochman, G., Ayal, S., and Glöckner, A. (2010). Physiological arousal in processing recognition information: ignoring or integrating cognitive cues? *Judgm. Decis. Mak.* 5, 285–299.

Holt, C. A., and Laury, S. K. (2002). Risk aversion and incentive effects. *Am. Econ. Rev.* 92, 1644–1655.

Holyoak, K. J., and Simon, D. (1999). Bidirectional reasoning in decision making by constraint satisfaction. *J. Exp. Psychol. Gen.* 128, 3.

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci. U.S.A.* 79, 2554–2558.

Horstmann, N., Ahlgrimm, A., and Glöckner, A. (2009). How distinct are intuition and deliberation? An eye-tracking analysis of instruction-induced decision modes. *Judgm. Decis. Mak.* 4, 335–354.

Janisse, M. P. (1974). Pupil size, affect and exposure frequency. *Soc. Behav. Pers.* 2, 125–146.

Johnson, E. J., Schulte-Mecklenbeck, M., and Willemsen, M. C. (2008). Process models deserve process data: comment on Brandstätter, Gigerenzer, and Hertwig (2006). *Psychol. Rev.* 115, 263–272.

Johnson, J. G., and Busemeyer, J. R. (2005). A dynamic, stochastic, computational model of preference reversal phenomena. *Psychol. Rev.* 112, 841–861.

Just, M. A., and Carpenter, P. A. (1976). Eye fixations and cognitive processes. *Cogn. Psychol.* 8, 441–480.

Krajbich, I., Armel, C., and Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nat. Neurosci.* 13, 1292–1298.

Krajbich, I., Lu, D., Camerer, C., and Rangel, A. (2012). The attentional drift-diffusion model extends to simple purchasing decisions. *Front. Psychol.* 3:193. doi:10.3389/fpsyg.2012.00193

Krajbich, I., and Rangel, A. (2011). Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions. *Proc. Natl. Acad. Sci. U.S.A.* 108, 13852–13857.

Luce, R. D. (2000). *Utility of Gains and Losses: Measurement-Theoretical and Experimental Approaches.* Mahwah, NJ: Erlbaum.

Luce, R. D., and Raiffa, H. (1957). *Games and Decisions: Introduction and Critical Survey.* New York: Wiley.

Milosavljevic, M., Malmaud, J., Huth, A., Koch, C., and Rangel, A. (2010). The drift diffusion model can account for the accuracy and reaction time of value-based choices under high and low time pressure. *Judgm. Decis. Mak.* 5, 437–449.

Newell, A., and Simon, H. A. (1972). *Human Problem Solving.* Oxford: Prentice-Hall Print.

Partala, T., and Surakka, V. (2003). Pupil size variation as an indication of affective processing. *Int. J. Hum. Comput. Stud.* 59, 185–198.

Payne, J. W., Bettman, J. R., Coupey, E., and Johnson, E. J. (1992). A constructive process view of decision making: multiple strategies in judgment and choice. *Acta Psychol. (Amst.)* 80, 107–141.

Payne, J. W., Bettman, J. R., and Johnson, E. J. (1988). Adaptive strategy selection in decision making. *J. Exp. Psychol. Learn Mem. Cogn.* 14, 534–552.

Payne, J. W., Bettman, J. R., and Johnson, E. J. (1993). *The Adaptive Decision Maker.* Cambridge: Cambridge University Press.

Preuschoff, K., Marius 't Hart, B., and Einhäuser, W. (2011). Pupil dilation signals surprise: evidence for noradrenaline's role in decision making. *Front. Neurosci.* 5:115. doi:10.3389/fnins.2011.00115

Raab, M., and Johnson, J. G. (2007). Expertise-based differences in search and option-generation strategies. *J. Exp. Psychol. Appl.* 13, 158–170.

Read, S. J., Vanman, E. J., and Miller, L. C. (1997). Connectionism, parallel constraint satisfaction processes, and Gestalt principles: (Re)introducing cognitive dynamics to social psychology. *Person. Soc. Psychol. Rev.* 1, 26–53.

Savage, L. J. (1954). *The Foundations of Statistics*, 2nd Edn. New York: Dover.

Schneider, W., and Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychol. Rev.* 84, 1–66.

Schulte-Mecklenbeck, M., Kühberger, A., and Ranyard, R. (2011). The role of process data in the development and testing of process models of judgment and decision making. *Judgm. Decis. Mak.* 6, 733–739.

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.

Shimojo, S., Simion, C., Shimojo, E., and Scheier, C. (2003). Gaze bias both reflects and influences preference. *Nat. Neurosci.* 6, 1317–1322.

Simon, D., Krawczyk, D. C., and Holyoak, K. J. (2004). Construction of preferences by constraint satisfaction. *Psychol. Sci.* 15, 331.

Simon, H. A. (1955). A behavioural model of rational choice. *Q. J. Econ.* 69, 99–118.

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychol. Rev.* 63, 129–138.

Thagard, P. (1989). Explanatory coherence. *Behav. Brain Sci.* 12, 435–467.

Tversky, A., and Kahneman, D. (1992). Advances in prospect theory: cumulative representation of uncertainty. *J. Risk Uncertain* 5, 297–323.

Ungemach, C., Chater, N., and Stewart, N. (2009). Are probabilities overweighted or underweighted when rare outcomes are experienced (rarely)? *Psychol. Sci.* 20, 473–479.

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

Usher, M., and McClelland, J. L. (2004). Loss aversion and inhibition in dynamical models of multialternative choice. *Psychol. Rev.* 111, 757–769.

Velichkovsky, B. M. (1999). From levels of processing to stratification of cognition. *Stratif. Cogn. Conscious.* 203–235.

Velichkovsky, B. M., Dornhoefer, S. M., Pannasch, S., and Unema, P. J. A. (2001). "Visual fixations and level of attentional processing," in *Proceedings of the Symposium on Eye Tracking Research and Applications*, ed. A. Duhowski (Palm Beach Gardens, NY: ACM Press), 79–85.

von Neumann, J., and Morgenstern, O. (1944). *Theory of Games and Economic Behavior*, 1st Edn. Princeton, NJ: Princeton University Press.

Wooldridge, J. M. (2001). *Econometric Analysis of Cross Section and Panel Data.* Cambridge, MA: MIT Press.

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 March 2012; accepted: 21 August 2012; published online: 01 October 2012.

Citation: Fiedler S and Glöckner A (2012) The dynamics of decision making in risky choice: an eye-tracking analysis. Front. Psychology ***3****:335. doi: 10.3389/fpsyg.2012.00335*

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

APPENDIX

SUMMARY STATISTICS

Table A1 provides the descriptive statistics for the core dependent measures.

Table A1 | Summary statistic for the main dependent variables (SD).

	Study 1	Study 2
Pupil size (radius in mm)	2.49 (0.43)	2.26 (0.30)
Pupil dilation (radius in mm)	0.122 (0.133)	0.095 (0.11)
Decision time (in s)	8.98 (6.27)	9.29 (5.99)
Proportion of inspected information (in %)[a]	93.92 (13.57)	96.01 (9.42)
Mean fixation duration (in s)	0.196 (0.040)	0.194 (0.033)
Number of fixations per decision	33.48 (21.01)	34.99 (21.26)
Number of transitions between information pieces	19.35 (11.56)	20.33 (11.99)
Transitions within-gambles (in %)[b]	74.15 (13.81)	69.48 (13.75)
Proportion of long fixations (in %)[c]	3.08 (1.73)	2.20 (1.4)

[a] *Proportion of AOIs fixated per trial, eight pieces of information are presented each trial.*

[b] *Reference class are all direct transitions between AOIs, re-fixations within the same AOI are dropped.*

[c] *Long fixations are fixation >500 ms and the reference class are all identified fixations.*

GAMBLES STUDY 1

Table A2 shows the 50 decision tasks used in study 1. Each Choice consists of two gambles A and B with two possible outcomes (and their probabilities). In the presentation, the positions of gambles and outcomes were varied according to a fixed random design.

Table A2 | Decision tasks used in study 1.

Decision	Gamble A	Gamble B	EV_A	EV_B	EVmean	EVdiff (abs)
1	0€ (0.60)/20€ (0.40)	10€ (0.40)/6.67€ (0.60)	8	8	8	0
2	4€ (0.50)/3€ (0.50)	0€ (0.30)/5€ (0.70)	3.5	3.5	3.5	0
3	1€ (0.45)/4€ (0.55)	4.4€ (0.60)/0€ (0.40)	2.65	2.64	2.645	0.01
4	10€ (0.7)/4€ (0.3)	10.25€ (0.80)/0€ (0.20)	8.2	8.2	8.2	0
5	10.5€ (0.50)/0.5€ (0.50)	11€ (0.75)/0€ (0.25)	8.25	8.25	8.25	0
6	0€ (0.30)/12€ (0.70)	10€ (0.60)/6€ (0.40)	8.4	8.4	8.4	0
7	7.5€ (0.50)/0€ (0.50)	2€ (0.65)/7€ (0.35)	3.75	3.75	3.75	0
8	3.3€ (0.40)/2.25€ (0.60)	4€ (0.67)/0€ (0.33)	2.67	2.68	2.675	0.01
9	0€ (0.40)/16€ (0.60)	22.5€ (0.40)/1€ (0.60)	9.6	9.6	9.6	0
10	4.5€ (0.70)/0€ (0.30)	0.3€ (0.50)/6€ (0.50)	3.15	3.15	3.15	0
11	5€ (0.60)/0€ (0.40)	0.4€ (0.55)/6.2€ (0.45)	3	3	3	0
12	11.5€ (0.70)/0.6€ (0.30)	10.3€ (0.80)/0€ (0.20)	8.23	8.24	8.235	0.01
13	0€ (0.25)/11.35€ (0.75)	16.25€ (0.50)/0.75€(0.50)	8.51	8.5	8.505	0.01
14	0.4€ (0.60)/7.3€ (0.40)	0€ (0.30)/4.5€ (0.70)	3.16	3.15	3.155	0.01
15	0€ (0.50)/4€ (0.50)	0.39€ (0.65)/5€ (0.35)	2	2	2	0
16	11.27€ (0.67)/0€ (0.33)	0.6€ (0.60)/18€ (0.40)	7.55	7.56	7.555	0.01
17	2€ (0.50)/3.4€ (0.50)	2.5€ (0.60)/3€ (0.40)	2.7	2.7	2.7	0
18	2€ (0.60)/5.25€ (0.40)	3€ (0.70)/4€ (0.30)	3.3	3.3	3.3	0
19	6.6€ (0.55)/11€ (0.45)	8.8€ (0.60)/8.25€ (0.40)	8.58	8.58	8.58	0
20	10€ (0.75)/6€ (0.25)	10€ (0.50)/8€ (0.50)	9	9	9	0
21	9€ (0.50)/7.5€ (0.50)	5€ (0.35)/10€ (0.65)	8.25	8.25	8.25	0
22	4.4€ (0.65)/15.4€ (0.35)	5.5€ (0.50)/11 € (0.50)	8.25	8.25	8.25	0
23	5.25€ (0.40)/2€ (0.60)	4€ (0.30)/3€ (0.70)	3.3	3.3	3.3	0
24	3.45€ (0.50)/1€ (0.50)	0.5€ (0.60)/4.8 € (0.40)	2.22	2.22	2.22	0
25	49.5€ (0.10)/5.5€ (0.90)	10€ (0.99)/0€ (0.01)	9.9	9.9	9.9	0
26	0€ (0.02)/9.44 € (0.98)	4.7€ (0.85)/35 € (0.15)	9.25	9.25	9.25	0
27	1.35€ (0.90)/12.5€ (0.10)	0€ (0.01)/2.5€ (0.99)	2.46	2.47	2.47	0.01
28	0€ (0.02)/3€ (0.98)	13.1€ (0.1)/1.8€ (0.9)	2.94	2.93	2.94	0.01
29	34€ (0.15)/3.5€ (0.85)	8.3€ (0.98)/0.1€ (0.02)	8.07	8.14	8.11	0.07
30	2.25€ (0.98)/0.1€ (0.02)	9€ (0.15)/1€ (0.85)	2.21	2.2	2.20	0.01
31	2.5€ (0.99)/0.3€ (0.01)	12.2€ (0.10)/1.4€ (0.90)	2.48	2.48	2.48	0
32	9€ (0.98)/0€ (0.02)	34.3€ (0.10)/6€ (0.90)	8.82	8.83	8.825	0.01
33	2.9€ (0.98)/3.2€ (0.02)	1.3€ (0.85)/12€ (0.15)	2.91	2.91	2.91	0
34	3.8.€ (0.85)/34 € (0.15)	8.3€ (0.98)/10 € (0.02)	8.33	8.33	8.33	0
35	1.44€ (0.85)/12€ (0.15)	3€ (0.98)/4€ (0.02)	3.02	3.02	3.02	0
36	4.58€ (0.80)/22€ (0.20)	15€ (0.01)/8€ (0.99)	8.07	8.07	8.07	0
37	17.2€ (0.15)/0.5€ (0.85)	3.2€ (0.01)/3€ (0.99)	3	3	3	0
38	31.2€ (0.20)/2.8€ (0.80)	8.5€ (0.99)/7.5€ (0.01)	8.48	8.49	8.485	0.1
39	0.9€ (0.80)/33.9€ (0.20)	8€ (0.01)/7.5€ (0.99)	7.5	7.5	7.5	0
40	22.5€ (0.15)/0.5€ (0.85)	2.15€ (0.98)/2.3€ (0.02)	3.8	2.153	2.98	1.65
41	3.2€ (0.90)/4€ (0.10)	7.7€ (0.10)/0.2€ (0.90)	3.28	0.95	2.12	2.33
42	3.2€ (0.80)/4€ (0.20)	7.7€ (0.20)/0.2€ (0.80)	3.36	1.7	2.53	1.66
43	3.2€ (0.70)/4€ (0.30)	7.7€ (0.30)/0.2€ (0.70)	3.44	2.45	2.95	0.99
44	3.2€ (0.60)/4€ (0.40)	7.7€ (0.40)/0.2€ (0.60)	3.52	3.2	3.36	0.32
45	3.2€ (0.50)/4€ (0.50)	7.7€ (0.50)/0.2€ (0.50)	3.6	3.95	3.78	0.35
46	3.2€ (0.40)/4€ (0.60)	7.7€ (0.60)/0.2€ (0.40)	3.68	4.7	4.19	1.02
47	3.2€ (0.30)/4€ (0.70)	7.7€ (0.70)/0.2€ (0.30)	3.76	5.45	4.61	1.69
48	3.2€ (0.20)/4€ (0.80)	7.7€ (0.80)/0.2€ (0.20)	3.84	6.2	5.02	2.36
49	3.2€ (0.10)/4€ (0.90)	7.7€ (0.90)/0.2€ (0.10)	3.92	6.95	5.44	3.03
50	3.2€ (0.00)/4€ (0.1)	7.7€ (1)/0.2€ (0.00)	4	7.7	5.85	3.70

GAMBLES STUDY 2

Table A3 shows the 50 decision tasks used in study 2. Each Choice consists of two gambles A and B with two possible outcomes (and their probabilities). In the presentation, the positions of gambles and outcomes were varied according to a fixed random design.

Table A3 | Decision tasks used in study 2.

Decision	Gamble A	Gamble B	EV_A	EV_B	EVmean	EVdiff (abs)
1	2.4€ (0.75)/0.8 € (0.25)	3.4€ (0.49)/0.8 € (0.51)	2	2.07	2.04	0.07
2	2.2€ (0.82)/2.8€ (0.18)	2€ (0.16)/2.3€ (0.84)	2.31	2.25	3.44	0.06
3	9.5€ (0.2)/0.6€ (0.8)	2.6€ (0.88)/1€ (0.12)	2.38	2.41	2.40	0.03
4	1€ (0.72)/7.8€ (0.28)	2.3€ (0.77)/4.6€ (0.23)	2.90	2.83	2.87	0.07
5	1.2€ (0.76)/5.3€ (0.24)	0.9€ (0.2)/2.4 € (0.8)	2.18	2.10	2.14	0.08
6	7€ (0.23)/1.5€ (0.77)	1.9€ (0.85)/8.1€ (0.15)	2.77	2.83	2.80	0.06
7	1.8€ (0.03)/2.4€ (0.97)	2.5€ (0.69)/2.1€ (0.31)	2.38	2.38	2.38	0
8	6.8€ (0.18)/1€ (0.82)	3.9€ (0.14)/1.8€ (0.86)	2.04	2.09	2.07	0.05
9	2.7€ (0.66)/2.9€ (0.34)	1.5€ (0.73)/6.3€ (0.27)	2.77	2.80	2.79	0.03
10	4€ (0.1)/1.9€ (0.9)	6.2€ (0.16)/1.3€ (0.84)	2.11	2.08	2.10	0.03
11	14.4€ (0.75)/4.8€ (0.25)	20.4€ (0.49)/4.8€ (0.51)	12	12.44	12.22	0.44
12	13.2€ (0.82)/16.8€ (0.18)	12€ (0.16)/13.8€ (0.84)	13.85	13.51	13.68	0.34
13	57€ (0.2)/3.6€ (0.8)	15.6€ (0.88)/6€ (0.12)	14.28	14.45	14.37	0.17
14	6€ (0.72)/46.8€ (0.28)	13.8€ (0.77)/27.6€ (0.23)	17.42	16.97	17.20	0.45
15	7.2€ (0.76)/31.8€ (0.24)	5.4€ (0.2)/14.4€ (0.8)	13.10	12.60	12.85	0.50
16	42€ (0.23)/9€ (0.77)	11.4€ (0.85)/48.6€ (0.15)	16.59	16.98	16.79	0.39
17	10.8€ (0.03)/14.4€ (0.97)	15€ (0.69)/12.6€ (0.31)	14.29	14.26	14.28	0.04
18	40.8€ (0.18)/6€ (0.82)	23.4€ (0.14)/10.8 € (0.86)	12.26	12.56	12.41	0.30
19	16.2€ (0.66)/17.4€ (0.34)	9€ (0.73)/37.8 € (0.27)	16.61	16.78	16.69	0.17
20	24€ (0.1)/11.4€ (0.9)	37.2€ (0.16)/7.8€ (0.84)	12.66	12.50	12.58	0.16
21	2.2€ (0.75)/0.6€ (0.25)	4.6€ (0.49)/2€ (0.51)	1.80	3.27	2.54	1.47
22	2€ (0.82)/2.6€ (0.18)	3.2€ (0.16)/3.5€ (0.84)	2.11	3.45	2.78	1.34
23	9.3€ (0.2)/0.4€ (0.8)	3.8€ (0.88)/2.2€ (0.12)	2.18	3.61	2.90	1.43
24	0.8€ (0.72)/7.6€ (0.28)	3.5€ (0.77)/5.8€ (0.23)	2.70	4.03	3.40	1.33
25	1€ (0.76)/5.1€ (0.24)	2.1€ (0.2)/3.6€ (0.8)	1.98	3.30	2.64	1.32
26	6.8€ (0.23)/1.3€ (0.77)	3.1€ (0.85)/9.3€ (0.15)	2.57	4.03	3.30	1.46
27	1.6€ (0.03)/2.2€ (0.97)	3.7€ (0.69)/3.3€ (0.31)	2.18	3.58	2.88	1.40
28	6.6€ (0.18)/0.8€ (0.82)	5.1€ (0.14)/3€ (0.86)	1.84	3.29	2.57	1.45
29	2.5€ (0.66)/2.7€ (0.34)	2.7€ (0.73)/7.5€ (0.27)	2.57	4	3.28	1.43
30	3.8€ (0.1)/1.7€ (0.9)	7.4€ (0.16)/2.5€ (0.84)	1.91	3.28	2.60	1.37
31	14.2€ (0.75)/4.6€ (0.25)	21.6€ (0.49)/6€ (0.51)	11.80	13.64	12.72	1.84
32	13€ (0.82)/16.6€ (0.18)	13.2€ (0.16)/15€ (0.84)	13.65	14.71	14.18	1.06
33	56.8€ (0.2)/3.4€ (0.8)	16.8€ (0.88)/7.2€ (0.12)	14.08	15.65	14.87	1.57
34	5.8€ (0.72)/46.6€ (0.28)	15€ (0.77)/28.8 € (0.23)	17.22	18.17	17.70	0.95
35	7€ (0.76)/31.6 € (0.24)	6.6€ (0.2)/15.6€ (0.8)	12.90	13.80	13.35	0.90
36	41.8€ (0.23)/8.8€ (0.77)	12.6€ (0.85)/49.8 € (0.15)	16.39	18.18	17.30	1.79
37	10.8€ (0.03)/14.2€ (0.97)	16.2€ (0.69)/13.8 € (0.31)	14.09	15.46	14.77	1.36
38	40.6€ (0.18)/5.8€ (0.82)	24.6€ (0.14)/12€ (0.86)	12.06	13.76	12.91	1.7
39	16€ (0.66)/17.2€ (0.34)	10.2€ (0.73)/39€ (0.27)	16.41	17.98	17.20	1.57
40	23.8€ (0.1)/11.2€ (0.9)	38.4€ (0.16)/9€ (0.84)	12.46	13.70	13.08	1.24
41	3.2€ (0.90)/4€ (0.10)	7.7€ (0.10)/0.2€ (0.90)	3.28	0.95	2.12	2.33
42	3.2€ (0.80)/4€ (0.20)	7.7€ (0.20)/0.2€ (0.80)	3.36	1.70	2.53	1.66
43	3.2€ (0.70)/4€ (0.30)	7.7€ (0.30)/0.2€ (0.70)	3.44	2.45	3	0.99
44	3.2€ (0.60)/4€ (0.40)	7.7€ (0.40)/0.2€ (0.60)	3.52	3.20	3.36	0.32
45	3.2€ (0.50)/4€ (0.50)	7.7€ (0.50)/0.2€ (0.50)	3.60	3.95	3.78	0.35
46	3.2€ (0.40)/4€ (0.60)	7.7€ (0.60)/0.2€ (0.40)	3.68	4.70	4.19	1.02
47	3.2€ (0.30)/4€ (0.70)	7.7€ (0.70)/0.2€ (0.30)	3.76	5.45	4.61	1.69
48	3.2€ (0.20)/4€ (0.80)	7.7€ (0.80)/0.2€ (0.20)	3.84	6.20	5.02	2.36
49	3.2€ (0.10)/4€ (0.90)	7.7€ (0.90)/0.2€ (0.10)	3.92	6.95	5.44	3.03
50	3.2€ (0.00)/4€ (0.1)	7.7€ (1)/0.2€ (0.00)	4	7.70	5.85	3.70

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 13 June 2012
doi: 10.3389/fpsyg.2012.00193

The attentional drift-diffusion model extends to simple purchasing decisions

Ian Krajbich[1,2], Dingchao Lu[1], Colin Camerer[1,3] and Antonio Rangel[1,3]*

[1] *Division of the Humanities and Social Sciences, California Institute of Technology, Pasadena, CA, USA*
[2] *Department of Economics, University of Zurich, Zurich, Switzerland*
[3] *Computational and Neural Systems, California Institute of Technology, Pasadena, CA, USA*

Edited by:
Konstantinos Tsetsos, Oxford University, UK

Reviewed by:
Adele Diederich, Jacobs University Bremen, Germany
Andreas Glöckner, Max Planck Institute for Research on Collective Goods, Germany

***Correspondence:**
Antonio Rangel, Division of the Humanities and Social Sciences, California Institute of Technology, 1200 East California Boulevard, Pasadena, CA 91125, USA.
e-mail: rangel@hss.caltech.edu

How do we make simple purchasing decisions (e.g., whether or not to buy a product at a given price)? Previous work has shown that the attentional drift-diffusion model (aDDM) can provide accurate quantitative descriptions of the psychometric data for binary and trinary value-based choices, and of how the choice process is guided by visual attention. Here we extend the aDDM to the case of purchasing decisions, and test it using an eye-tracking experiment. We find that the model also provides a reasonably accurate quantitative description of the relationship between choice, reaction time, and visual fixations using parameters that are very similar to those that best fit the previous data. The only critical difference is that the choice biases induced by the fixations are about half as big in purchasing decisions as in binary choices. This suggests that a similar computational process is used to make binary choices, trinary choices, and simple purchasing decisions.

Keywords: drift-diffusion, decision-making, neuroeconomics, decision neuroscience, eye-tracking, valuation, choice, purchasing

INTRODUCTION

A basic goal of decision neuroscience and neuroeconomics is to characterize the computations carried out by the brain to make different types of decisions (Busemeyer and Johnson, 2004; Smith and Ratcliff, 2004; Bogacz, 2007; Gold and Shadlen, 2007; Rangel et al., 2008; Kable and Glimcher, 2009; Hare and Rangel, 2010; Rushworth et al., 2011). Over the last decade, a sizable number of studies have found that standard drift-diffusion-models (DDM; Ratcliff, 1978, 2002; Busemeyer and Rapoport, 1988; Ratcliff and McKoon, 1997, 2008; Ratcliff and Rouder, 2000; Ratcliff and Smith, 2004; Leite and Ratcliff, 2010), as well as closely related versions such as the leaky competing accumulator (LCA) model (Usher and McClelland, 2001; Smith and Ratcliff, 2004; Tsetsos et al., 2011) provide quantitative explanations of the psychometrics, chronometrics, and neurometrics of perceptual choices. More recently, it has been shown that these models also provide good accounts of value-based choice (Basten et al., 2010; Krajbich et al., 2010; Milosavljevic et al., 2010; Philiastides et al., 2010; Hare et al., 2011; Krajbich and Rangel, 2011).

This class of models assumes that decisions are made by accumulating noisy evidence in favor of the different options. The combined evidence for each option is compared to that for the other options, and when the relative evidence for any option exceeds a pre-defined threshold, that option is chosen. One can think of the relative evidence signals as measures of the individual's confidence that each option is the correct choice, and thus the model implies that choices are made only when the subject is confident enough. For perceptual discrimination, the source of the noisy evidence comes from the stimulus itself. For value-based decision-making, the noisy evidence derives from how item values are computed and compared. Note, in particular, that the decision process involves the sequential and repeated sampling of the attractiveness of each option's individual attributes or characteristics. This introduces two sources of noise in the process: noise intrinsic to the sampling of attribute values, and noise due to random shifts in attention between the options which affect how the attribute values are sampled (Busemeyer and Townsend, 1993; Diederich, 1997; Roe et al., 2001; Busemeyer and Diederich, 2002; Usher and McClelland, 2004; Johnson and Busemeyer, 2005; Usher et al., 2008; Tsetsos et al., 2010).

In previous work we have shown that a variant of the DDM, which we refer to as the attentional drift-diffusion model (aDDM), provides quantitatively accurate predictions of the relationship between choices, reaction times, and visual fixations in experiments where subjects make either binary or trinary snack food choices (Krajbich et al., 2010; Krajbich and Rangel, 2011). A critical feature of the aDDM is that the evidence accumulation process depends on where the subject is looking, so that on average a subject accumulates more evidence for an item when it is being looked at than when it is not. A fitting of the model to the data found that subjects only accumulate about a third as much evidence for an item when it is not being looked at. This difference in the accumulation rate has important implications for the pattern and quality of decisions: choices are biased by their fixation patterns, i.e., the more time a subject spends looking at an appetitive item, the more likely he is to choose it. Importantly, in our previous work we found that a single model with common parameters was able to account for the data in both binary and trinary food choices, which suggests that the underlying processes exhibit some robustness.

This study seeks to advance our understanding of the properties, advantages, and limitations of the aDDM by investigating

if it can also provide an accurate description of purchasing decisions, and the extent to which the model's parameters need to change to explain this new class of decisions. In the simple purchasing decisions studied here, subjects see a product and a price, and have to decide whether or not to buy the item at that price. This type of decision is interesting because the subject needs to combine information from two very different types of stimuli: real world or rich pictorial representations (e.g., pictures of snack food) and symbolic/numerical information (i.e., the price tag). It is not obvious *a priori* if the aDDM will be able to account for the data in simple purchasing decisions, or if changes in the underlying model parameters will be required. For example, since the price is presented in numerical form, it is not obvious if the price information integrates dynamically and stochastically as in the DDM, or if in contrast it is incorporated using a more deterministic algorithm.

We present a computational model of the aDDM applied to simple purchasing decisions, and data from an eye-tracking experiment designed to address the following questions. First, how do subjects allocate fixation time between products and prices? Second, to what extent, if any, do visual fixations influence choices? Third, can the aDDM explain purchasing decisions with reasonable quantitative accuracy?

COMPUTATIONAL MODEL

In previous work we have proposed and tested a model of how visual fixations interact with the choice process to make choices between pairs of stimuli (e.g., apple or orange?; Krajbich et al., 2010). We refer to that model as the aDDM. Here we show that a simple parameter change in the model is also able to provide a good quantitative account of simple purchasing decisions.

In a simple purchasing decision, a subject is shown a product and a price, and has to decide whether or not he wants to purchase it at that price. Economists typically assume that an individual knows his value for a good, i.e., the amount that he is willing to pay for it in dollars. In this case the optimal strategy involves purchasing the item when its net value is greater than zero, and not purchasing the item otherwise. Net value is defined as the difference between the item's value and its price.

In contrast, the aDDM assumes that every purchasing decision involves the dynamic computation of a relative decision value (RDV) variable, which starts at zero at the beginning of each decision, and evolves over time until a choice is made (see **Figure 1A** for a graphical illustration). At any instant t within the decision process, the RDV variable (denoted by V_t) measures the current estimate of the relative value of purchasing the item minus the value of not purchasing it. The RDV evolves as a Markov Gaussian process until it reaches a barrier located at either $+1$ or -1. If it crosses the $+1$ barrier then the item is purchased, if it crosses the -1 barrier the item is not purchased.

A critical feature of the aDDM is that the mean rate of change (drift rate) of the RDV depends on the fixation location at that instant in time (as illustrated in **Figure 1A**). In particular, when the subject is looking at the item, the evolution of the RDV is given by

$$V_t = V_{t-1} + d\left(v - \theta p\right) + \varepsilon_t$$

and when he is looking at the price, it is given by

$$V_t = V_{t-1} - d\left(p - \theta v\right) + \varepsilon_t$$

where V_t is the value of the RDV at time t within the decision trial, v denotes the subject's value for the product being considered for purchase, p denotes the price of the purchase, d is a constant controlling the speed of integration (in units of $\$^{-1}\text{ms}^{-1}$), θ between 0 and 1 is a parameter reflecting the bias toward the fixated option, and ε is white Gaussian noise with variance σ^2 (randomly sampled once every ms). Note that at any point in a trial the RDV is evolving according to one of these two drift rates, depending on where the subject is looking. When the subject shifts his gaze to the other option, the RDV continues to evolve from its current value, but with the other drift rate.

The model assumes that fixations are generated by a stochastic process that is independent of the value of the path that the RDV takes during the trial. Note that this does not rule out the possibility that the fixation process could be affected by the latent value of the product or by the price. In fact, as is described in the methods section below, all of our analyses assume that fixation locations and durations are drawn from the observed empirical distribution, and that the integration process within each fixation continues until the end of the fixation, in which case another fixation is drawn, or until a barrier is crossed, in which case the trial ends. This approach allows us to investigate the comparator process in detail, while taking the fixation process as exogenously given.

Several features of the model are worth highlighting. First, the model has three free parameters: the bias parameter θ, the slope parameter d, and the noise parameter σ. As discussed below and in our previous work (Krajbich et al., 2010; Krajbich and Rangel, 2011) changes in the value of these parameters have important qualitative and quantitative implications for the accuracy and speed of choice. Second, the variables v and p reflect properties of the stimuli being considered, and thus are not parameters of the decision process. In particular, the experimental design described below allows us to measure v and p for each subject and trial independently of the actual purchase decisions, and this information is used to estimate the free parameters of the aDDM. Third, the model includes the standard DDM, in which fixations do not matter, as a special case when $\theta = 1$ (Ratcliff and McKoon, 2008; Milosavljevic et al., 2010). Fourth, the model is almost identical to the one that we have previously proposed for binary choice (Krajbich et al., 2010), except for a slight change in the nature of the stimuli and responses. In particular, in our previous work subjects were shown pictures of two food items, one of the left and one on the right, and they had to choose one of them with a button press. In contrast, here subjects were shown a more complex screen consisting of a picture of an item and a price, and had to indicate with a button press whether they wanted to buy the item at that price. Fifth, when $\theta < 1$ the model predicts that random fluctuations in fixations affect choices. In particular, items are more likely to be purchased when subjects fixate more on the product relative to the price, and less likely to be bought when they fixate relatively more on the price. The intuition is simple

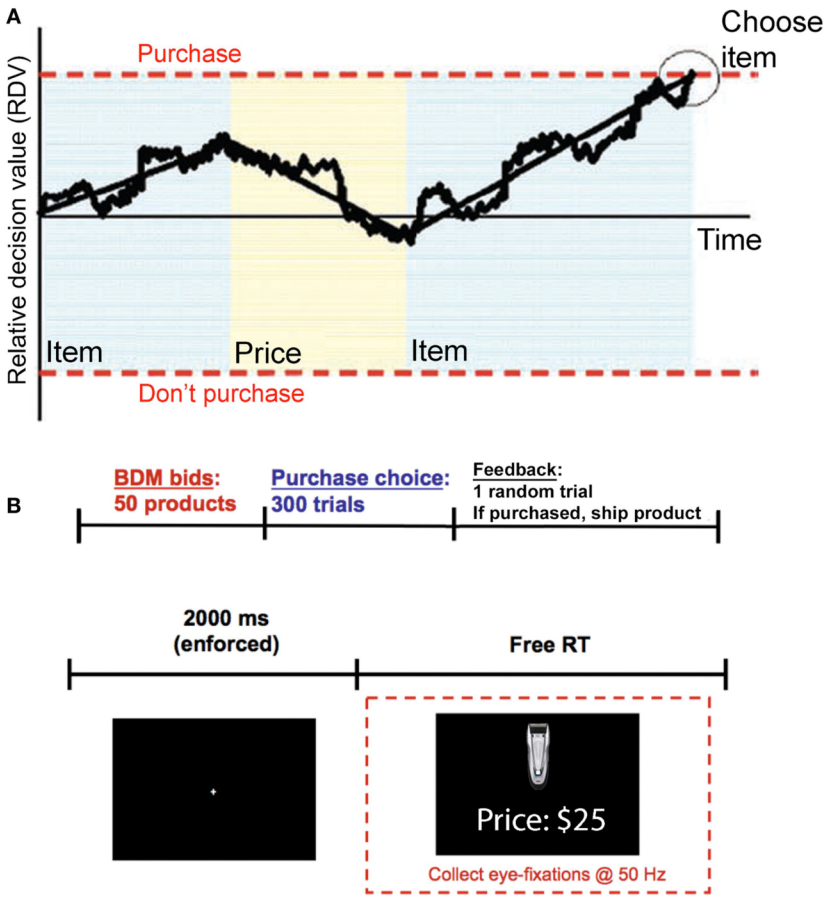


FIGURE 1 | Model and experiment. (A) Model. A relative decision value (RDV) evolves over time with a slope that depends on what the subject is looking at. In addition to the average drift, there is also Gaussian noise. When the RDV reaches one of the two barriers the subject makes the corresponding choice. The shaded regions indicate what the subject is currently looking at, blue for the product and yellow for the price. **(B)** Timeline. (Top) Subjects first reveal how much they are willing to pay for each of the 50 products, using a BDM auction. Then subjects make 300 purchasing choices (the 50 products at six different prices). At the end of the experiment one trial from the combined tasks is randomly chosen and the subject is paid and/or is shipped the chosen product. (Bottom) Within a choice trial, subjects must first fixate at the center of the screen for 2 s. They are then presented with an item and a price and given unlimited self-paced time to decide whether to buy the item at that price.

and can be easily seen from the two equations above. During fixations on the item, the RDV evolves with an average rate that underestimates the size of the price, which entails a temporary overestimation of the net value, and makes it more likely that the "purchase barrier" is reached. The opposite is true during price fixations.

MATERIALS AND METHODS

SUBJECTS

Thirty Caltech students participated in the experiment. Subjects received $50 for their participation, which they either kept or used to purchase items using the task described below. Caltech's Human Subjects Internal Review Board approved the experiment. Subjects provided informed consent prior to their participation.

TASK

At the beginning of the experiment subjects were endowed with $50 that they could use to purchase items in two subsequent tasks: a bidding task followed by a purchasing task. The subjects were told that at the end of the experiment one trial from the whole experiment would be randomly selected and implemented. Subjects kept whatever funds they did not spend.

Every subject performed two different tasks in the same order. First, they carried out a bidding task designed to measure the values (i.e., the v in the model) for each product. Second, they carried out a purchasing task that provides the data used to test the aDDM.

In the bidding task, subjects placed bids for the right to purchase 50 different consumer goods including mostly consumer electronics and household items. The task followed the rules of a Becker–Degroot–Marshak (BDM) auction, which is a tool widely used in economics to incentivize subjects to reveal their true values for products (Becker et al., 1964). In brief, a subject is asked to state their willingness-to-pay (WTP) for each of the products. If one of these trials is randomly selected to count, then the experimenter generates a random selling price (from a known uniform distribution from $0 to $50). If the random selling price is less than

or equal to the subject's stated WTP, then the subject purchases the product at the random price and keeps the rest of the $50. If the selling price is greater than the subject's WTP, then the subject does not purchase the product and keeps the entire $50. Note that the subject cannot influence the price that they pay for the product. They can only indicate whether they would be willing to buy at different prices. Thus, their unique best strategy is to state their true WTP. The order of item presentation was randomized for each subject. Each trial, subjects had unlimited time to examine a high-resolution photograph of the item along with a brief written description of the product on the computer screen and then submit a WTP from $0 to $50 by typing their bid.

The purchasing task is depicted in **Figure 1B**. Every trial subjects decided whether to buy the shown item at the stated price, or to keep their entire $50. On half of the trials the product was displayed on the top half of the screen with the price on the bottom half of the screen, while on the other half of the trials the locations were switched. Subjects had unlimited time to fixate back and forth between the product and price before indicating their choice to purchase (left arrow) or not purchase (right arrow). All 50 items were randomly presented six times, each time coupled with one of the following six prices: $3, $10, $18, $25, $33, $40. These prices were selected based on piloting data to span the mean WTPs for most items.

At the end of the experiment if one of these trials was randomly selected, then the subject was entitled to receive the product for that trial, and if purchased, we immediately ordered and mailed the product to the subject. The remaining money (up to $50) was immediately paid to the subject in cash.

EYE-TRACKING

Subjects' eye movements were recorded at 50 Hz using a Tobii desktop-mounted eye-tracker. Before each choice trial subjects were required to maintain a fixation at the center of the screen for 2 s before the stimuli would appear. This ensured that subjects began every trial fixating on the same location.

The eye-tracking data were preprocessed using the same procedure as in our previous work (Krajbich et al., 2010; Krajbich and Rangel, 2011). In particular, square regions of interest (ROI) were constructed around the product and price. A fixation was defined as the time from when at least one eye entered the ROI to the moment when both eyes left the ROI. For all measurements following the first fixation and preceding the last fixation of the trial, non-ROI fixations were dealt with according to the following rules:

(1) If the non-ROI fixations were recorded between fixations on the same ROI, then those non-ROI fixations were changed to that ROI. So for example a fixation pattern of "Product," "Blank," "Product" would become "Product," "Product," "Product." These non-ROI fixations are typically very short and are likely due to blinks.
(2) If the blank fixations were recorded between fixations on different ROIs, then those blank fixations were recorded as non-decision time and discarded from further analysis. Again, these non-ROI fixations are typically just one eye-tracker measurement (20 ms) and due to transitions between ROIs.

DATA CLEANING

Two types of trials were excluded from further analysis. First, for every subject, we excluded trials with items that received a bid of exactly $0 in the first task. These trials are problematic because there are an unusually large number of them (29%) compared to items with small but positive values ($1, $2, $3, etc.) which suggests that these items are being treated differently by the subjects. This could be the case, for example, if these are items that they already own and thus are not seriously considering purchasing. In this case such items could not be used to study the aDDM since they do not involve purchasing decisions. Nevertheless, including these trials does not qualitatively change any of the results.

Second, trials with a net value smaller than −$20, or larger than $20, were also excluded (an additional 28% of trials). These choices are extremely easy for the subjects and so not of particular interest to us. It is only in the −$20 to $20 range that choices are difficult and the aDDM makes interesting predictions. Beyond these net values, subjects are close to 100% accurate and the choice probabilities and reaction times asymptote. Because we were not interested in modeling these very easy choices, we designed the task to minimize trials in this range and so the data in these regions look noisy (due to missing observations) and uninteresting. Additionally, the bins outside this range are scarcely populated, which interferes with our ability to estimate the aDDM accurately.

In principle, this issue could have been avoided through experimental design by choosing prices in the purchasing task close to the values reported in the BDM task. Unfortunately, this is not a feasible solution, because it invalidates the incentive compatibility of the BDM procedure: subjects would have an incentive to bid low amounts in order to ensure low prices in the purchasing task.

MODEL SIMULATIONS AND FITTING

The model was fit to the choice and reaction time data using only the even numbered trials of the pooled data set from all the subjects. We then tested the quality of the model fit and predictions by simulating the model using the fitted parameters, and comparing them with the actual data from the odd trials. Thus, the predictions of the model are tested quantitatively and out-of-sample.

The model was fit using a maximum likelihood estimation (MLE) procedure implemented in several steps. First, we ran 3000 simulations for each combination of the model parameters in the grid described below, the six different prices, and the product values (sampled in $2 increments from $0 to $50). In the simulations we randomly sampled fixation times from the empirical distribution conditional on their fixation type (product or price). First-fixation price fixations were sampled separately from the rest of the price fixations because they are statistically shorter than non-first price fixations. Item fixations were sampled conditional on their fixation number (1, 2, 3, 4, 5, 6, 7, or later), and binned on the absolute net value ($|v - p|$, in $4 increments). We also used the empirical fact that subjects looked at the item first 53.3% of the time. We also took account of the latency for the first fixation, as well as for the transition time between fixations. We did this by computing the difference between the average reaction time and the average total fixation time (time spent looking at either the product or the price), and adding this trial-independent estimate to the simulated reaction times. This average "non-decision" time

was 336 ms, consistent with non-decision times in our previous work (355 ms in Krajbich et al., 2010) and in other DDM studies of binary choice (365 ms in Milosavljevic et al., 2010 and 338 ms in Ratcliff, 2002, for example).

Second, we computed the probability of each observation in the data, for each set of parameters, as follows. The empirical spread of reaction times ranged from 495 ms to 47.1 s so in the fitting procedure we discarded any simulation trials below 400 ms or above 48 s. The rest of the reaction times were separated into bins of 400 ms except for the final bin, which went from 4800 ms to 48 s. For each net value bin ($4 increments), we split the data into buy and no-buy trials, then for each group counted the number of data trials in each reaction time bin, and similarly calculated the probability of a simulation trial occurring in each reaction time bin. Note that the first two steps are necessary to compute the likelihood function of the data because there is no known closed-form solution for the aDDM.

Third, we identified the set of parameters that maximized the log-likelihood of the data by taking the logarithms of each of these probabilities and summing them up. The resulting number is used to assess how well the model fit the data, with larger numbers (closer to zero) indicating better fits. The search was performed over a grid of values for d, σ, and θ. The MLE search was carried out in two steps. First we searched over a coarse grid with d in (0.00006, 0.000065, 0.00007), σ in (0.0228, 0.0237, 0.0247), and θ in (0.7, 0.8, 0.9, 0.99). Second, we used the results from the first search to define a narrower search with d in (0.00006, 0.000065), σ in (0.0218, 0.0228, 0.0237), and θ in (0.5, 0.6, 0.7, 0.8). The initial parameters were chosen based on comparisons to our previous results (Krajbich et al., 2010). The best-fitting parameters at this stage were $d = 0.000065$, $\sigma = 0.02275$, and $\theta = 0.7$.

The previous procedure assumes that there is no motor latency, namely that the subject enters his choice instantaneously once the RDV crosses a barrier. To address this problem, we carried out the following additional step in the model fitting process. We simulated the model with the best MLE parameters described above, and then compared the average duration of the final fixations with those from the actual data. Note that this comparison is important because, to the extent that subjects maintain fixation during the motor latency period, our measured durations for final fixations include the actual duration of the fixation and the motor latency. Consistent with this concern, our predicted final fixation durations were on average 73 ms shorter than actual measured fixations. Taking this value as the mean motor latency, we redid the model fitting procedure with the added feature that there is a 73 ms delay after the decision is made. As before, we began with a coarse grid search, this time with d in (0.00006, 0.000065, 0.00007), σ in (0.0228, 0.0237, 0.0247), and θ in (0.6, 0.7, 0.8, 0.9). Second, we used the results from the first search to define a narrower search with d in (0.000065), σ in (0.0218, 0.0228, 0.0233, 0.0237), and θ in (0.6, 0.65, 0.7, 0.75, 0.8, 0.9). The reported model fits below include the motor latency correction, which is also used in all simulations reported in the paper.

GOODNESS-OF-FIT CALCULATIONS

For **Figures 2B,C**, and **5** we could not compute χ^2 goodness-of-fit statistics because the dependent variables are not binary. Instead we devised the following alternative goodness-of-fit statistic: (1) for each bin of the independent variable we "correct" the dependent variable by subtracting the average simulated value from each subject's average value. (2) We then run a weighted least-squares (WLS) regression with the "corrected" dependent variable. The weights in the regression are equal to the inverse of the variance. Note that if the simulations fit the data well, then the "corrected" data should be a flat line at 0. On the other hand, if the simulation fits poorly, then the WLS coefficients should be non-zero. So, for goodness-of-fits, we report the p-values for the coefficients of these WLS regressions. If the p-values are less than 0.05 then we reject that the model accurately fits the data.

RESULTS

MODEL FIT

We fitted the model to the even numbered trials from the group data using MLE. The model has three free parameters: the constant determining the speed of integration d, the discount parameter θ, and the noise parameter σ. The model was fit under the assumption that time evolves in 1 ms discrete steps. We selected the parameters that maximized the probability of the observed choices and reaction times, conditional on the net value of the offer (see Materials and Methods for details). The best-fitting model had parameters $d = 0.000065\ \$^{-1}\text{ms}^{-1}$, $\sigma = 0.0233$, and $\theta = 0.7$.

For comparison purposes, in our previous work on binary and trinary choice we found that the parameters $d = 0.0002\ \text{ms}^{-1}$, $\sigma = 0.02$, and $\theta = 0.3$ provided the best fit for the data (Krajbich et al., 2010; Krajbich and Rangel, 2011). Since these parameters were fitted using methods very similar to those used here, their comparison provides some insight about the computational differences between binary choices and yes-no purchasing decisions. Consider some of the most salient relationships. First, we found that the noise parameter σ was nearly identical in the two cases, which suggests a similar amount of computational noise in both problems. Second, we found that the slope parameter d was about 1/3 smaller in the current data set. This difference needs to be interpreted with caution, however, because the slope depends on the scale with which values are measured. In particular, multiplying the net values by a constant a results in a new fit that decreases d by a factor of $1/a$, and leaves the fits of the other two parameters unchanged. The difference is then easily explained by the fact that values in our previous paper were measured using liking-rating differences that ranged from -5 to $+5$, whereas here the value differences ranged from -20 to $+20$. Third, the bias is significantly smaller than in our previous work: $\theta = 0.7$ vs. $\theta = 0.3$ previously (recall that $\theta = 1$ is the case of no bias, so that as theta increases from 0 to 1 the size of the bias goes down). Nevertheless, as we will see in more detail below, with $\theta = 0.7$ there is still a non-trivial effect of visual fixations on the choice process.

MODEL SIMULATION

In order to investigate the ability of the model to predict the data quantitatively, we then simulated the model 10,000 times per value and price combination (binning values every $2), using the estimated maximum likelihood parameters, and by sampling fixation lengths from the actual empirical fixation data (see Materials and Methods for details). Throughout, we assume that fixations always

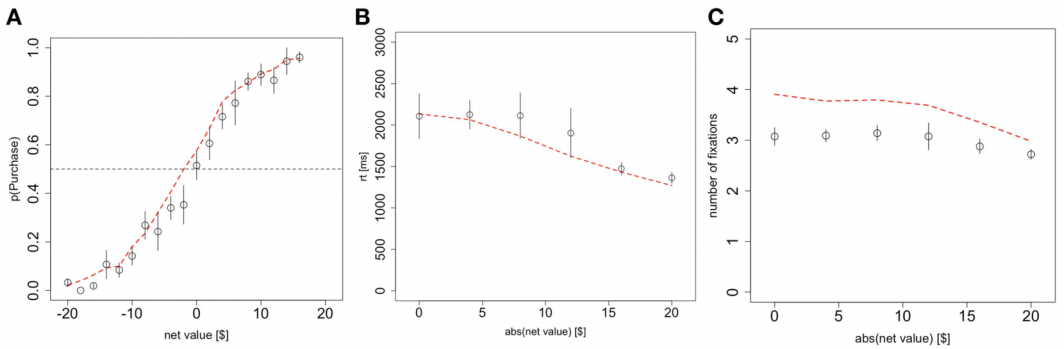


FIGURE 2 | Basic psychometrics. (A) Probability of purchasing the product as a function of the net value (product value – price). **(B)** Reaction times as a function of the magnitude of the net value. **(C)** The number of fixations in a trial, as a function of the magnitude of the net value. Black circles indicate data from the odd-numbered trials of the subject data, and red dashed lines indicate the simulated data from the aDDM. Bars are standard error bars, clustered by subject.

alternate between the product and price, and that the location of the first fixation is chosen probabilistically to match the empirical data (look at the product first with probability 53.3%). The results of the simulations are described below.

Note that all comparisons of the model to the data were made out-of-sample using only on the odd-numbered trials, since the model was fitted to the even-numbered trials. Also, in all of the figures below red curves represent the simulations, and black symbols and curves represent the data.

Unless otherwise noted, throughout the results, goodness-of-fit p-values are based on two-sided t-tests of the regression parameters against zero (see Materials and Methods for goodness-of-fit details), and p-values for trends in the subject data are based on two-sided t-tests of the mixed-effects regression parameters against zero. Each mixed-effects regression contains random effects for the intercept and all other regressors.

BASIC PSYCHOMETRICS

In this section we investigate how well the aDDM predicted the choice and reaction time data on the half of the data that it was not fitted to, and additionally estimate the effects of choice difficulty on average reaction times and number of fixations. Our measure of choice difficulty is the absolute (unsigned) net value, which tells us the difference in value between the item and the price. With this measure, the hardest decision is one where the item value and price are identical, leading to a net value of $0.

The fitted model accounted for the choice and reaction time curves well. Choices were a logistic function of net value in both the data and the simulations (**Figure 2A**; χ^2 goodness-of-fit: $p = 0.8$). As expected, subjects were more likely to purchase on trials with a positive net value and less likely to purchase on trials with a negative net value.

Reaction times significantly increased with difficulty at an average rate of 42 ms/$ ($p < 10^{-10}$ mixed-effects) and the aDDM also provided a close fit to the data (**Figure 2B**; goodness-of-fit slope: $p = 0.9$, intercept: $p = 0.7$).

The average number of fixations in a trial exhibits a small but significant ($p < 10^{-5}$ mixed-effects) response to increasing difficulty at a rate of 0.016 fixations/$, though the aDDM does not fit as well here (**Figure 2C**; goodness-of-fit slope: $p = 0.0005$, intercept: $p = 10^{-13}$). Even after correcting for the difference in scale, the size of effect is about 1/3 of the one that we previously observed with choices between food items (Krajbich et al., 2010).

As shown in **Figure 2C**, the model systematically predicts 0.7 excess fixations. This mismatch has been consistently observed with the aDDM (Krajbich et al., 2010; Krajbich and Rangel, 2011), is an unavoidable consequence of the procedures used to carry out the simulations, and does not reflect an inherent limitation of the model. In fact, this bias is present even if one simulates a dataset using the aDDM, and then carries out the model fitting exercise on the simulated data. More concretely, the problem is due to the fact that we have to sample fixations from the empirical distribution of non-final fixations, but many of those fixations are cut short by a barrier crossing and become final fixations in the simulations. The longer the fixation, the more likely it is to cross a barrier, and so the average middle fixation duration is shorter in the simulations; this means that more fixations are required to achieve the same reaction times in our simulations as in the actual data. Nevertheless, the mismatch is larger here than previously observed and the aDDM also over-estimates the effect of difficulty on the number of fixations, though the difference is quite small (0.03 fixations/$).

MODEL PREDICTIONS AND CHOICE BIASES

The model with $\theta = 0.7$ makes several stark predictions about the relationship between fixations, choices, and reaction times that we test using the eye-tracking data.

First, the model predicts that the last fixation of the trial is more likely to be to the product when the subject decides to purchase it, unless the BDM value of the item is sufficiently smaller than the price. It also predicts that the last fixation is likely to be to the price when the subject decides not to purchase the item, unless the price is sufficiently lower than the value of the product. To see the intuition for this effect, consider the case in which the last fixation is to the product. In this case, the RDV tends to climb toward the purchase barrier, unless the value of the item is smaller than $\theta * p$. **Figure 3A** looks at the probability that the last fixation is to the chosen item (product = purchase or price = no purchase) as a

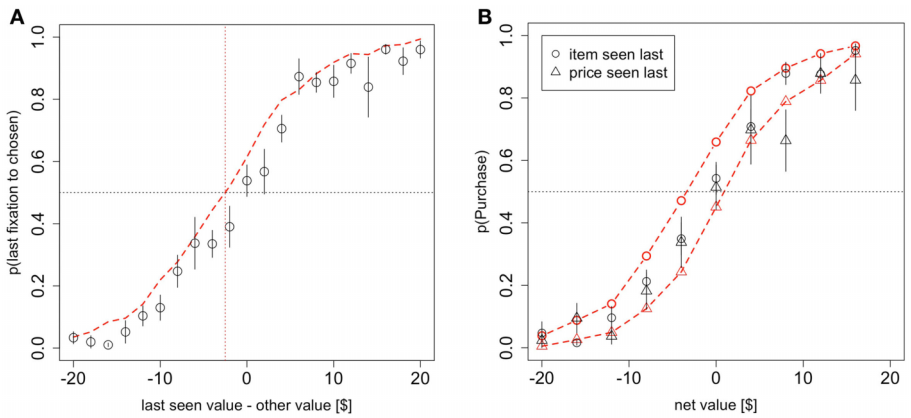


FIGURE 3 | Model predictions and results. (A) The probability that the last fixation of the trial is to the chosen stimulus (product or money) as a function of the difference in value between the last-seen stimulus and the other stimulus. **(B)** The probability of purchasing the item as a function of net value, contingent on whether the last fixation was to the product or to the price. Black circles indicate data from the odd-numbered trials of the subject data, and red dashed lines indicate the simulated data from the aDDM. Bars are standard error bars, clustered by subject.

function of the difference in value between the last-seen stimulus and the other stimulus. The figure shows that a small but noticeable bias of this type is present in both the data and the simulations (χ^2 goodness-of-fit: $p = 0.4$).

For an additional test of the prediction that there is an overall bias toward choosing the item that is looked at last, we ran a logit regression using the entire dataset and variables in **Figure 3A** to test whether the intercept was greater than 0. The logit confirmed that indeed the intercept was greater than 0 ($p = 0.03$ based on a one-sided t-test on the entire dataset), indicating that for a net value of \$0, the probability of choosing the item that is looked at last is significantly greater than 0.5. We also ran an additional version of this analysis with a different logit regression for each subject and found that the average intercept was marginally significantly greater than zero ($p = 0.07$ one-sided t-test), providing further evidence for this effect.

Second, the model predicts that the choice curves should be different for trials where the last fixation was to the product or to the price. Specifically, it predicts that subjects should be more likely to purchase the item if the last fixation was to the product than if it was to the price. This follows directly from the fact that with $\theta < 1$, the net value of the trade is overestimated when fixating on the product, and underestimated when fixating on the price. This effect is small but noticeable in both the simulations and the data (**Figure 3B**; χ^2 goodness-of-fit item seen last: $p = 0.01$, price seen last: $p = 0.3$). To test the significance of this effect in the data, we ran two logistic regressions on the entire dataset and variables from **Figure 3B**, one for trials where the product was seen last and one for trials where the price was seen last. For trials where the product was seen last the logit intercept was significantly greater than zero ($p = 0.03$ one-sided t-test) while for trials where the price was seen last the logit intercept was significantly less than zero ($p = 0.05$ one-sided t-test). Subject-level analyses were inconclusive ($p = 0.53$), likely due to many cases of perfect separation and small numbers of observations for other subjects in some of the bins.

Third, the model predicts that more time spent looking at the product over the course of the trial will bias subjects toward purchasing the good, and that is what we see in the data (**Figure 4A**; χ^2 goodness-of-fit: $p = 0.4$). Here the total fixation time advantage for the item is just the total amount of time spent looking at the item minus the total amount of time spent looking at the price, in that trial. One important concern with this result has to do with the exogeneity of the fixation lengths. In particular, it could be that the choice bias shown in **Figure 4A** is due to a positive relationship between the total fixation time advantage for the item and the net value. To address this concern we ran a mixed-effects logit regression including net value and total fixation time advantage for the item as independent variables and purchasing as the dependent variable. The effect of the total fixation time advantage was highly significant ($p < 10^{-6}$), ruling out this alternative explanation. Furthermore, a regression of the total fixation time advantage for the item on net value shows no significant relationship (5.2 ms/\$, $p = 0.09$ mixed-effects linear regression, two-sided t-test).

An alternative way to investigate this effect is to split trials based on whether the subject spent more time looking at the product or the price. As expected, controlling for net value, subjects were more likely to purchase the product if they spent more time looking at it than if they spent more time looking at the price (**Figure 4B**). To test for a significant difference between these curves, we ran two logistic regressions (analogous to the analysis for **Figure 3B**) on the entire dataset and variables in **Figure 4B**, one for trials where the product was looked at more and one for trials where the price was looked at more. For trials where the subjects looked longer at the product, the logit intercept was significantly greater than zero ($p < 0.01$ one-sided t-test) while for trials where the subjects looked longer at the price, the logit intercept was significantly less than zero ($p = 0.02$ one-sided t-test). As a further test, we ran identical subject-level logits and again found a significant difference between the intercepts for trials where the product was looked at more and the intercepts for trials where the price was

looked at more ($p = 0.03$, one-sided paired t-test; 10 subjects were excluded due to perfect separation).

Finally, the model also predicts a precise quantitative relationship between reaction times and net values, as a function of the type of decision made. As shown in **Figure 5**, the model provides a fairly good description of the associated patterns (**Figure 5A** goodness-of-fit slope: $p = 0.5$, intercept: $p = 0.8$; **Figure 5B** goodness-of-fit slope: $p = 0.11$, intercept: $p = 0.12$). The intuition behind these patterns goes as follows. In this experiment, we can define a mistake as either a trial where the subject purchased an item with a negative net value or a trial where the subject didn't purchase an item with a positive net value. With $\theta < 1$, mistakes tend to occur when the subject has spent more time looking at the worse option. Therefore, mistakes take longer than correct choices because the average drift rate when the subject is looking at the worse item is always smaller than when the subject is looking at the better item, and a smaller drift rate requires more time for the RDV to reach the choice barrier. Furthermore, the bigger the mistake, the longer the decision should take, up to a point. There is a counteracting force, which is that big mistakes *shouldn't* happen and so if the subject takes some time to make his decision then he won't make these big mistakes. In other words, given a reasonable amount of time, the overwhelming average drift rate in favor of the correct option should overwhelm the noise in the decision process. Therefore, big mistakes must occur due to large spikes in the noise, and these spikes must occur early in the trial before there is overwhelming evidence for the correct choice. Without a formal model it would be unclear at what point these two counteracting forces should shift in power, but our model predicts the trend in the data quite accurately.

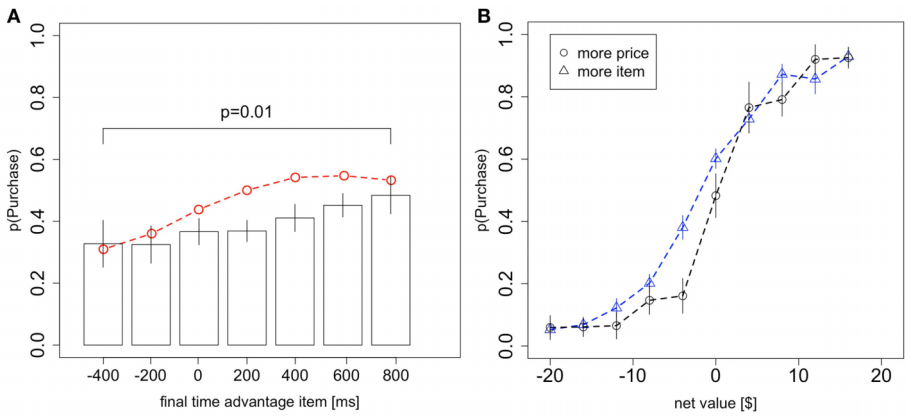


FIGURE 4 | Choice biases. (A) The probability of purchasing the item as a function of the difference in total fixation time (over the whole trial) between the product and the price. Black circles indicate data from the odd-numbered trials of the subject data, and the red dashed line indicates the simulated data from the aDDM. The p-value is from a one-sided t-test. **(B)** The probability that the product is chosen as a function of the net value, conditional on whether more time was spent looking at the product or the price in that trial. Bars are standard error bars, clustered by subject.

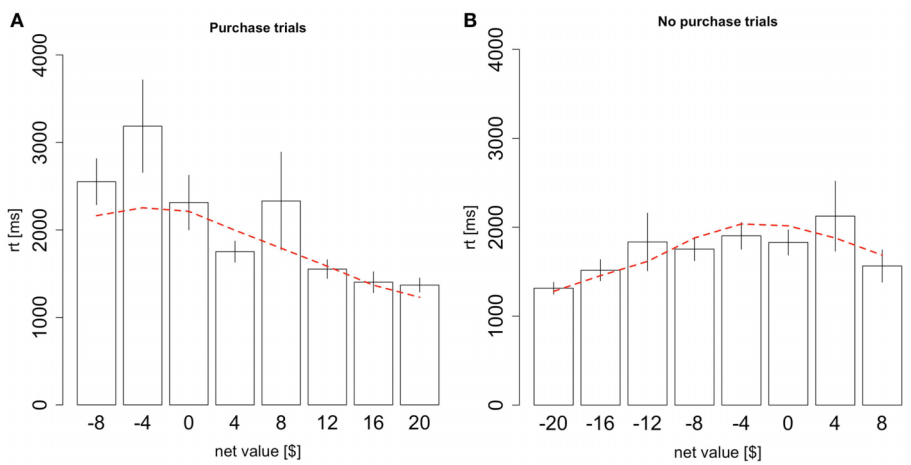


FIGURE 5 | Reaction times conditional on choice. (A) Reaction times as a function of the net value, conditional on purchasing the product. **(B)** Reaction times as a function of the net value, conditional on not purchasing the product. Black bars indicate data from the odd-numbered trials of the subject data, and the red dashed lines indicate the simulated data from the aDDM. Bars are standard error bars, clustered by subject.

FIXATION PROPERTIES

Finally, we investigate the extent to which the fixation process resembles the assumptions and predictions of the model. To do so we look at the first fixations of each trial, as well as the "middle fixations" which are any fixations that are not the first fixation or the last fixation of the trial. We treat the last fixations of the trials separately since in the model these fixations are cut short by the choice.

Consistent with previous findings on choice between products, we found no significant correlation between product value and mean fixation duration for either first fixations or middle fixations. There was a small effect of price and a large effect of choice difficulty on middle item mean fixation duration (**Figures 6A–C**, price: 10 ms/\$, $p = 0.02$, product value: 6.7 ms/\$, $p = 0.3$, net value: -35 ms/\$, $p < 10^{-6}$, two-sided t-tests based on a mixed-effects regression with all three factors). However, we found no such effect for middle price fixations (**Figures 6D–F**, price: 0.85 ms/\$, $p = 0.08$; product value: -0.018 ms/\$, $p = 0.98$; net value: -1.38 ms/\$, $p = 0.08$, two-sided t-tests based on a mixed-effects regression with all three factors), or for first fixations to either products (price: 0.089 ms/\$, $p = 0.9$; product value: 0.37 ms/\$, $p = 0.5$; net value: -0.9 ms/\$, $p = 0.2$, two-sided t-tests based on a mixed-effects regression with all three factors) or prices (price: 0.11 ms/\$, $p = 0.8$; product value: 0.71 ms/\$; $p = 0.2$, net value: -0.51 ms/\$, $p = 0.5$, two-sided t-tests based on a mixed-effects regression with all three factors).

Consistent with the predictions of the model, and with previous findings (Krajbich et al., 2010; Krajbich and Rangel, 2011), we also found that first and last fixations were significantly shorter than middle fixations both for products (first: $p = 10^{-7}$, last: $p = 0.0001$, two-sided paired t-tests) and prices (first: $p = 10^{-5}$, last: $p = 0.03$, two-sided paired t-tests; **Figure 7A**). Our model does not predict the first fixation effect, but we have seen it consistently in all of our previous related studies. The model does predict shorter final fixations, since final fixations are just middle fixations cut short by the RDV crossing a decision barrier.

Figure 7B compares the distribution of price and product total fixation times, i.e., summed over the whole trial. The distribution of total fixation times for items has a larger mean and standard deviation than that for prices (items: $M = 1064$ ms, SD $= 1266$ ms; prices: $M = 471$ ms, SD $= 390$ ms; M: $p = 10^{-8}$, SD: $p = 10^{-7}$; statistics computed first at the individual level, and then compared across individuals using paired t-tests).

Finally, **Figure 7C** shows the average number of item fixations and price fixations per trial. We see that both types of fixations follow the same trend as seen in the aggregate **Figure 2C**, but there are consistently about 0.5 fewer price fixations than item fixations. However, this result is not surprising, given that item fixations are longer than price fixations and so are more likely to be the last fixation of the trial. For trials with an even number of fixations there is always an equal number of fixations to item and price, regardless of the last fixation location. But for trials with an odd number of fixations there will always be one more fixation for the last-seen stimulus. Since the last fixation of the trial is to the item in 79% of trials, we indeed expect there to be ∼0.4 fewer price fixations than item fixations.

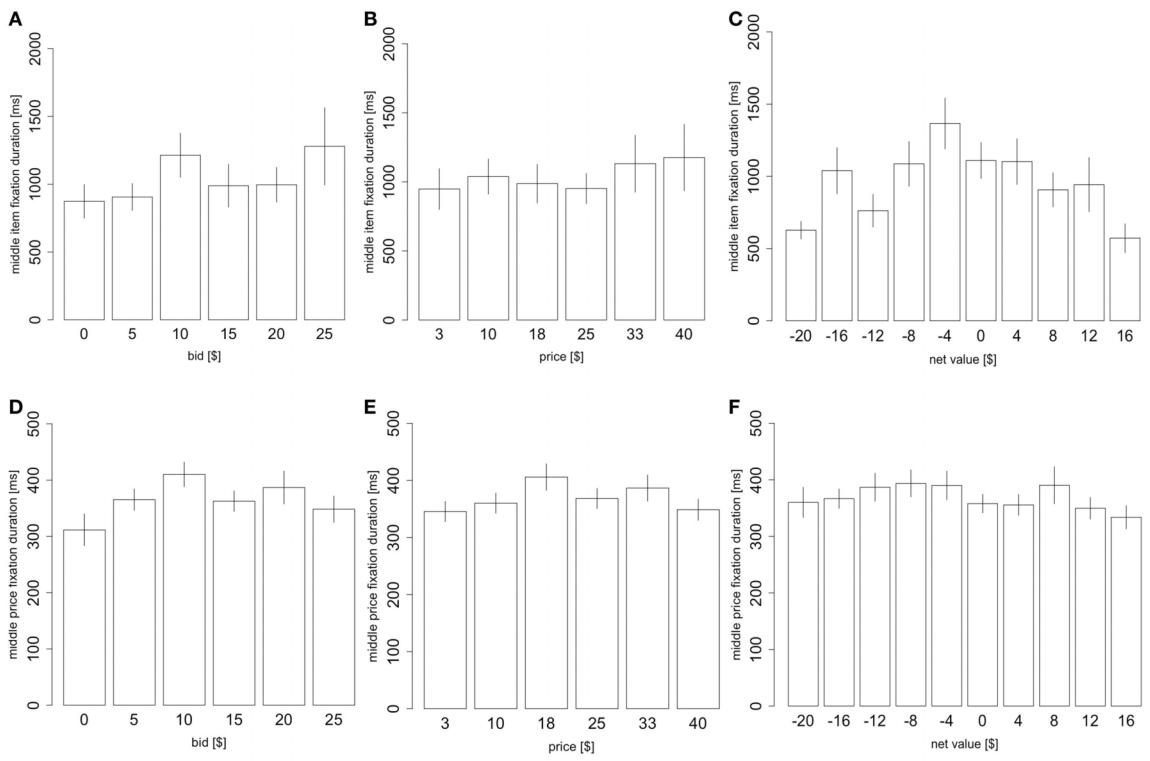


FIGURE 6 | Fixation properties. (A) The duration of middle item fixations as a function of the item value, **(B)** price, and **(C)** net value. **(D)** The duration of middle price fixations as a function of the product value, **(E)** price, and **(F)** net value. Bars are standard error bars, clustered by subject.

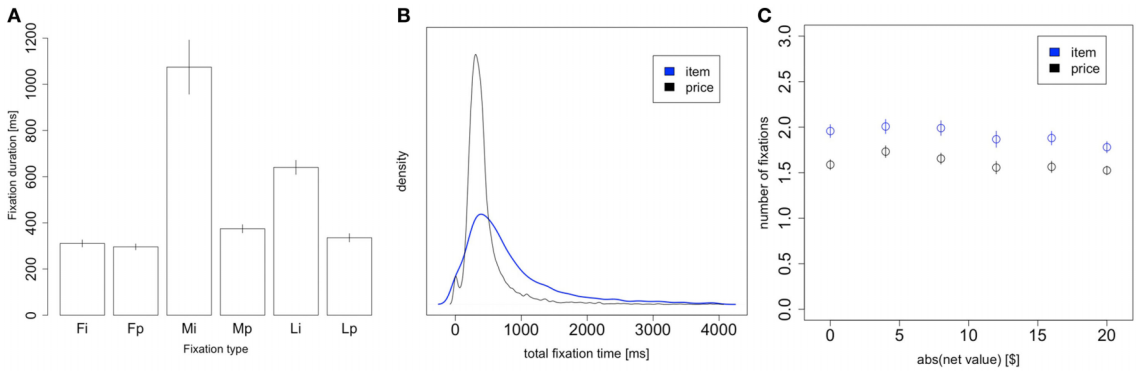


FIGURE 7 | Item vs. price fixations. (A) Fixation duration as a function of fixation type: first fixation to item, first fixation to price, middle fixation to item, middle fixation to price, last fixation to item, and last fixation to price. Note that "first/last fixation" means the first/last fixation of the trial, not the first/last fixation to each stimulus. **(B)** Density plot of the total trial time spent looking at the item (blue) and price (black). **(C)** The number of item and price fixations in a trial, as a function of the magnitude of the net value. Bars are standard error bars, clustered by subject.

DISCUSSION

We have described the results of an eye-tracking experiment of purchasing decisions designed to investigate if the aDDM is able to provide a reasonable quantitative description of the relationship between the fixation, choice, and reaction time data in the case of simple purchasing decisions. The motivation for doing this is that in previous work we have found that the model provides a remarkably accurate description of these variables and their interrelationship in the case of binary and trinary food choices (Krajbich et al., 2010; Krajbich and Rangel, 2011). Thus, the research agenda here is to investigate the extent to which the aDDM also applies to other types of decisions, what changes are needed to account for new aspects of the task, and more generally when it breaks down and why.

We find that the model provides a reasonably accurate description of the purchasing decision data, although of significantly lower accuracy than in our previous work. This shows that purchasing decisions introduce new aspects into the problem that are not well captured by the simple aDDM and that need to be investigated in future research (see below for some conjectures). However, we find that the best-fitting aDDM has parameters similar to what we found in our previous work, with the exception of the parameter controlling the magnitude of the visual fixation bias in the value integration process. In particular, the bias parameter went from $\theta = 0.3$ in our previous work to $\theta = 0.7$ in the current dataset, which constitutes a sizable reduction in the size of the fixation-driven choice biases. Nevertheless, many of the key biases predicted by the model were still present in the purchasing data, albeit of a smaller magnitude than those found in the previous studies and generally not as big as the model predicted. Furthermore, once again we found that these effects are not due to subjects looking longer at products (or less at prices) in trials with a high net value.

In judging the reduced accuracy of the model it is important to keep in mind that we are imposing some very strict tests on the model. First, we are fitting the model on one half of the data and then predicting on the other half of the data, rather than merely showing fits to the data. Second, we are fitting the model on only the choice and reaction time curves (**Figures 2A,B**). It is quite likely that we could have achieved nicer results by fitting to the fixation trends as well (e.g., **Figure 4A**), but that would detract from a main feature of the model, which is the ability to predict fixation trends using only choice and reaction time data.

These results are important for several reasons.

First, they provide additional evidence that the aDDM provides a reasonably accurate and robust characterization of how the brain computes value-based choices of different types. An important difference with our previous paradigms is that here subjects had to integrate the value of two very different types of stimuli.

Second, the results provide some new insights to the literature on decision field theory (DFT) by Busemeyer and others (Busemeyer and Townsend, 1993; Diederich, 1997, 2003; Roe et al., 2001; Busemeyer and Diederich, 2002; Johnson and Busemeyer, 2005; Tsetsos et al., 2010). See also the closely related models of Usher and McClelland (2004) and Usher et al. (2008). These models have also investigated the impact that random fluctuations in attention have on choice accuracy and reaction time. In particular, although the DFT model has not been previously applied to the type of simple purchasing decisions studied here, it is easily extended to this case. Such an extension highlights an important difference between the two models: DFT assumes that attention (and thus the drift rate) fluctuates continuously across time according to either a stationary or Markov process, while we assume that there are only two states of attention and that the current state is indicated by the subject's fixation location. This is an important difference because, as a result, although it can account for the basic choice and reaction time profiles, DFT cannot account for many of the fixation patterns described here and in our previous work on multi-option choice (Krajbich et al., 2010; Krajbich and Rangel, 2011).

Our model is closer in spirit to multiattribute decision field theory (MDFT; Diederich, 1997, 2003; Busemeyer and Diederich, 2002) but in that model it is assumed that the attributes are processed in a serial manner, equivalent to the case of $\theta = 0$ in our model. Again, that model cannot account for the trends in our data. For purchasing decisions, our aDDM represents something like a hybrid between DFT and MDFT, with the additional

specification that gaze location determines the weights on the attributes. It is also worth noting that analytic solutions have been derived for those models, under certain assumptions about the evolution of the attribute weights.

However, we emphasize that our purpose here is not to rule out these alternative models and explanations, but to argue that the aDDM provides a simpler explanation of value-based decision-making that can quantitatively account for the effects of visual attention on choice in both purchasing and multi-option decisions. Obviously this simplicity comes at a cost: we do not account for possible shifts of attention to positive or negative attributes *within* a fixation. Instead we assume that evidence for an option is accumulated with an average rate that is proportional to the latent value (i.e., the BDM value), which should reflect the mixture of positive and negative attributes. Shifts of attention to different attributes *within* a fixation are therefore only taken into account by the Gaussian random noise that is added to the average evidence accumulation.

Second, the results show that the basic mechanisms at work in the aDDM also seem to apply to cases in which one of the options is numeric or symbolic, instead of a more complex visual stimulus. This is important because it provides some new hints about the nature of the processes at work. It has been previously speculated that the process of value integration and comparison is noisy because the brain needs to take repeated noisy samples of the value of the stimuli being evaluated, and that visual attention matters because it guides the sampling process (Busemeyer and Townsend, 1993; Krajbich et al., 2010; Glockner and Herbold, 2011; Krajbich and Rangel, 2011). This is a natural interpretation for stimuli that are visually complex, but it is not obvious if the same holds true for numerical price representations. The results here show that this is indeed the case: although fixations to prices are shorter and less variable, the results suggest that they are also integrated noisily over time. This suggests that the process of noisy dynamic integration of value might be a widespread aspect of the choice process, and not just applicable to complex visual stimuli.

To lend further support to this claim, we carried out model fittings of an additional model that allows for different values of the visual fixation bias parameter θ, depending on whether the product or the price is being fixated on. We found that adding this additional parameter did indeed significantly improve the fit of the model, but that the best fitted θ for the item was still 0.7, and the best fitted theta for the price θ changed only slightly to 0.6. This provides further support for the finding that the integration processes for the product and price information have similar properties.

Third, the difference between the magnitude of the visual biases raises the following interesting puzzle: why is it the case then that the visual bias for the products during the purchasing decisions is much lower than the bias for the foods during binary choice? After all, the complexity of the pictures used to display the products here is very similar to that of the food pictures used in our previous work. One potential but speculative explanation is that there is a visual difference in the two displays. During binary choice two similar food pictures are displayed, while two different types of visual images – a visually complex food picture and a simple number – are displayed during purchasing decisions. Looking at one picture might inhibit working memory for the other picture (including low-level rehearsal captured by θ) more strongly than looking at a number, and vice-versa (Baddeley, 2003). The reduced inhibition in the picture-number paper would be manifested by a higher value of θ, just as we infer from the data. If this hypothesis is correct, then when prices are presented in a complex visual display (e.g., stacks of money and coins) rather than as a simple number, we might expect a θ lower than what was measured in this task.

Fourth, the fact that price fixations were shorter and less variable than those for products is consistent with the idea that the deployment of visual attention is based on the utility of information (Gottlieb and Balan, 2010). An important open question for future work is to develop and test a full optimal model of visual attention deployment. Note that this imposes more stringent restrictions on the fixation process than just having shorter fixations on stimuli that are easier to process, or having less noise associated with them. One place where the aDDM was noticeably inaccurate was in the prediction of how the number of fixations varies with choice difficulty, and so clearly we need a better understanding of how fixations work in purchasing decisions.

Fifth, the results have obvious implications for marketing and public policy. In particular, they show that sellers may be able to use strategic product and price placement, as well as manipulation of saliency (Milosavljevic et al., 2012), to encourage consumers to purchase their products. Also, sellers of inferior products may benefit from deliberately putting people under unnaturally high time pressure (Diederich, 1997). In the aDDM, time pressure creates a large rate of "mistakes" (choices with a negative net value) that decreases in longer trials. Indeed, some high-pressure phone and door-to-door sales tactics could be construed as attempts to bring a premature halt to a slow DDM process, in order to create large consumer mistakes that benefit the sellers. The aDDM approach gives a new way to study this process scientifically and can suggest policy remedies (e.g., "cooling off periods," during which buyers can costlessly renege on an agreement to sell; Camerer et al., 2003).

Of course such manipulations have their limits. Compared to previous results with multi-option choice, the effects of visual attention are quite reduced in our purchasing task. Therefore, if the product under consideration is substantially worse than its price then the drift rate will always tend toward the "do not buy" threshold, regardless of what the subject is looking at. In addition, noticeable attention manipulations such as forcing the subject to look at a product for a long time may alert the subject and alter the decision-making process. Nevertheless, previous research has shown that it is indeed possible to influence choices by exogenously manipulating fixation times (Shimojo et al., 2003; Armel et al., 2008), consistent with the idea that visual fixations are influencing the choices and not vice-versa.

Finally, the study provides additional evidence of the value of utilizing eye-tracking data in conjunction with carefully designed

decision tasks to test process models of decision-making. In this sense, it builds on the seminal pioneering work of Johnson and colleagues (Johnson et al., 1988, 2002, 2007; Camerer and Johnson, 2004), as well as more recent applications (Russo et al., 2006, 2008; Raab and Johnson, 2007; Horstmann et al., 2009; Glockner and Herbold, 2011; Russo and Yong, 2011; Glockner et al., 2012). For an outstanding recent review, see Russo (2010).

ACKNOWLEDGMENTS

This research was supported by grants to Antonio Rangel from the NSF (SES-0851408, SES-0926544, SES-0850840), NIH (R01 AA018736), and Lipper Foundation. Ian Krajbich was supported by a National Science Foundation Integrative Graduate Education and Research Traineeship. Dingchao Lu was supported by a Summer Undergraduate Research Fellowship from the California Institute of Technology.

REFERENCES

Armel, K. C., Beaumel, A., and Rangel, A. (2008). Biasing simple choices by manipulating relative visual attention. *Judgm. Decis. Mak.* 3, 396–403.

Baddeley, A. (2003). Working memory: looking back and looking forward. *Nat. Rev. Neurosci.* 4, 829–839.

Basten, U., Biele, G., Heekeren, H., and Fieback, C. J. (2010). How the brain integrates costs and benefits during decision making. *Proc. Natl. Acad. Sci. U.S.A.* 107, 21767–21772.

Becker, G. M., Degroot, M. H., and Marschak, J. (1964). Measuring utility by a single-response sequential method. *Behav. Sci.* 9, 226–232.

Bogacz, R. (2007). Optimal decision-making theories: linking neurobiology with behaviour. *Trends Cogn. Sci. (Regul. Ed.)* 11, 118–125.

Busemeyer, J. R., and Diederich, A. (2002). Survey of decision field theory. *Math. Soc. Sci.* 43, 345–370.

Busemeyer, J. R., and Johnson, J. G. (2004). "Computational models of decision making," in *Blackwell Handbook of Judgment and Decision Making*, eds D. J. Koehler and N. Harvey (Malden, MA: Blackwell Publishing Ltd), 133–154.

Busemeyer, J. R., and Rapoport, A. (1988). Psychological models of deferred decision making. *J. Math. Psychol.* 32, 91–134.

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.

Camerer, C., Issacharoff, S., Loewenstein, G., O'Donoghue, T., and Rabin, M. (2003). Regulation for conservatives: behavioral economics and the case for "asymmetric paternalism." *Univ. PA. Law Rev.* 151, 1211–1254.

Camerer, C., and Johnson, E. J. (2004). "Thinking about attention in games: backward and forward induction," in *The Psychology of Economic Decisions*, eds J. Carillo and I. Brocas (New York: Oxford University Press), 111–129.

Diederich, A. (1997). Dynamic stochastic models for decision making under time constraints. *J. Math. Psychol.* 41, 260–274.

Diederich, A. (2003). MDFT account of decision making under time pressure. *Psychon. Bull. Rev.* 10, 157–166.

Glockner, A., Heinen, T., Johnson, J. G., and Raab, M. (2012). Network approaches for expert decisions in sports. *Hum. Mov. Sci.* 31, 318–333.

Glockner, A., and Herbold, A.-K. (2011). An eye-tracking study on information processing in risky decisions: evidence for compensatory strategies based on automatic processes. *J. Behav. Decis. Mak.* 24, 71–98.

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

Gottlieb, J., and Balan, P. (2010). Attention as a decision in information space. *Trends Cogn. Sci. (Regul. Ed.)* 14, 240–248.

Hare, T., and Rangel, A. (2010). Neural computations associated with goal-directed choice. *Curr. Opin. Neurobiol.* 20, 262–270.

Hare, T., Schultz, W., Camerer, C., O'Doherty, J. P., and Rangel, A. (2011). Transformation of stimulus value signals into motor commands during simple choice. *Proc. Natl. Acad. Sci. U.S.A.* 108, 18120–18125.

Horstmann, N., Ahlgrimm, A., and Glockner, A. (2009). How distinct are intuition and deliberation? An eye-tracking analysis of instruction-induced decision modes. *Judgm. Decis. Mak.* 4, 335–354.

Johnson, E. J., Camerer, C., Sen, S., and Rymon, T. (2002). Detecting failures of backward induction: monitoring information search in sequential bargaining. *J. Econ. Theory* 104, 16–47.

Johnson, E. J., Haubl, G., and Keinan, A. (2007). Aspects of endowment: a query theory of value construction. *J. Exp. Psychol. Learn. Mem. Cogn.* 33, 461–474.

Johnson, E. J., Payne, J. W., and Bettman, J. R. (1988). Information displays and preference reversals. *Organ. Behav. Hum. Decis. Process* 42, 1–21.

Johnson, J. G., and Busemeyer, J. R. (2005). A dynamic, stochastic, computational model of preference reversal phenomena. *Psychol. Rev.* 112, 841–861.

Kable, J., and Glimcher, P. (2009). The neurobiology of decision: consensus and controversy. *Neuron* 63, 733–745.

Krajbich, I., Armel, K. C., and Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nat. Neurosci.* 13, 1292–1298.

Krajbich, I., and Rangel, A. (2011). Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions. *Proc. Natl. Acad. Sci. U.S.A.* 108, 13852–13857.

Leite, F. P., and Ratcliff, R. (2010). Modeling reaction time and accuracy of multiple-alternative decisions. *Atten. Percept. Psychophys.* 72, 246–273.

Milosavljevic, M., Malmaud, J., Huth, A., Koch, C., and Rangel, A. (2010). The drift diffusion model can account for the accuracy and reaction time of value-based choices under high and low time pressure. *Judgm. Decis. Mak.* 5, 437–449.

Milosavljevic, M., Navalpakkam, V., Koch, C., and Rangel, A. (2012). Relative visual saliency differences induce sizable bias in consumer choice. *J. Consum. Psychol.* 22, 67–74.

Philiastides, M., Biele, G., and Heekeren, H. (2010). A mechanistic account of value computation in the human brain. *Proc. Natl. Acad. Sci. U.S.A.* 107, 9430–9435.

Raab, M., and Johnson, J. G. (2007). Expertise-based differences in search and option-generation strategies. *J. Exp. Psychol. Appl.* 13, 158–170.

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

Ratcliff, R. (2002). A diffusion model account of response time and accuracy in a brightness discrimination task: fitting real data and failing to fit fake but plausible data. *Psychon. Bull. Rev.* 9, 278–291.

Ratcliff, R., and McKoon, G. (1997). A counter model for implicit priming in perceptual word identification. *Psychol. Rev.* 104, 319–343.

Ratcliff, R., and McKoon, G. (2008). The diffusion model: theory and data for two-choice decision tasks. *Neural Comput.* 20, 873–922.

Ratcliff, R., and Rouder, J. N. (2000). A diffusion model account of masking in two-choice letter identification. *J. Exp. Psychol. Hum. Percept. Perform.* 26, 127–140.

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

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.

Rushworth, M. F. S., Noonan, M. P., Boorman, E. D., Walton, M. E., and Behrens, T. E. (2011). Frontal cortex and reward-guided learning and decision-making. *Neuron* 70, 1054–1069.

Russo, J. E. (2010). "Eye fixation as a process trace," in *A Handbook of Process Tracing Methods for Decision Research*, eds M. Schulte-Mecklenbeck, A. Kuehberger, and R. Ranyard (New York: Psychology Press), 43–64.

Russo, J. E., Carlson, K. A., and Meloy, M. G. (2006). Choosing an inferior alternative. *Psychol. Sci.* 17, 899–904.

Russo, J. E., Carlson, K. A., Meloy, M. G., and Yong, K. (2008). The goal of consistency as a cause of information distortion. *J. Exp. Psychol. Gen.* 137, 456–470.

Russo, J. E., and Yong, K. (2011). The distortion of information to support an emerging evaluation of risk. *J. Econom.* 162, 132–139.

Shimojo, S., Simion, C., Shimojo, E., and Sheier, C. (2003). Gaze bias both reflects and influences preference. *Nat. Neurosci.* 6, 1317–1322.

Smith, P. L., and Ratcliff, R. (2004). Psychology and neurobiology of simple decisions. *Trends Neurosci.* 27, 161–168.

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

Tsetsos, K., Usher, M., and McClelland, J. L. (2011). Testing Multi-Alternative Decision Models with Non-Stationary Evidence. *Front. Neurosci.* 5:63. doi:10.3389/fnins.2011.00063

Usher, M., Elhalal, A., and Mcclelland, J. L. (2008). "The neurodynamics of choice, value-based decisions, and preference reversal," in *The Probabilistic Mind: Prospects for Bayesian Cognitive Science*, eds N. Chater and M. Oaksford (New York: Oxford University Press Inc.), 277–300.

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

Usher, M., and McClelland, J. L. (2004). Loss aversion and inhibition in dynamical models of multialternative choice. *Psychol. Rev.* 111, 757–769.

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 February 2012; accepted: 25 May 2012; published online: 13 June 2012.

*Citation: Krajbich I, Lu D, Camerer C and Rangel A (2012) The attentional drift-diffusion model extends to simple purchasing decisions. Front. Psychology **3**:193. doi: 10.3389/fpsyg.2012.00193*

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

APPENDIX

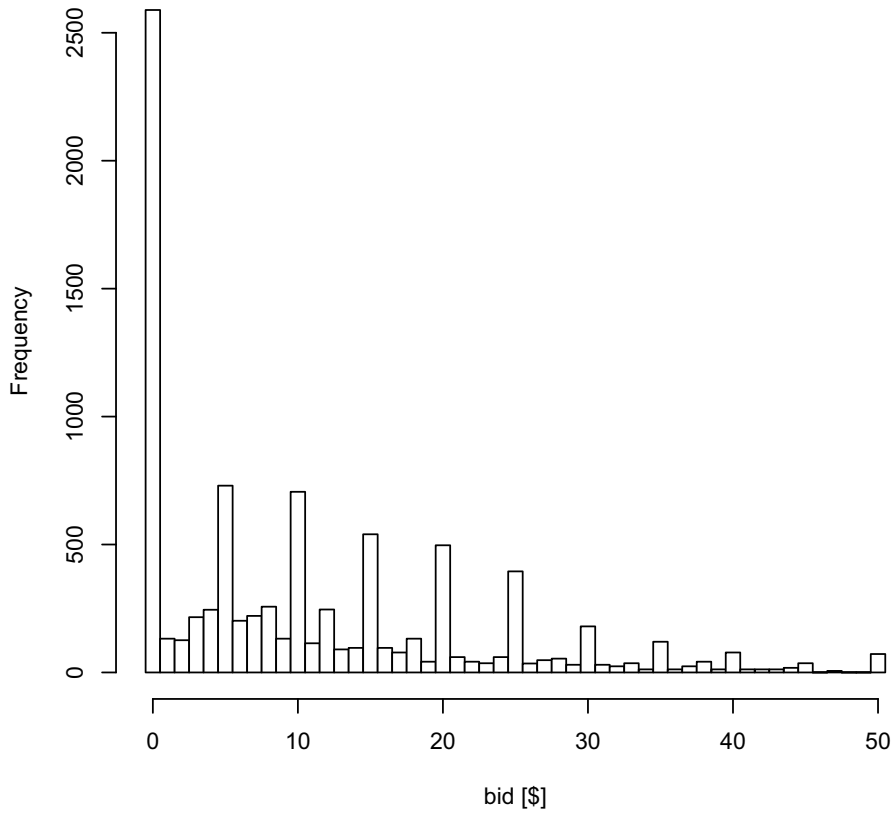


FIGURE A1 | Histogram of the bids for the various products in the BDM task, using all the data.

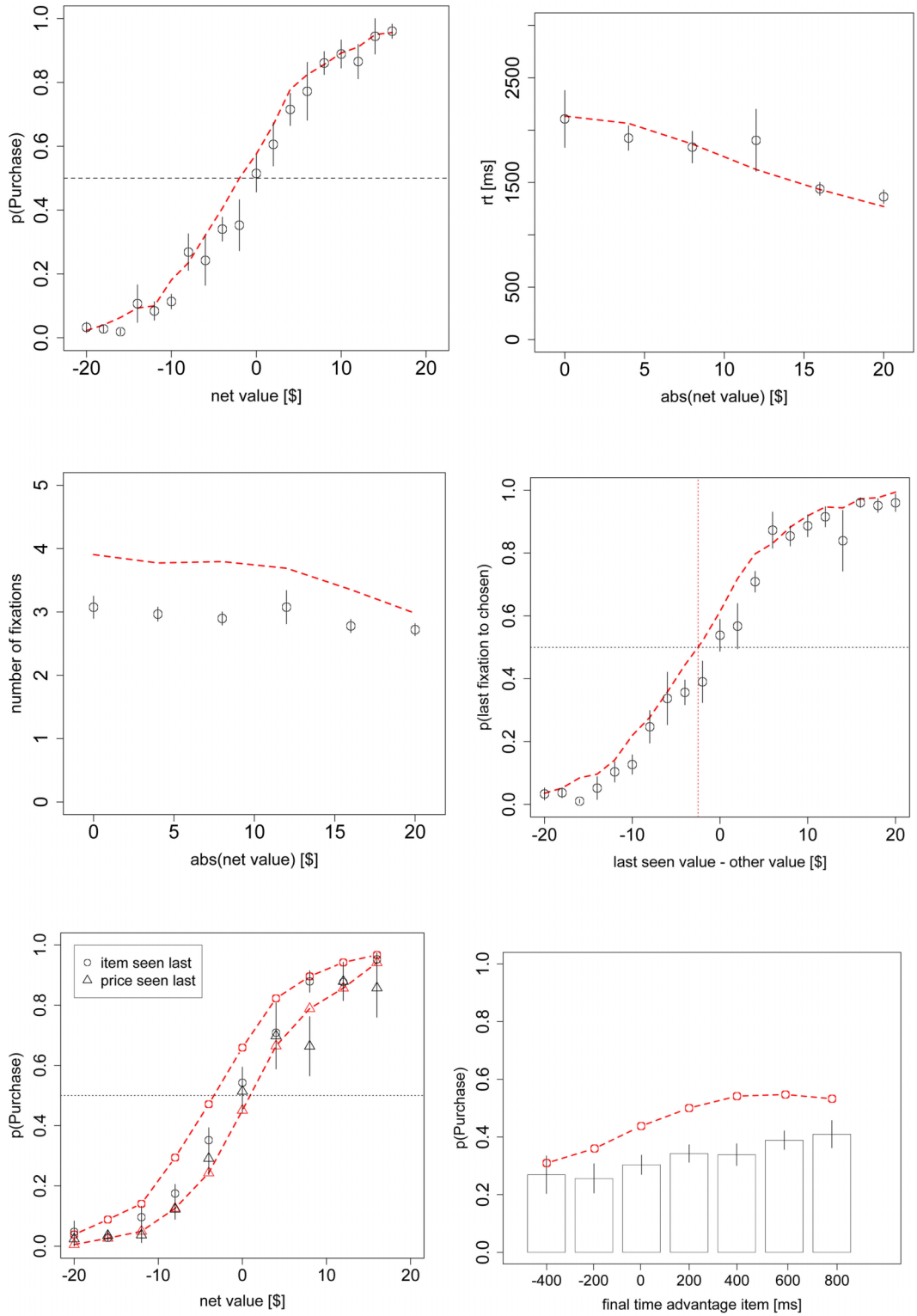


FIGURE A2 | Continued

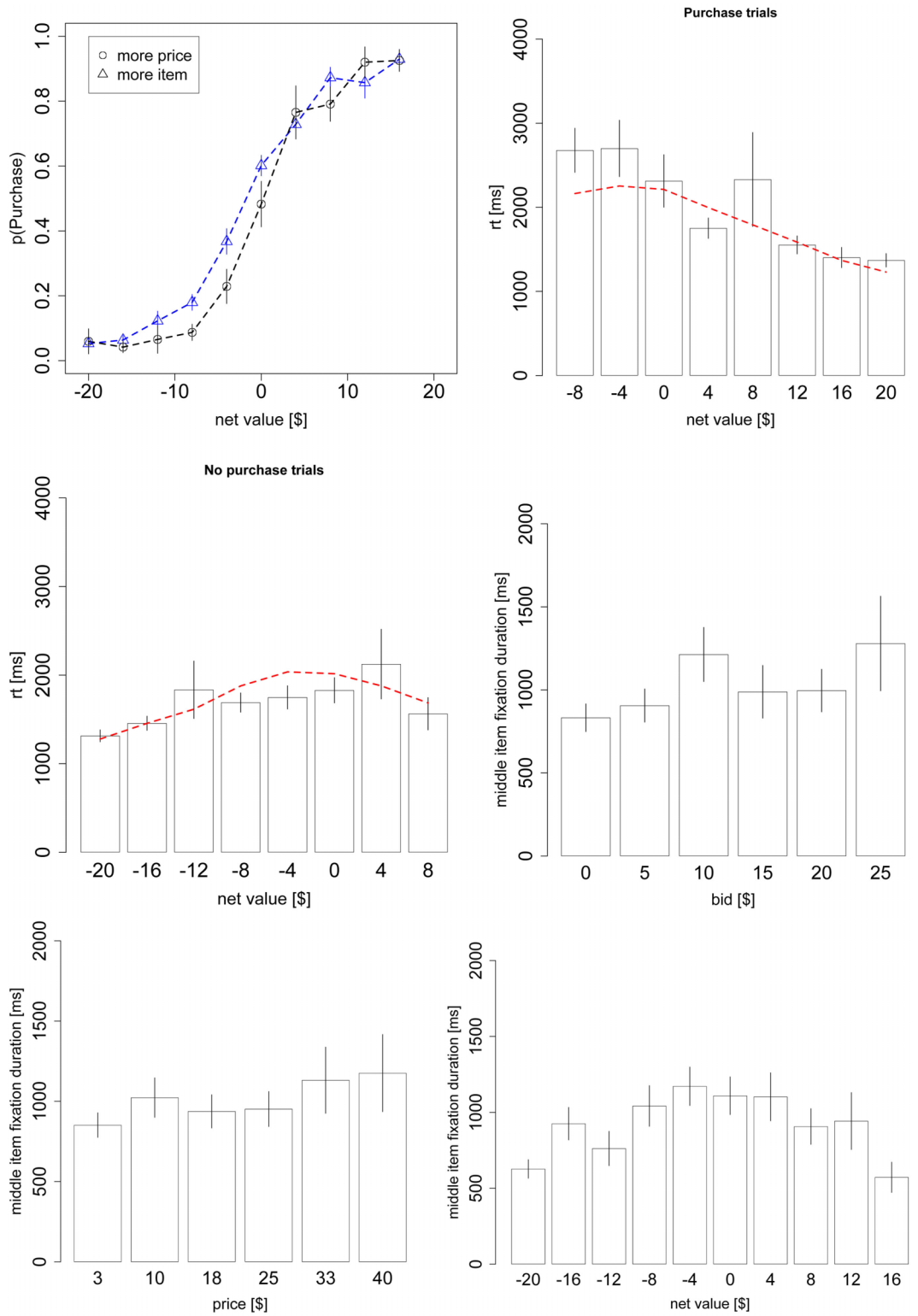


FIGURE A2 | Continued

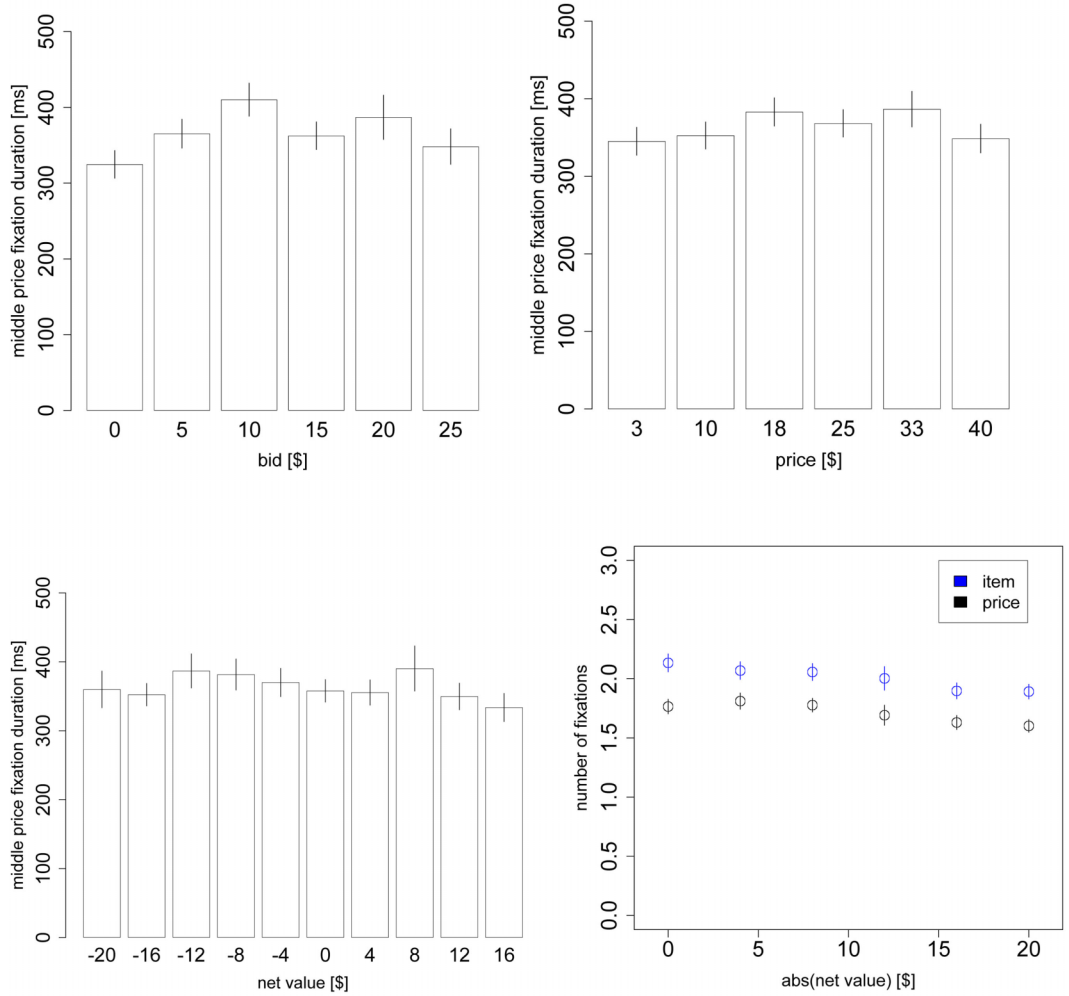


FIGURE A2 | Replication of all the figures from the text but including the $0 products.

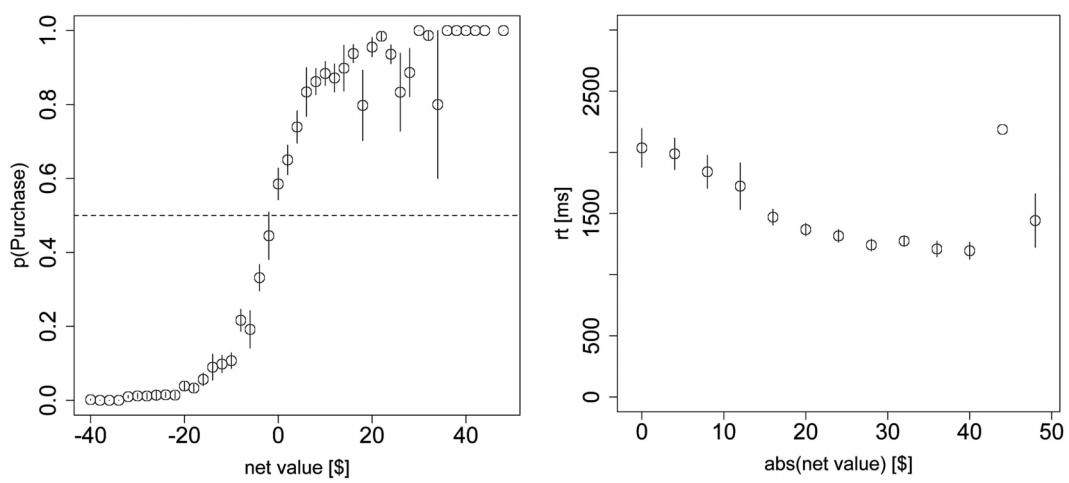


FIGURE A3 | Replication of Figures 2A,B choice and reaction time curves but using all of the data, including the $0 bids and net values beyond +/−$20.

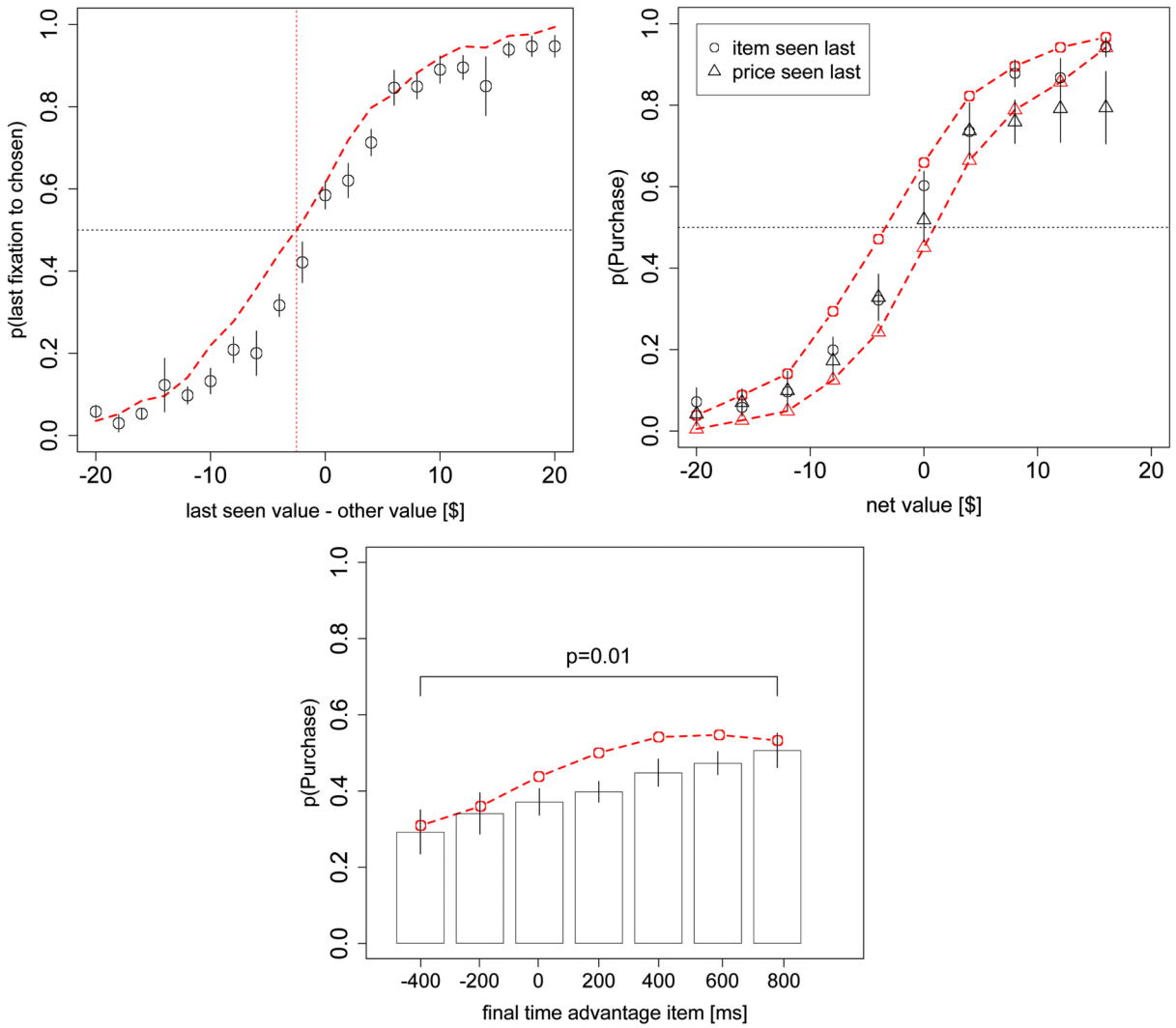


FIGURE A4 | Replication of Figures 3A,B, and 4A but using both the even and odd-numbered trials.

frontiers in
PSYCHOLOGY

METHODS ARTICLE
published: 20 June 2012
doi: 10.3389/fpsyg.2012.00189

The 2*N*-ary choice tree model for *N*-alternative preferential choice

Lena M. Wollschläger and Adele Diederich*

School of Humanities and Social Sciences, Jacobs University Bremen, Bremen, Germany

Edited by:
Marius Usher, Tel-Aviv University, Israel

Reviewed by:
Maarten Speekenbrink, University College London, UK
Jerome Busemeyer, Indiana University, USA

****Correspondence:***
Adele Diederich, School of Humanities and Social Sciences, Jacobs University Bremen, Research IV, Campus Ring 1, 28759 Bremen, Germany.
e-mail: a.diederich@jacobs-university.de

The 2*N*-ary choice tree model accounts for response times and choice probabilities in multi-alternative preferential choice. It implements pairwise comparison of alternatives on weighted attributes into an information sampling process which, in turn, results in a preference process. The model provides expected choice probabilities and response time distributions in closed form for optional and fixed stopping times. The theoretical background of the 2*N*-ary choice tree model is explained in detail with focus on the transition probabilities that take into account constituents of human preferences such as expectations, emotions, or socially influenced attention. Then it is shown how the model accounts for several context-effects observed in human preferential choice like similarity, attraction, and compromise effects and how long it takes, on average, for the decision. The model is extended to deal with more than three choice alternatives. A short discussion on how the 2*N*-ary choice tree model differs from the multi-alternative decision field theory and the leaky competing accumulator model is provided.

Keywords: 2*N*-ary choice tree model, preferential choice, multiple choice alternatives, multi-attribute choice alternatives, response times, choice probabilities, elimination of choice alternatives, computational model

1. INTRODUCTION

Life is full of decisions: Be it the selection of clothing in the morning or of menu for lunch, the question which car to buy or if taking cold medication is necessary. This type of decisions is called preferential choice and has been subject of numerous investigations within the field of decision theory (Koehler and Harvey, 2007, for a review). Several effects have been observed when the decision maker has more than two-choice options (multi-alternative preferential choice). Hick's Law (Hick, 1952; Hyman, 1953), originally defined in the context of stimulus detection paradigms, postulates a dependency of deliberation time on the number of alternatives. In particular, it states that a linear increase of the number of equally attractive alternatives to choose from leads to a logarithmic increase of the time that passes until the decision is made. Furthermore, a decision maker who is indifferent between two-choice alternatives from a given choice set may change the preference for one or the other alternative when the choice set is enlarged, i.e., the local context may affect the decision and generate preference reversals. Similarity effects (Tversky, 1972), attraction effects (Huber et al., 1982), and compromise effects (Simonson, 1989), for instance, depend on a third alternative that is added to a choice set of two equally attractive but dissimilar alternatives. If the third alternative is very similar to one of the others, the two similar alternatives share their choice frequency and are both chosen less often than the dissimilar one (similarity effect). If the third alternative is similar to one of the others but slightly inferior, it promotes the similar one and increases its choice frequency compared to the dissimilar one (attraction effect). If the third alternative is a compromise between the other two, the decision maker will prefer the compromise to the other alternatives (compromise effect). Besides those preference reversals that emerge from local context, there might also be influence from background context (Tversky and Simonson, 1993) like a reference point outside of the choice set which – together with the loss-aversion principle (Kahneman and Tversky, 1979) – affects evaluation of the given alternatives.

One challenge for (cognitive) modelers is to think of a model which predicts decision making behavior for multi-alternative preferential choice tasks in general but also accounts for all the aforementioned effects. Another challenge is to formulate the model such that (expected) response times and choice probabilities can be calculated and the model parameters conveniently estimated from the observed choice times and choice frequencies.

Decision field theory (DFT, Busemeyer and Townsend, 1992, 1993) and its multi-attribute extension (Diederich, 1997) predict choice response times and choice probabilities for binary choice tasks. Both approaches provide closed form solutions to calculate these entities. Since then, several attempts have been made to extend this kind of models to multi-alternative preferential choice tasks: multi-alternative DFT (Roe et al., 2001) and the leaky competing accumulator (LCA) model (Usher and McClelland, 2001, 2004) predict choice probabilities for three alternative choice tasks but cannot account for optional choice times, i.e., the time the decision maker needs to come to a decision. Both approaches, however, do account for fixed stopping times, i.e., for an externally determined time limit. Furthermore, multi-alternative DFT and the LCA model both account for the similarity, attraction, and compromise effects using computer simulations to predict the patterns. To do so Roe et al. (2001) interpret DFT as a connectionist network and implement distance-dependent inhibition between the alternatives. Usher and McClelland (2001, 2004) add insights from perceptual choice and neuropsychology to the multi-alternative DFT and propose for their LCA model direct implementation of

loss-aversion by means of an asymmetric value function and global inhibition instead of distance-dependent inhibition.

Our 2N-ary choice tree model builds on the previous approaches and tries to overcome some of their problems. It is a general model for choice probabilities and response times in choice between N alternatives with D attributes. As such, it provides a way to calculate expected response times, response time distributions, and choice probabilities in closed form by determining the time course of an information sampling process via a random walk on a specific tree. It is able to account for similarity, attraction, and compromise effects which have been most challenging for previous models. In contrast to previous approaches, the 2N-ary choice tree model accounts for these effects without additional mechanisms like inhibition or loss-aversion and is thus more parsimonious. However, it is possible to implement these mechanisms if the situation requires it.

First, we describe the structure of the 2N-ary choice tree and the implementation of the random walk on it in general, including a discussion of initial values and stopping rules. Then we define expected choice probabilities and reaction times and state that these exist and can be calculated in finite time. The proof of this statement is given later in the paper. It is not essential for understanding the theory; we provide it rather as completing the theoretical derivations. Next, we show how to derive transition probabilities from the given alternatives in a specific choice set and therewith define the random walk for that set. A psychological interpretation of their constituents is given afterward. Finally, we demonstrate the predictive power of our model by showing several simulations for choice situations producing the similarity, attraction, and compromise effect and calculate expected hitting times and choice probabilities. We conclude with a comparison of the 2N-ary choice tree with its closest competitors, the multi-alternative DFT and the LCA model.

2. THE 2N-ARY CHOICE TREE MODEL

Making an informed decision usually implies sampling of information about the alternatives under consideration. In Psychology, information sampling processes (e.g., Townsend and Ashby, 1983; Luce, 1986, for review, LaBerge, 1962; Laming, 1968; Link and Heath, 1975; Townsend and Ashby, 1983; Luce, 1986; Ratcliff and Smith, 2004) have a long tradition and proven to be an adequate tool for detailed interpretation of decision making processes, mostly in perception as they provide insight about accuracy and time course of these processes. Poisson counter models (e.g., Pike, 1966; Townsend and Ashby, 1983; LaBerge, 1994; Diederich, 1995; Smith and Van Zandt, 2000; Van Zandt et al., 2000) are a special class of information sampling models that assume constant amounts of information being sampled at Poisson distributed points in time. (Multi-alternative) DFT (Busemeyer and Townsend, 1993; Roe et al., 2001) and the LCA model (Usher and McClelland, 2004) make use of information sampling principles in modeling preferential choice under uncertainty. Both models assume one counter per alternative and all of these counters are updated once per fixed time interval until one of them reaches a threshold. The amounts to update the counters depend on comparison of the alternatives and on already sampled information. In our 2N-ary choice tree model, only one counter per fixed time interval is updated with a fixed amount, but the probability for each counter to be updated depends on comparison of the alternatives and on already sampled information. With regard to its constituents it is thus based on the same principles as both DFT and the LCA model. As only one counter is updated per iteration, the next time for a specific counter to be updated depends on the given probabilities. Hence the technical component of the 2N-ary choice tree model resembles a counter model.

2.1. 2N-ARY CHOICE TREES

In contrast to the aforementioned models, the 2N-ary choice tree model assigns two counters to each of N alternatives in a given choice set. One of them samples positive information, i.e., information in favor of the respective alternative, the other one samples negative information, i.e., information against it. Their difference describes the actual preference state relating to that alternative. As an example, consider two alternatives A and B. The four counters are labeled A^+, A^-, B^+, and B^- and yield the preference states $Pref(A) = A^+ - A^-$ for alternative A and $Pref(B) = B^+ - B^-$ for alternative B. Beginning at a fixed point in time, the model chooses one counter and increases its state by one whenever a specific time interval h (e.g., 1 ms) has passed. The length h of the time interval can be chosen arbitrarily with a shorter time interval leading to more precision in the calculation of expected choice probabilities and choice response times. Due to limitations of recording devices, experimental data will be discrete as well and it is thus not necessary to aim for a continuous model. Note that increasing only one counter state at a time with a fixed amount of evidence equal to one is equivalent to increasing all counter states at the same time with an amount of evidence equal to the probability with which these counters are chosen and which also sum up to one (see below). Updating counters at discrete points in time creates a discrete structure of possible combinations of counter states which can be interpreted as graph or, more precisely, as (b-ary) tree[1].

Definition 1 (b-ary tree): A b-ary tree is a rooted tree $T = (V, E, r)$ with vertices V, edges $E \subseteq V \times V$ and root $r \in V$ where all vertices are directed away from r and each internal vertex has b children.

For N choice alternatives, consider a 2N-ary tree $T = (V, E, r)$. **Figure 1** depicts the 4-ary tree for the two-alternative example. The topmost vertex is the root r with outgoing edges directing to four vertices that represent the counters A^+, A^-, B^+, and B^-. Each of these vertices has four outgoing edges and thus four children itself, and so forth. The information sampling process is mapped to this tree as a walk, i.e., a sequence of edges and vertices, beginning with the root r that takes one step, i.e., passes from one vertex through an edge to another vertex, per time interval h. Whenever the walk reaches a vertex, the counter with the same label is updated by $+1$. **Figure 2** shows an example for a walk on the 4-ary tree where first counter B^+ (information in favor of choice alternative B), then counter A^- (unfavorable information for choosing alternative A) and then again counter B^+ is updated.

[1]Definitions of graph-related terms not defined here can be found in Korte and Vygen (2002).

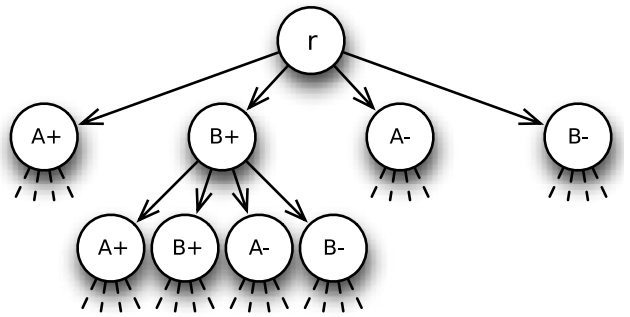


FIGURE 1 | 4-ary tree for choice between two alternatives *A* and *B*. The root *r* has four outgoing edges directing to four vertices that represent the counters A^+, A^-, B^+, and B^-. Each of these vertices has four outgoing edges and thus four children itself, and so forth.

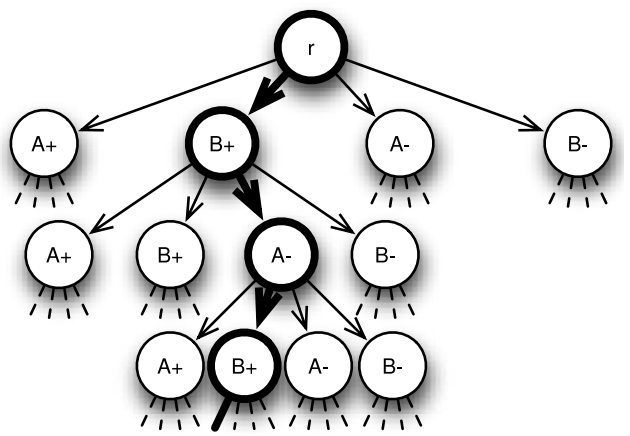


FIGURE 2 | 4-ary tree for choice between two alternatives *A* and *B* with highlighted sample path $r \rightarrow B^+ \rightarrow A^- \rightarrow B^+ \rightarrow \ldots$

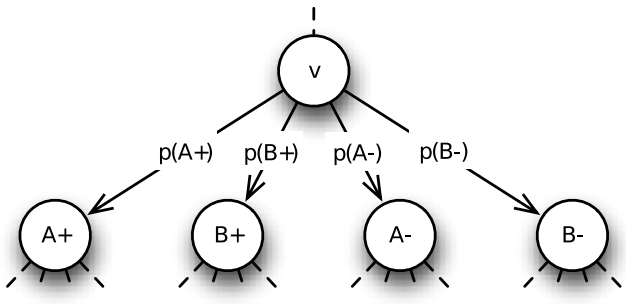


FIGURE 3 | Transition probabilities for the two-alternative choice problem with no counter-dependent and random component. $p(A^+)$, $p(B)^+$, $p(A)^+$, and $p(B^-)$ sum up to one and are the same for the outgoing edges of each vertex $v \in V$.

Three features of the model are of specific interest: (a) when and how the walk starts after presentation of choice alternatives (in an experimental trial), (b) how the walk chooses the next edge to pass through in each step, and (c) when and how the walk stops. Without an *a priori* bias toward any of the choice alternatives, we assume that all counter states are set to zero at the outset of the information sampling process and hence, the process starts with presentation of the choice alternatives. Biases toward one or several of the alternatives can be implemented by either defining initial values unequal to zero for these alternatives or by independently sampling information for the alternatives from predefined distributions for some time before the actual information sampling process starts (cf. Diederich and Busemeyer, 2006; Diederich, 2008). For simplicity, we assume no biases here, i.e., initial values are set to zero for all alternatives. Note that for the 2*N*-ary choice tree, initial values are counter states at the root *r*. Then the walk moves away from there step by step, choosing the next edge to pass through by means of so-called transition probabilities p_e, $e \in E$. The transition probabilities are built up of the comparison of the alternatives the decision maker considers and supplemented with a counter-dependent component and a random component. For each vertex, the transition probabilities for all outgoing edges sum up to one, so that the walk does not stay still at any vertex it reaches throughout the information sampling process. We show the structure of the model first; a detailed description of the transition probabilities is presented in the next section. For simplicity consider a choice situation with two alternatives *A* and *B*; the counter-dependent component and the random component are set to zero. As shown in **Figure 3**, transition probabilities are the same for the outgoing edges of each vertex $v \in V$, i.e., $p_{(v,v(A^+))} = p(A^+)$ for each edge $(v,v(A^+)) \in E$ leading to a vertex with label A^+, $p_{(v,v(B^+))} = p(B^+)$ for each edge leading to a vertex with label B^+ and so on for the other counters A^- and B^-. The probability for walking along a specific path is the product of transition probabilities of all edges on that path. In our example, the probability *p* for making the first three steps as shown in **Figure 2** is $p = p(B^+) \cdot p(A^-) \cdot p(\mathrm{B}^+)$.

The third topic addresses the stopping rule, that is, when the decision maker stops sampling information and chooses a choice alternative. A specific stopping rule depends on the preference states associated with the alternatives, i.e., the differences of their respective two counters which are compared to certain thresholds θ. The thresholds can be defined in several ways, their suitability depending on the definition of transition probabilities and initial values. They are (1) one single positive threshold $\theta^+ > 0$ for all alternatives, (2) one positive and one negative threshold $\theta^+ > 0$ and $\theta^- < 0$ for all alternatives, (3) a positive threshold θ_i^+ for each alternative $i \in \{1, \ldots, N\}$, and (4) a positive and a negative threshold θ_i^+ and θ_i^- for each alternative $i \in \{1, \ldots, N\}$.

Obviously, the simplest setup is a single positive threshold $\theta^+ > 0$ for all alternatives, which is hit as soon as the information sampled in favor of any/one of the alternatives exceeds the information against it by θ^+ for the first time, i.e., when $Pref(i) = \theta^+$ for one alternative $i \in \{1, \ldots, N\}$. Sometimes, however, the probability for collecting negative information may be greater than the probability for sampling information in favor of these alternatives and reaching a positive threshold θ^+ is very unlikely. For those situations it is useful to introduce a second, negative threshold, $\theta^- < 0$, which is hit when negative information of one alternative

exceeds the positive information of this alternative by $-\theta^-$, i.e., $Pref(i) = \theta^-$. In this case the respective alternative is not chosen but withdrawn from the choice set and the sampling process continues with one alternative less as described in the next paragraph. Note, that in both cases the thresholds are global in the sense that the same thresholds apply for all choice alternatives. Finally, θ^+ (and θ^-) may vary from alternative to alternative, yielding one (or two) thresholds θ_i^+ (and θ_i^-) for each of N alternatives, $i \in \{1, \ldots, N\}$. Here the thresholds are local in the sense that each alternative has its own threshold(s). This is an alternative way to implement biases when the initial values are zero. That is, biases do not affect transition probabilities through the counter-dependent component and can thus be interpreted as the decision maker's stable opinion about the presented alternatives.

Withdrawal of alternatives from a choice set traces back to the model of elimination by aspects (EBA model, Tversky, 1972). But whereas elimination is the only means to come to a decision in the EBA model, the 2*N*-ary choice tree model like the multi-alternative DFT (Roe et al., 2001) provides several ways to reach a decision. An alternative i is chosen either if its preference state exceeds θ_i^+ or if all other alternatives have been withdrawn from the choice set or a combination of these two. **Figure 4** shows three examples of walks that lead to the choice of alternative B from a set of three alternatives A, B, and C with global thresholds $\theta^+ = 2$ and $\theta^- = -2$. The leftmost walk represents direct choice of alternative B, the rightmost withdrawal of alternative C and subsequent choice of option B and the middle walk illustrates withdrawal of alternative A first and then of option C. Note that after withdrawal of one alternative, there are two-outgoing edges less from the respective vertex downward. Transition probabilities change accordingly, i.e., the withdrawn alternative is removed from the comparison procedure and its counter states no longer contribute to the counter-dependent component (cf. next section). This corresponds to an anew started information sampling process between the remaining alternatives and their previous counter states as initial values.

For a choice set with N alternatives and given thresholds $\theta_i^\pm$ this defines the structure of the 2*N*-ary choice tree. For each alternative $i \in \{1, 2, \ldots, N\}$ we can thus completely identify the set $V_i \subseteq V$ of vertices where alternative i is chosen. Defining $P_v := P_{\{r,\, v\}}$ to be the unique path from the root r to a vertex $v \in V$ and given transition probabilities p_e for all edges $e \in E$ we can identify the probability for walking along a path P_v as the product $p_v = \Pi_{e \in P_v} p_e$ and therewith define:

Definition 2 (expected choice probability): The expected probability for choosing alternative $i \in \{1, 2, \ldots, N\}$ is the probability for reaching the set V_i:

$$p_i = \sum_{v \in V_i} \prod_{e \in P_v} p_e. \tag{1}$$

The length $|P_v|$ of the path P_v from r to $v \in V$ indicates the number of steps that the random walk has to take to reach v. Multiplied by the length h of the time interval, this yields the time it takes to cover the distance from r to v. Thus $T_v = h \cdot |P_v|$.

Definition 3 (expected hitting time): The expected time for choosing alternative $i \in \{1, 2, \ldots, N\}$ is the sum of expected hitting times for each vertex $v \in V_i$ weighted by the probability for reaching v:

$$\mathbb{E}[T_i] = h \cdot \sum_{v \in V_i} |P_v| \cdot p_v = h \cdot \sum_{v \in V_i} |P_v| \cdot \prod_{e \in P_v} p_e. \tag{2}$$

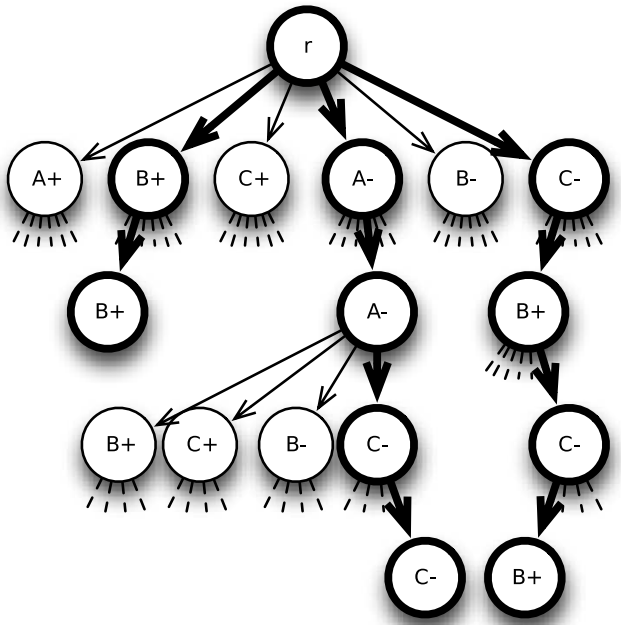


FIGURE 4 | 6-Ary tree for choice between three alternatives *A*, *B*, and *C* with decision thresholds $\theta^+ = 2$ and $\theta^- = 2$ and three-different sample paths that lead to choice of alternative *B*.

The expected choice probabilities and hitting times can be approximated up to absolute accuracy in finite time. See below for the formal statement and proof of this property.

2.2. TRANSITION PROBABILITIES

Having defined the skeletal structure of our theory, we can now proceed to its heart, the transition probabilities. The main components are (a) weighted comparison of alternatives, (b) mutual or global inhibition, (c) decay of already sampled information over time, and (d) a random part. The transition probabilities control the information sampling process and thus describe the development of human preferences in specific choice situations. Throughout this section we will consider such situations with N choice alternatives that are evaluated with respect to the same D attributes. For each alternative, the decision maker is provided with one non-negative value per attribute, representing an objective evaluation of that alternative with respect to the attributes. The $N \cdot D$ values in total can be stored in a $N \times D$-matrix $L = (l_{ij})$ with $i = 1, \ldots, N$ and $j = 1, \ldots, D$.

The definition of transition probabilities is based on weighted integration of results of an attribute-wise comparison of alternatives. To ensure equally significant impact of the weight parameters, preprocessing of the values of the alternatives with respect to the attributes is necessary and we do so by separately normalizing them to one for each attribute. This yields a new

matrix $M = (m_{ij}) = (l_{ij}/\sum_{k=1}^{N} l_{kj})$ with $i = 1, \ldots, N$ and $j = 1, \ldots, D$ and thus column sum $\sum_{i=1}^{N} m_{ij} = 1$ for all $j \in \{1, \ldots, D\}$.

2.2.1. Comparison of alternatives

At first we focus on one attribute j only. The easiest way to define transition probabilities is simply to use the entries from the respective column j of M and assign them to the edges that affiliate with the counters for positive information. The counters for negative information get transition probabilities zero. This corresponds to a framework where the alternatives are compared to an inferior external reference point (e.g., cars A, B, and C are compared to not having a car at all). Because the values for each attribute sum up to one already, no further normalization is needed.

We differentiate between external reference points that are not part of the choice set and internal reference points that are part of the choice set and available for the decision maker. For instance, if someone moves to a new city and has to choose between several available apartments, she will probably compare them to her old apartment which is no longer available in the new city and thus an example for an external reference point. Or consider a choice of dessert in a restaurant when the decision maker is told that the chocolate cake she ordered is no longer available because someone just had the last piece. An internal reference point, however, is part of the choice set, actually several or even all available alternatives can be used as internal reference points at the same time, possibly in combination with an external reference point.

Having decided which reference points to use, the alternatives $i \in \{1, \ldots, N\}$ are compared to them. For each alternative i, favorable and unfavorable comparisons are handled separately and the absolute values of their differences are summed up to obtain measures of evidence for and against alternative i, respectively. This yields two non-negative values per alternative and thus $2N$ values in total that are then normalized to one in order to obtain probabilities. In the car example where three cars are compared to not having a car at all, probabilities associated with negative counters are set to zero as each car is better than, presumably, no car. Actually, whenever one single reference point is used, at least half of the probabilities are zero because each alternative is either favored over the reference point or not and hence there cannot be evidence for and against one alternative at the same time. However, our main focus is on situations where each of at least three alternatives in the choice set is used as internal reference point for all the other alternatives and thus there are at least two-reference points.

In this case we obtain a vector

$$P_j = \begin{pmatrix} p_{1j} \\ \vdots \\ p_{Nj} \\ p_{(N+1)j} \\ \vdots \\ p_{(2N)j} \end{pmatrix} = \begin{pmatrix} p_{1j}^{+} \\ \vdots \\ p_{Nj}^{+} \\ p_{1j}^{-} \\ \vdots \\ p_{Nj}^{-} \end{pmatrix} / \sum_{i=1}^{N} \left(p_{ij}^{+} + p_{ij}^{-} \right)$$

with

$$p_{ij}^{+} = \sum_{k \neq i} \left(m_{ij} - m_{kj} \right) \cdot \mathbb{I} \left(m_{ij} > m_{kj} \right),$$

$$p_{ij}^{-} = \sum_{k \neq i} \left(m_{kj} - m_{ij} \right) \cdot \mathbb{I} \left(m_{ij} < m_{kj} \right)$$

for $k = 1, \ldots N$ and

$$\mathbb{I}(x) = \begin{cases} 1, & \text{if } x \text{ is true} \\ 0, & \text{else.} \end{cases}$$

Especially with an external reference point at hand, the actual choice may lead to a loss of some kind. For instance, in the apartment example above a loss could be a further way to the workplace or a smaller bathroom. People usually try to avoid losses more than they seek gains while overrating small losses and gains compared to larger ones (Kahneman and Tversky, 1991; Tversky and Simonson, 1993). The $2N$-ary choice tree model can account for the loss-aversion principle (Kahneman and Tversky, 1979) with an asymmetric value function (Kahneman and Tversky, 1979; Tversky and Simonson, 1993) by increasing probabilities for sampling negative information compared to probabilities for gathering positive information.

In their LCA model, Usher and McClelland (2004) use an asymmetric value function

$$V(x) = \begin{cases} \log(1 + x), & \text{for } x > 0 \\ -\left[\log(1 + |x|) + \log(1 + |x|)^2 \right], & \text{for } x < 0 \end{cases}$$

and apply it to the relative advantages ($x > 0$) and disadvantages ($x < 0$) of alternatives compared to each other on one dimension. $V(x)$ is steeper for losses than for gains but flattens for both advantages and disadvantages when they become bigger. This favors similar pairs of alternatives over dissimilar ones and allows the LCA model to account for attraction and compromise effects.

Adopting it to our $2N$-ary choice tree model this yields an asymmetric value function

$$A(x) = \begin{cases} \log(1 + x), & \text{for favorable comparisons} \\ \log(1 + x) & \\ \quad + \log(1 + x)^2, & \text{for unfavorable comparisons,} \end{cases}$$

which can be applied to the absolute differences from the comparison process before normalizing them to one:

$$p_{ij}^{+} = \sum_{k \neq i} A \left(m_{ij} - m_{kj} \right) \cdot \mathbb{I} \left(m_{ij} > m_{kj} \right),$$

$$p_{ij}^{-} = \sum_{k \neq i} A \left(m_{kj} - m_{ij} \right) \cdot \mathbb{I} \left(m_{ij} < m_{kj} \right).$$

The asymmetric value function $A(x)$ is not necessary for explanation of similarity, attraction, or compromise effects in the $2N$-ary choice tree model but moderates the strength of the compromise effect (see below). In cases where θ^{+} and θ^{-} have the same

order of magnitude, application of $A(x)$ leads to faster withdrawal of alternatives and hence, more decisions are based on withdrawal of all but one alternative.

In summary, the comparison of alternatives provides us with a set of $2N$ transition probabilities for each attribute $j \in \{1, \ldots, D\}$ that form a vector P_j. Each of these vectors can be used to model an information sampling process based on a single attribute. As the probabilities are derived from comparison of alternatives only, they remain constant during the whole process.

2.2.2. Weighting of attributes

So far we have only focused on one attribute but choice alternatives in real life are most often described by several attributes and thus require more elaboration. In the following, we consider choice sets with N alternatives characterized by $D \geq 2$ attributes. Especially situations with three alternatives where similarity, attraction or compromise effects have been observed, require at least two attributes to distinguish the different alternatives from each other. Note that it is difficult to construct a choice set with two equally attractive but different alternatives due to the decision maker's individual salience. Diederich (1997) accounts for subjective salience by defining a Markov process on the attributes giving probabilities for switching attention from one attribute to the other. This process can be directly implemented into the transition probabilities by using a stationary distribution on the attributes. Each attribute $j \in \{1, \ldots, D\}$ is assigned a weight w_j that corresponds to the probability for considering this attribute during the information sampling process. For each alternative $i \in \{1, \ldots, N\}$ weighted positive and negative evidence is added up and normalized to obtain a proper probability distribution (the probabilities add up to one), that is

$$P = \begin{pmatrix} p_1 \\ \vdots \\ p_N \\ p_{(N+1)} \\ \vdots \\ p_{(2N)} \end{pmatrix} = \begin{pmatrix} p_1^+ \\ \vdots \\ p_N^+ \\ p_1^- \\ \vdots \\ p_N^- \end{pmatrix} / \sum_{i=1}^{N} \left(p_i^+ + p_i^- \right)$$

with $p_i^+ = \sum_{j=1}^{D} (p_{ij}^+ \cdot w_j)$ and $p_i^- = \sum_{j=1}^{D} (p_{ij}^- \cdot w_j)$ for $i \in \{1, \ldots, N\}$.

The weights account for subjective salience that in turn may be influenced by several internal and external factors such as personal preferences, social influences, characteristics of the choice set, or the experimenter's instructions. Personal preferences like, for instance, the preference of time over money or of tastiness to healthiness may be learned from friends, family, or other people in our surrounding. They are generally independent from the choice situation and hence, their impact on the information sampling is indirect. On the contrary, the choice set itself has a direct influence on the subjective saliences. For example, the decision maker may primarily focus on those attributes where alternatives are very similar to each other, because this information may be crucial for the choice. Or she concentrates on attributes with somehow outstanding values. It is therefore important to normalize the values for each attribute as described before because this guarantees representation of these effects by the attention weights. People that are present during the deliberation process like sales people or immediately prior to it like the experimenter in a laboratory context can also have a direct influence on the saliences by drawing the decision maker's attention to a specific attribute. This can be used to verify influence of attention weights by instructing decision makers to focus on certain attributes while choosing between different cars, salad dressings, chocolate bars or shoes. Corresponding experiments are under way.

2.2.3. Noise

In order to account for random fluctuations in the decision maker's attention (cf. Busemeyer and Townsend, 1993) which are independent of the characteristics of the choice alternatives, we add a constant to each transition probability. This makes every outgoing edge of a vertex $v \in V$ available for the next (random) step because it guarantees non-zero transition probabilities for all of them. Let $\mathcal{N}$ be a vector of length $2N$ with all entries equal to $1/2N$. Weighting the transition probabilities P from the weighted comparison of alternatives by $(1 - \xi)$ with $0 \leq \xi \leq 1$ and adding the product $\xi \cdot \mathcal{N}$ yields noisy transition probabilities where ξ moderates the strength of the uniformly distributed noise:

$$P_{\mathcal{N}} = (1 - \xi) \cdot P + \xi \cdot \mathcal{N}.$$

The vector $P_{\mathcal{N}}$ of noisy transition probabilities integrates comparison of alternatives on all present attributes. Related to the $2N$-ary choice tree, this information is global as it is independent of the local counter states and thus the transition probabilities are the same for the edges emanating from each vertex.

2.2.4. Leakage

During their development of DFT (Busemeyer and Townsend, 1993) introduce a factor s for serial positioning effects, called "growth-decay rate." It produces recency effects for $0 < s < 1$ and primacy effects for $s < 0$. In their multi-alternative version of DFT (Roe et al., 2001) the reverse $(1 - s)$ of this factor reappears as "self-feedback loop" and accounts for the memory of previous preference states. $(1 - s) = 1$ denotes perfect memory of the previous state $(1 - s) = 0$ no memory at all. For their simulations (Roe et al., 2001) use $(1 - s) = 0.94$ or $(1 - s) = 0.95$. Usher and McClelland (2001, 2004) adopted the idea of the self-feedback loop, but call it "leakage" λ and – based on findings from neuroscience – interpret it as "neural decay."

In order to account for decay of already sampled information over time, we implement leakage $\mathcal{L}$ into our transition probabilities. Leakage obviously depends on already sampled information and thus we normalize the current states of our $2N$ counters to $1 - \lambda$ and for each alternative $i \in \{1, \ldots, N\}$ add the result for the positive (negative) counter of alternative i to the transition probability associated with the negative (positive) counter for alternative i weighted by λ. Like this, the overall sum of the transition probabilities remains 1 and only $100 \cdot \lambda\%$ of the sampled information is actually memorized. The greater λ, the longer it takes until the process reaches a threshold. Overall, this yields

$$P_{\mathcal{N}\mathcal{L}} = (1 - \lambda) \cdot [(1 - \xi) \cdot P + \xi \cdot \mathcal{N}] + \lambda \cdot \mathcal{L}.$$

2.2.5. Inhibition

To account for the similarity, attraction, and compromise effect, DFT (Roe et al., 2001) and the LCA model (Usher and McClelland, 2004) both rely on inhibition. Whereas distance-dependent inhibition enables DFT to account for the attraction and compromise effect, global inhibition produces the similarity effect in the LCA model. We can implement both types of inhibition into the $2N$-ary choice tree model to explore their impact on the aforementioned effects. We define weights for all pairs of alternatives by either using the same weight for all pairs like Usher and McClelland (2004) do with their "global inhibition" parameter β or different weights like Roe et al. (2001) do with their distance-dependent weights (i.e., higher weights for more similar alternatives). Those weights can be stored in a symmetric $N \times N$-matrix with zeros on the diagonal.

Taking into account the basic concept of inhibition, we assume that the state of the positive counter for each alternative $i \in \{1, \ldots, N\}$ reduces sampling of positive information for all other alternatives $j \in \{1, \ldots, N\} - \{i\}$. Because this is equivalent to increasing sampling of negative information for these alternatives and vice versa for states of negative counters, we implement inhibition $\mathscr{I}$ into our model as follows: Multiplying the symmetric $N \times N$-matrix with both the vector of states of positive counters and negative counters yields two vectors with weighted sums of counter states. We concatenate them in inverted order and normalize the resulting vector of length $2N$ to μ before adding it to the vector of transition probabilities now weighted by $(1 - \lambda - \mu)$. This completes the final definition of transition probabilities

$$P_{\mathscr{N}\mathscr{L}\mathscr{I}} = (1 - \lambda - \mu) \cdot [(1 - \xi) \cdot P + \xi \cdot \mathscr{N}] + \lambda \cdot \mathscr{L} + \mu \cdot \mathscr{I}.$$

In a nutshell, the transition probabilities consist of a global part that is independent from the current counter states of the random walk and a local part that depends on already sampled information. The global parts are weighted sums of comparative values P and noise $\mathscr{N}$ that remain constant during the whole process. They are complemented with leakage $\mathscr{L}$ and inhibition $\mathscr{I}$ which may change from step to step and hence, are local in the terminology of the $2N$-ary choice tree model.

3. PREDICTIONS OF THE 2N-ARY CHOICE TREE MODEL

To show the predictions of the $2N$-ary choice tree model and how it accounts for similarity, attraction, and compromise effects in choice settings with three alternatives characterized by two attributes, we run several simulations. An extension to more alternatives is straightforward. We will define values l_{ij} that range between 0 and 10. As values of choice alternatives are normalized to one on each dimension before comparison, only the relative amount of these values is of importance. Unless stated otherwise, we run 1000 trials per simulation with threshold $\theta = 20$, noise factor $\xi = 0.01$ and leakage factor $\lambda = 0.05$, but without inhibition (i.e., $\mu = 0$).

In order to meet the assumptions of the similarity, attraction, and compromise effect, we constructed two equally attractive but dissimilar alternatives $A = (9, 1)$ and $B = (1, 9)$ that are both evaluated with respect to two attributes. The choice probabilities were 0.52, 0.51, and 0.47 for alternative A and 0.48, 0.49, and 0.53 for option B in three simulations with the above mentioned parameters and attribute weight 0.5 for both attributes.

To reproduce the similarity effect (Simonson, 1989) we add a third alternative to the choice set that is equal or similar to either option A or B, i.e., $C = (1, 9)$, $C_2 = (0.9, 9.1)$, or $C_3 = (1.1, 8.9)$. To prevent a combination of the similarity effect with a slight compromise effect (cf. Usher and McClelland, 2004), we will use only C for demonstration, but the results for options C_2 and C_3 are very similar to the ones presented here. The alternatives are put together in a 3×2-matrix L, whose columns are normalized to one, resulting in matrix M:

$$L = \begin{pmatrix} 9 & 1 \\ 1 & 9 \\ 1 & 9 \end{pmatrix} \quad \text{and} \quad M = \begin{pmatrix} 0.818 & 0.053 \\ 0.091 & 0.474 \\ 0.091 & 0.474 \end{pmatrix}.$$

M already shows smaller values for alternatives B and C on the second dimension than for alternative A on dimension one which characterizes the similarity effect. In the next step, the values on each dimension are compared to each other, resulting in a 6×2-matrix that is then multiplied by $W = \begin{pmatrix} 0.4 \\ 0.6 \end{pmatrix}$ before being normalized to one again:

$$P' = \begin{pmatrix} 1.4545 & 0 \\ 0 & 0.4211 \\ 0 & 0.4211 \\ 0 & 0.8421 \\ 0.7273 & 0 \\ 0.7273 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0.4 \\ 0.6 \end{pmatrix} = \begin{pmatrix} 0.5818 \\ 0.2526 \\ 0.2526 \\ 0.5053 \\ 0.2909 \\ 0.2909 \end{pmatrix}.$$

and

$$P = \begin{pmatrix} 0.5818 \\ 0.2526 \\ 0.2526 \\ 0.5053 \\ 0.2909 \\ 0.2909 \end{pmatrix} / 2.1742 = \begin{pmatrix} 0.2676 \\ 0.1162 \\ 0.1162 \\ 0.2324 \\ 0.1338 \\ 0.1338 \end{pmatrix}$$

Finally noise is added to this constant part of the transition probabilities. In contrast to leakage that depends on the respective counter states and has to be computed anew for every step, $P_{\mathscr{N}}$ remains constant over time. The only occasion where it changes is after withdrawal of one alternative from the choice set.

The most interesting parameters in this attempt to model a similarity effect are the attribute weights as they control the strength of the effect. **Figure 5** demonstrates this by means of choice probabilities from simulations with different sets of attribute weights but otherwise unchanged parameters. It starts with $W = \begin{pmatrix} 0.6 \\ 0.4 \end{pmatrix}$ and $W = \begin{pmatrix} 0.55 \\ 0.45 \end{pmatrix}$ on the left side and gradually changes by 0.05 to $W = \begin{pmatrix} 0.25 \\ 0.75 \end{pmatrix}$ on the right side. The relative frequency of choices for alternatives A, B, and C including the mean number of steps leading to these choices can be found in **Table 1**.

The same mechanisms account for the attraction effect (Huber et al., 1982) that occurs during choice between two equally attractive but dissimilar alternatives A and B and a third alternative C

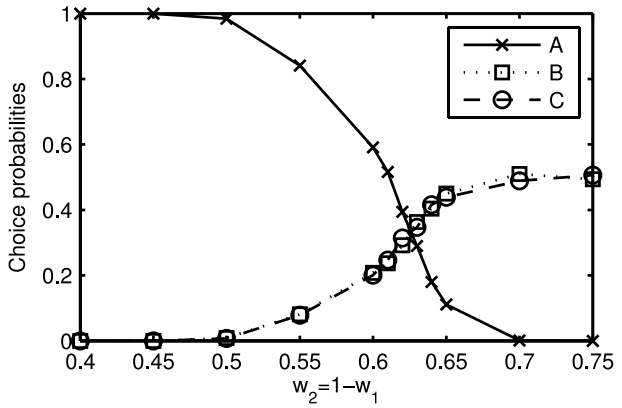


FIGURE 5 | Choice probabilities for choice between three alternatives A = (9, 1), B = (1, 9), and C = (1, 9) and different attention weights w_1 and w_2 for the two attributes. The abscissa is labeled with increasing values of w_2 corresponding to decreasing values of w_1. For $w_2 < 0.625$ a similarity effect can be observed.

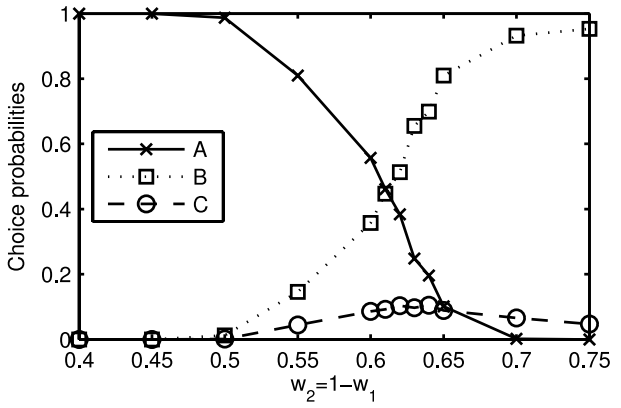


FIGURE 6 | Choice probabilities for choice between three alternatives A = (9, 1), B = (1, 9), and C = (1, 8.5) and different attention weights w_1 and w_2 for the two attributes. The abscissa is labeled with increasing values of w_2 corresponding to decreasing values of w_1. For $0.61 \leq w_2 < 0.65$ an attraction effect can be observed.

Table 1 | Relative number of choices and mean response times (arbitrary unit, in parentheses) for alternatives A = (9, 1), B = (1, 9), and C = (1, 9) from 1000 simulations with $\theta = 20$, $\xi = 0.01$, $\lambda = 0.05$, $\mu = 0$, and $w_2 = 1 - w_1$ ranging from 0.4 to 0.75 as indicated in the first row.

w_2	0.4	0.45	0.5	0.55
A	1 (92.4)	1 (116.4)	0.984 (163.0)	0.841 (205.5)
B	0 (–)	0 (–)	0.009 (596.4)	0.081 (555.6)
C	0 (–)	0 (–)	0.007 (475.4)	0.078 (524.6)
w_2	0.6	0.61	0.62	0.63
A	0.591 (316.5)	0.516 (358.3)	0.394 (401.0)	0.29 (383.7)
B	0.208 (554.7)	0.237 (622.1)	0.292 (602.8)	0.363 (666.6)
C	0.201 (568.9)	0.247 (590.1)	0.314 (634.0)	0.347 (666.9)
w_2	0.64	0.65	0.7	0.75
A	0.180 (418.9)	0.111 (430.7)	0.001 (124.0)	0 (–)
B	0.404 (685.3)	0.450 (679.4)	0.510 (631.8)	0.494 (658.5)
C	0.416 (689.4)	0.439 (678.6)	0.489 (664.2)	0.506 (666.1)

Table 2 | Relative number of choices and mean response times (arbitrary unit, in parentheses) for alternatives A = (9, 1), B = (1, 9), and C = (1, 8.5) from 1000 simulations with $\theta = 20$, $\xi = 0.01$, $\lambda = 0.05$, $\mu = 0$, and $w_2 = 1 - w_1$ ranging from 0.4 to 0.75 as indicated in the first row.

w_2	0.4	0.45	0.5	0.55
A	1 (93.5)	1 (118.4)	0.987 (168.7)	0.81 (204.7)
B	0 (–)	0 (–)	0.012 (593.3)	0.146 (523.4)
C	0 (–)	0 (–)	0.001 (499.0)	0.044 (601.7)
w_2	0.6	0.61	0.62	0.63
A	0.557 (313.1)	0.461 (343.6)	0.384 (391.2)	0.248 (396.5)
B	0.357 (466.2)	0.447 (476.1)	0.513 (473.6)	0.655 (487.3)
C	0.086 (558.7)	0.092 (568.1)	0.103 (569.2)	0.097 (605.3)
w_2	0.64	0.65	0.7	0.75
A	0.196 (395.5)	0.101 (342.5)	0.002 (155.5)	0 (–)
B	0.699 (448.9)	0.810 (414.6)	0.932 (237.2)	0.953 (159.5)
C	0.105 (483.7)	0.089 (459.0)	0.066 (243.7)	0.047 (174.6)

that is similar to one of these but slightly less attractive. For $A = (9, 1)$, $B = (1, 9)$, and $C = (1, 8.5)$ this yields

$$L = \begin{pmatrix} 9 & 1 \\ 1 & 9 \\ 1 & 8.5 \end{pmatrix} \quad \text{and} \quad M = \begin{pmatrix} 0.818 & 0.054 \\ 0.091 & 0.487 \\ 0.091 & 0.459 \end{pmatrix}.$$

As shown in **Figure 6**, the attraction effect occurs between $W = \begin{pmatrix} 0.39 \\ 0.61 \end{pmatrix}$ and $W = \begin{pmatrix} 0.35 \\ 0.65 \end{pmatrix}$ Note that the deviation from weights $w_1 = w_2 = 0.5$ is due to a higher salience of attribute two because the values on this attribute differentiate between the alternatives. The relative frequency of choices for alternatives *A*, *B*, and *C* including the mean number of steps leading to these choices can be found in **Table 2**.

For the compromise effect (Simonson, 1989), two equally attractive but dissimilar alternatives $A = (9, 1)$ and $B = (1, 9)$ compete against a compromise option $C = (5, 5)$. Note that the defined values for each alternative sum up to ten and thus all three alternatives objectively are equally attractive provided the attributes are equally weighted. We get

$$L = \begin{pmatrix} 9 & 1 \\ 1 & 9 \\ 5 & 5 \end{pmatrix} \quad \text{and} \quad M = \begin{pmatrix} 0.6 & 0.067 \\ 0.067 & 0.6 \\ 0.333 & 0.333 \end{pmatrix},$$

and restricting $w_1 = w_2 = 0.5$ to be equal, this yields

$$P = \begin{pmatrix} 0.4 \\ 0.4 \\ 0.2667 \\ 0.4 \\ 0.4 \\ 0.2667 \end{pmatrix}.$$

So far, these transition probabilities do not seem to induce any compromise effect but as the probabilities for sampling negative information are comparatively high, withdrawal of one alternative from the choice set frequently occurs in that setting. After withdrawal of one alternative, comparison of the remaining alternatives is renewed. In the cases where alternative A or B are withdrawn, the new probabilities clearly favor the compromise option C, yielding an overall preference for that alternative:

$$M_{-A/-B} = \begin{pmatrix} 0.643 & 0.167 \\ 0.357 & 0.833 \end{pmatrix}, \; P_{-A/-B} = \begin{pmatrix} 0.15 \\ 0.35 \\ 0.35 \\ 0.15 \end{pmatrix}.$$

In 1000 trials with decision threshold $\theta = 20$, noise factor $\xi = 0.01$, leakage factor $\lambda = 0.05$, and no inhibition, alternatives A and B were chosen 247 (24.7%) and 250 (25%) times, respectively, and option C won 503 (50.3%) decisions. Decreasing θ to 10 yields choice frequencies of 243 (24.3%) for alternative A, 267 (26.7%) for option B, and 490 (49%) for alternative C. $\theta = 5$ leads to 253 (36.9%) choices with an average step number of 36.9 for alternative A, 269 (26.9%) choices with 38.3 steps on average for option B, and 478 (47.8%) choices with 43.8 steps on average for alternative C. **Figure 7** shows the response time distribution for alternative A for $\theta = 5$, $\xi = 0.01$, $\lambda = 0.05$, and $\mu = 0$. The expected response time, i.e., the mean of the distribution is 36.6. The magnitude of the compromise effect can be influenced by application of an asymmetric value function after comparison of alternatives.

4. COMPARISON WITH OTHER MODELS

Multi-alternative DFT (Roe et al., 2001) and the LCA model (Usher and McClelland, 2004) both account for similarity, attraction, and compromise effects in three-alternative preferential choice and thus build the theoretical background for the 2N-ary choice tree model. Nevertheless there are some important differences and the first one to set the new model apart from the previous approaches is the attribute-wise normalization of the initially provided evaluations of alternatives. This preprocessing of input values makes them comparable over attributes. Effects that originate from differing orders of magnitude of the input values can thus be controlled by influencing the attention weights for the attributes. The comparison of alternatives on single attributes is basically the same in all three models but only the LCA model and the 2N-ary choice tree model allow for external reference points that are not present in the choice set to influence the resulting values. Application of an asymmetric value function allows the LCA model to implement the loss-aversion principle (Kahneman and Tversky, 1979) and addition of a positive constant avoids negative activations and thus negated inhibition which was crucial for some of the results of multi-alternative DFT. Both concepts (asymmetric value function and positive constant) can be implemented into the 2N-ary choice tree model as well but do not affect its ability to account for the aforementioned effects (except for the magnitude of the compromise effect). Whereas all three models use leakage to account for decay of already sampled information over time and have a random part that implements noise in human decision making, inhibition is another crucial difference between them. In multi-alternative DFT, local inhibition explains the attraction and compromise effect, the LCA model uses global inhibition to account for the similarity effect. Both types of inhibition can be implemented in the 2N-ary choice tree model but are not necessary for explanation of the three effects.

Beside some similarities and dissimilarities between the models, in particular with respect to some underlying psychological concepts the 2N-ary choice tree model is the first to provide expected choice probabilities and response times in closed form and thus allows for convenient estimation of the model parameters from the observed choice times and frequencies in experimental settings. Furthermore, it can be extended to more than three-choice alternatives in a straightforward way to account for choice behavior in more complex, and possibly more realistic choice situations.

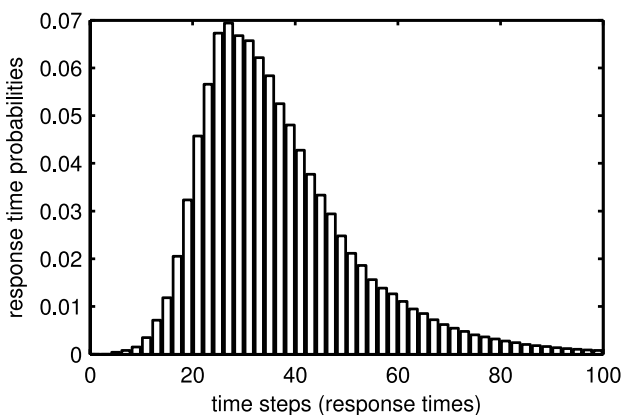


FIGURE 7 | Response time distribution for alternative $A = (9, 1)$ in the compromise setup with $\theta = 5$, $\xi = 0.01$, $\lambda = 0.05$, and $\mu = 0$. The expected response time, i.e., the mean of the distribution is 36.6.

5. CONCLUDING REMARKS

The 2N-ary choice tree model provides alternative explanations for the similarity, attraction, and compromise effect that can be experimentally tested as suggested before. Especially the manipulation of attention weights is of interest, because it differentiates the model on hand from former approaches and should allow to experimentally produce similarity and attraction effects which has been proven to be difficult in the past. One problem, however, we are currently encountering is limited machine accuracy which leads to accumulation of rounding errors during calculation of expected choice probabilities and response times.

6. FORMAL STATEMENT AND PROOF

We can approximate the expected choice probabilities and hitting times up to absolute accuracy in finite time. This follows from Theorem 1:

Theorem 1: Each random walk Y_n on the above defined tree $T = (V, E, r)$ with transition probabilities p_e, ends in finite time with probability one.

Corollary 1: With probability one only finitely many addends in equations (1) and (2) are unequal zero.

Considering expected values, i.e., limits of infinite sums, it is helpful to make use of a concept that allows for propositions about asymptotic behavior. For each alternative, the difference of the two counters that are associated with this alternative resemble a birth-death chain:

Definition 4 (birth-death chain): A sequence of random variables $X_1, X_2, \ldots$ with values in a countable state space $\mathscr{S} \equiv \{0, 1, 2, \ldots\} \subseteq \mathbb{N}$ is called *Markov chain*, if it satisfies the Markov property

$$\mathbb{P}[X_{n+1} = x | X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n] = \mathbb{P}[X_{n+1} = x | X_n = x_n].$$

A Markov chain is called *time homogeneous*, if $\mathbb{P}[X_{n+1} = x | X_n = y] = p(x, y)$ for all n, i.e., the probability for going from x to y is independent from n. A *birth-death chain* is a time homogeneous Markov chain that does not skip any state. Its transition probabilities $p(x, y)$ are equal to zero for all $x, y \in \mathscr{S}$ with $|y - x| > 1$. A *non-homogeneous birth-death chain* is a birth-death chain that is not time homogeneous.

Durrett (2010) proves the following theorem for (non-homogeneous) birth-death chains as special case of Markov chains:

Theorem 2: Let X_n be a Markov chain and suppose

$$\mathbb{P}\left[\cup_{m=n+1}^{\infty} \{X_m \in B_m\} | X_n\right] \geq \delta > 0 \text{ on } \{X_n \in A_n\}.$$

Then $\mathbb{P}[\{X_n \in A_n \text{ infinitely often}\} - \{X_n \in B_n \text{ infinitely often}\}] = 0$.

For each alternative $i \in \{1, 2, \ldots, N\}$, the above mentioned difference of its two counters can be interpreted as non-homogeneous birth-death chain X_n with absorbing states θ_i^- and θ_i^+ and state space $\mathscr{S} = \{\theta_i^-, (\theta_i^- + 1), (\theta_i^- + 2), \ldots, (\theta_i^+ - 2), (\theta_i^+ - 1), \theta_i^+\}$. Its transition probabilities are

$$\left.\begin{array}{ll} p^n(x, x+1) & = p_x^n = p_i^n(x), \\ p^n(x, x-1) & = q_x^n = p_{N+i}^n(x), \\ p^n(x, x) & = r_x^n = 1 - p_x^n - q_x^n, \end{array}\right\} \quad \text{for } x \in \mathscr{S} - \left\{\theta_i^-, \theta_i^+\right\}$$

and

$$\left.\begin{array}{ll} p_x^n = q_x^n & = 0, \\ r_x^n & = 1, \end{array}\right\} \quad \text{for } x \in \left\{\theta_i^-, \theta_i^+\right\}.$$

Due to the noise in the transition probabilities, $p_x^n > 0$ and $q_x^n > 0$ for all $x \in \mathscr{S} - \{\theta_i^-, \theta_i^+\}$. It follows that the probability for walking the direct way from $x \in \mathscr{S} - \{\theta_i^-, \theta_i^+\}$ to either θ_i^- or θ_i^+ is

$$\delta_x := \left(\prod_{y=\theta_i^-+1}^{x} q_i^{n+x-y}(y) + \prod_{z=x}^{\theta_i^+-1} p_i^{n+z-x}(z) \right) > 0$$

and thus

$$\delta := \min_{\theta_i^- < x < \theta_i^+} \delta_x > 0.$$

Define $T_{\theta_i^-} = \inf\{n : X_n = \theta_i^-\}$, $T_{\theta_i^+} = \inf\{n : X_n = \theta_i^+\}$, $T_i = T_{\theta_i^-} \wedge T_{\theta_i^+}$, $A_n = \mathscr{S}$, and $B_n = \{\theta_i^-, \theta_i^+\}$. Then $\{X_m \in B_m\}$ is equivalent to $T_i \leq m$ and

$$\mathbb{P}\left[\cup_{m=n+1}^{\infty} \{X_m \in B_m\} | X_n\right] = 1 > \delta$$

for $X_n \in B_n$. The probability for walking from any $x \in \mathscr{S}$ to either θ_i^- or θ_i^+ on every possible way is

$$\mathbb{P}\left[\cup_{m=n+1}^{\infty} \{X_m \in B_m\} | X_n\right] \geq \delta > 0$$

and thus fulfills the assumptions of Theorem 2. It follows that

$$\mathbb{P}\left[\left\{X_n \in A_n \text{infinitely often}\right\} - \left\{X_n \in B_n \text{infinitely often}\right\}\right] = 0$$

which is equivalent to

$$\mathbb{P}\left[\left\{X_n \in A_n - B_n \text{ finitely often}\right\}\right] = \mathbb{P}[T_i < \infty] = 1$$

and as this is true for every alternative $i \in \{1, 2, \ldots, N\}, \mathbb{P}[T, \infty] = 1$ holds for $T := \min T_1, T_2, \ldots, T_N$. This proves Theorem 1.

REFERENCES

Busemeyer, J. R., and Townsend, J. T. (1992). Fundamental derivations from decision field theory. *Math. Soc. Sci.* 23, 255–282.

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, 432–455.

Diederich, A. (1995). Intersensory facilitation of reaction time: evaluation of counter and diffusion coactivation models. *J. Math. Psychol.* 39, 197–215.

Diederich, A. (1997). Dynamic stochastic models for decision making under time constraints. *J. Math. Psychol.* 41, 260–274.

Diederich, A. (2008). A further test of sequential-sampling models that account for payoff effects on response bias in perceptual decision tasks. *Percept. Psychophys.* 70, 229–256.

Diederich, A., and Busemeyer, J. R. (2006). Modeling the effects of payoff on response bias in a perceptual discrimination task: bound-change, drift-rate-change, or two-stage-processing hypothesis. *Percept. Psychophys.* 68, 194–207.

Durrett, R. (2010). *Probability: Theory and Examples*, 4th Edn. New York: Cambridge University Press.

Hick, W. E. (1952). On the rate of gain of information. *Q. J. Exp. Psychol.* 4, 11–26.

Huber, J., Payne, J. W., and Puto, C. (1982). Adding asymmetrically dominated alternatives: violations of regularity and the similarity hypothesis. *J. Consum. Res.* 9, 90–98.

Hyman, R. (1953). Stimulus information as a determinant of reaction time. *J. Exp. Psychol.* 45, 188–196.

Kahneman, D., and Tversky, A. (1979). Prospect theory: an analysis of decision under risk. *Econometrica* 47, 263–291.

Kahneman, D., and Tversky, A. (1991). Loss aversion in riskless choice: a reference-dependent model. *Q. J. Econ.* 106, 1039–1061.

Koehler, D. J., and Harvey, N. (2007). *Blackwell Handbook of Judgment and Decision Making*, 2nd Edn. Oxford: Blackwell Publishing.

Korte, B., and Vygen, J. (2002). *Combinatorial Optimization: Theory and Algorithms*, 2nd Edn. Berlin: Springer.

LaBerge, D. (1962). A recruitment theory of simple behavior. *Psychometrika* 27, 375–396.

LaBerge, D. (1994). Quantitative models of attention and response processes in shape identification tasks. *J. Math. Psychol.* 38, 198–243.

Laming, D. R. J. (1968). *Information Theory of Choice Reaction Times*. New York: Academic Press.

Link, S. W., and Heath, R. A. (1975). A sequential theory of psychological discrimination. *Psychometrika* 40, 77–105.

Luce, D. (1986). *Response Times*. New York: Oxford University Press.

Pike, A. R. (1966). Stochastic models of choice behaviour: response probabilities and latencies of finite Markov chain systems. *Br. J. Math. Stat. Psychol.* 19, 15–32.

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

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.

Simonson, I. (1989). Choice based on reasons: the case of attraction and compromise effects. *J. Consumer Res.* 16, 158–174.

Smith, P. L., and Van Zandt, T. (2000). Time-dependent Poisson counter models of response latency in simple judgment. *Br. J. Math. Stat. Psychol.* 53, 293–315.

Townsend, J. T., and Ashby, F. G. (1983). *The Stochastic Modeling of Elementary Psychological Processes*. London: Cambridge University Press.

Tversky, A. (1972). Choice by elimination. *J. Math. Psychol.* 9, 341–367.

Tversky, A., and Simonson, I. (1993). Context-dependent preferences. *Manage. Sci.* 39, 1179–1189.

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

Usher, M., and McClelland, J. L. (2004). Loss aversion and inhibition in dynamical models of multialternative choice. *Psychol. Rev.* 111, 757–769.

Van Zandt, T., Colonius, H., and Proctor, R. W. (2000). A comparison of two response time models applied to perceptual matching. *Psychon. Bull. Rev.* 7, 208–256.

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 February 2012; accepted: 23 May 2012; published online: 20 June 2012.

Citation: Wollschläger LM and Diederich A (2012) The 2N-ary choice tree model for N-alternative preferential choice. Front. Psychology ***3****:189. doi: 10.3389/fpsyg.2012.00189*

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

frontiers in PSYCHOLOGY

HYPOTHESIS AND THEORY ARTICLE
published: 14 May 2012
doi: 10.3389/fpsyg.2012.00138

A quantum probability model of causal reasoning

Jennifer S. Trueblood [1] **** and Jerome R. Busemeyer*** [1,2]

[1] *Cognitive Science Program, Indiana University, Bloomington, IN, USA*
[2] *Department of Psychological and Brain Sciences, Indiana University, Bloomington, IN, USA*

Edited by:
David Albert Lagnado, University College London, London

Reviewed by:
Marius Usher, Tel-Aviv University, Israel
Philip M. Fernbach, University of Colorado Leeds School of Business, USA

****Correspondence:***
Jennifer S. Trueblood, Department of Psychological and Brain Sciences, Indiana University, 1101 East 10th Street, Bloomington, IN 47405, USA.
e-mail: jstruebl@indiana.edu

People can often outperform statistical methods and machine learning algorithms in situations that involve making inferences about the relationship between causes and effects. While people are remarkably good at causal reasoning in many situations, there are several instances where they deviate from expected responses. This paper examines three situations where judgments related to causal inference problems produce unexpected results and describes a quantum inference model based on the axiomatic principles of quantum probability theory that can explain these effects. Two of the three phenomena arise from the comparison of predictive judgments (i.e., the conditional probability of an effect given a cause) with diagnostic judgments (i.e., the conditional probability of a cause given an effect). The third phenomenon is a new finding examining order effects in predictive causal judgments. The quantum inference model uses the notion of incompatibility among different causes to account for all three phenomena. Psychologically, the model assumes that individuals adopt different points of view when thinking about different causes. The model provides good fits to the data and offers a coherent account for all three causal reasoning effects thus proving to be a viable new candidate for modeling human judgment.

Keywords: causal reasoning, quantum theory, order effects

1. INTRODUCTION

People can perform remarkably well at causal reasoning tasks that prove to be extremely difficult for statistical methods and machine learning algorithms. For example, Gopnik et al. (2001) demonstrated that individuals can infer causal relationships even when sample sizes are too small for statistical tests. Further, people can infer hidden causal structures that are difficult for computer scientists or statisticians to uncover (Kushnir et al., 2003). Even though people can infer rich causal representations of the world based on limited data, human causal reasoning is not infallible. Like many other types of subjective probability judgments, judgments about causal events often deviate from the normative rules of classic probability theory. This paper describes a quantum inference model previously developed in Trueblood and Busemeyer (2011) and demonstrates how it can account for judgment phenomena in causal reasoning problems.

The quantum inference model provides a general framework for updating probabilities about a hypothesis given a sequence of information, and it was first developed to account for order effects. One of the oldest and most reliable findings regarding human inference is that the order in which evidence is presented affects the final inference (Hogarth and Einhorn, 1992). For example, a juror's belief that a criminal suspect is guilty might depend on the order of presentation of the prosecution and defense. More generally, an order effect occurs when a judgment about the probability of a hypothesis given a sequence of information A followed by B, does not equal the probability of the same hypothesis when the given information is reversed, B followed by A. Because of the commutative nature of events in classical probability, order effects are difficult to explain using Bayesian models. Classical probability theory requires $p(A \cap B \mid H) = p(B \cap A \mid H)$ which according to Bayes rule implies $p(H \mid A \cap B) = p(H \mid B \cap A)$ (Trueblood and Busemeyer, 2011).

The quantum inference model is based on the axiomatic principles of quantum probability theory. This theory is a generalized approach to probability that relaxes some of the axioms or assumptions of standard probability theory in order to account for violations of the latter. Quantum probability theory is one of many generalized approaches to probability. Specifically, quantum probability theory is a geometric approach using subspaces and projections. Other generalized probability theories include Dempster-Shafer belief function theory (Fagin and Halpern, 1991) and intuitionist probability theory (Narens, 2003).

Models constructed from quantum probability theory do not make assumptions about biological substrates. Rather, quantum probability models provide an alternative mathematical approach for generating theories of how an observer processes information. The quantum approach has been used to account for a number of phenomena in cognitive science including violations of rational decision-making principles (Pothos and Busemeyer, 2009), conjunction and disjunction fallacies (Busemeyer et al., 2011), paradoxes of conceptual combination (Aerts, 2009), bistable perception (Atmanspacher et al., 2004), and interference effects in perception (Conte et al., 2009).

There are at least four reasons for considering a quantum approach to human judgments: (1) human judgment is not a simple read out from a pre-existing or recorded state, instead it is constructed from the current context and question. Quantum probability theory postulates that an individual's belief state is undetermined before measurement, and it is the process of

imposing measurements that forces a resolution of the indeterminacy. (2) Before measurement, cognition behaves more like a wave than a particle allowing for individuals to feel a sense of ambiguity about different belief states simultaneously. According to quantum probability theory, beliefs remain in a superimposed state until a final decision must be reached, which resolves the uncertainty and produces a collapse of the wave to a specific position like a particle. (3) Changes in context produced by one judgment can affect later judgments. Quantum probability theory captures this phenomenon though the notion of incompatibility allowing for one event to disturb and generate uncertainty about another. (4) Cognitive logic does not necessarily obey the rules of classic logic such as the commutative and distributive axioms. Quantum logic is more generalized than classic logic and can model human judgments that do not obey Boolean logic.

1.1. QUANTUM INFERENCE MODEL

The quantum inference model was first developed to account for order effects in a number of different problems including medical diagnostic tasks and jury decision-making problems. The following example of a medical inference task (Bergus et al., 1998) is one of the problems accounted for by the model. Physicians (N = 315) were initially informed about a particular woman's health complaint, and they were asked to estimate the likelihood that she had an infection on the basis of (a) her medical history and findings of the physical exam and (b) laboratory test results, presented in different orders. For one order, the physicians' initial estimate started out at 0.67; after they had seen the patient's history and findings of the physical exam, the estimate increased to 0.78; and then after they had also seen the lab test results, it decreased to 0.51. For the other order, the initial estimate again started at 0.67; after they had seen the lab test results, the estimate decreased to 0.44; and then after they also had seen the history and findings of the physical exam, it increased to 0.59. Because the final judgments were significantly different (0.51 versus 0.59, p = 0.03), an order effect is said to have occurred. Specifically, this type of order effect is called a recency effect, because the same evidence had a larger effect when it appeared at the end as opposed to the beginning of the sequence.

The quantum inference model uses the concept of incompatibility to account for order effects. The concept of compatibility is one of the most important new ideas introduced to cognitive science by quantum probability theory. Specifically, the model assumes that different pieces of information can be incompatible in the following sense: The set of feature patterns used to evaluate one piece of information is not shared by the set used to think of another so that no common set of features can be used to evaluate both pieces of information. For example, a physician needs to use knowledge about one set of features concerning a patient's history and physical exam, and the physician needs to use knowledge about another set of features concerning laboratory tests, but knowledge about all of the combinations from these two sets is not accessible to the physician. Psychologically, this corresponds to adopting different perspectives when evaluating different pieces of information. For example, in the medical inference task, we assume a physician has different representations for beliefs depending on three different points of view: a point of view determined by the woman's initial health complaint, a point of view determined by the medical history and findings of the physical exam, and a point of view determined by the laboratory test results.

The quantum model is able to account for the medical inference data (Bergus et al., 1998) and also a similar type of data from the domain of jury decision-making (Trueblood and Busemeyer, 2011). In these experiments, subjects read fictitious criminal cases and made a sequence of three judgments for each case: one before the presentation of any evidence, and two more judgments after presentations of evidence by a prosecutor and a defense. For a random half of the cases, the prosecution was presented before the defense, and for the other half, the defense was presented first.

In one version of the experiment (N = 291) the strength of the prosecution and defense was also manipulated. For example, subjects might be asked to judge the probability that a defendant was guilty based on a weak prosecution and a strong defense. Combining the order manipulation with two levels of strength (strong and weak) allowed for eight different order conditions, and as far as we know, this is the largest existing study of order effects on inference. Because of the many different conditions, this experiment provided a rich testing ground for the quantum inference model. Specifically, we compared the quantum inference model to two previously proposed models of order effects from the judgment and decision-making literature. All of the models had the same number of parameters, and the quantum model provided the best fits.

Because the quantum model provides a general way to calculate probabilities in inference problems, it is natural to apply it to situations involving causal reasoning. We begin by describing two recently discovered phenomena in causal reasoning, and illustrate how the model can account for them. Then, we introduce an *a priori* prediction about order effects which we test and confirm through a new experimental study.

1.2. PREDICTIVE AND DIAGNOSTIC CAUSAL JUDGMENTS

There are two possible ways to frame a causal reasoning problem. As formalized by Fernbach et al. (2011), a predictive probability judgment is represented by the conditional probability $p(\text{Effect} \mid \text{Cause})$ and a diagnostic probability judgment is represented by the conditional probability $p(\text{Cause} \mid \text{Effect})$.

Fernbach et al. (2011) illustrate these two different framings through an example about the transmission of a drug-addiction between a mother and a child. More specifically, a predictive causal reasoning problem could be formulated as "A mother has a drug-addiction. How likely is it that her newborn baby has a drug-addiction?" and a diagnostic causal reasoning problem could be formulated as "A newborn baby has a drug-addiction. How likely is it that the baby's mother has a drug-addiction?"

One recently discovered finding arises from the comparison of predictive and diagnostic judgments when there are strong and weak alternative causes (Fernbach et al., 2011). The drug-addiction scenario described above is an example of a weak alternative causes scenario because there are few alternatives to a child being drug addicted when the mother is not. On the other hand, a strong alternative causes scenario might be one involving the transmission of dark skin from a mother to a child (Fernbach et al., 2011). In such

a scenario, a father with dark skin provides a strong alternative to a child having dark skin when the mother does not. The results of experiment 1 by Fernbach et al. (2011) show that subjects are sensitive to the strength of alternative causes when making probability judgments about diagnostic problems but not when making probability judgments about predictive problems. As expected, the probability judgments for diagnostic problems with strong alternative causes (e.g. "A newborn baby has dark skin. How likely is it that the baby's mother has dark skin?") are significantly lower than the probability judgments for diagnostic problems with weak alternative causes (e.g. "A newborn baby has a drug-addiction. How likely is it that the baby's mother has a drug-addiction?"). One might expect that predictive problems with strong alternative causes should produce higher probability judgments than predictive problems with weak alternative causes because alternative causes increase the likelihood that the effect was brought about by different mechanisms (Fernbach et al., 2011). However, the experimental data shows no significant difference between probability judgments for the two types of predictive problems.

A second finding arises from the comparison of predictive and diagnostic judgments in cases where there are full conditionals and no-alternative conditionals (Fernbach et al., 2010). The term full conditional is used to describe situations in which alternative causes are implicit. For example, the following predictive question is a full conditional used in experiment 1 of Fernbach et al. (2010): "Ms. Y has depression. What is the likelihood she presents with lethargy?" The term no-alternative conditional is used to describe situations in which subjects are told that there are no-alternative causes. For example, a no-alternative conditional for the same depression problem might be "Ms. Y has depression. She has not been diagnosed with any other medical or psychiatric disorder that would cause lethargy. What is the likelihood she presents with lethargy?" One might expect that the following two inequalities should hold

$$p\,(\text{Effect}\,|\,\text{Cause}) > p\,(\text{Effect}\,|\,\text{Cause},\ \text{No Alternative Causes}) \tag{1}$$

$$p\,(\text{Cause}\,|\,\text{Effect}) < p\,(\text{Cause}\,|\,\text{Effect},\ \text{No Alternative Causes})\,. \tag{2}$$

The first inequality is expected because alternative causes should increase the likelihood of an effect. Even though alternative causes are not specifically mentioned in a full conditional, the alternative causes are still present. Thus, the full conditional should be judged as more likely than the no-alternative conditional in predictive problems. On the other hand, the second inequality is expected because alternative causes compete to explain an effect. Thus, the full conditional should be judged as less likely than the no-alternative conditional in diagnostic problems. Experimental results from (Fernbach et al., 2010) show that the probability judgments of subjects obey the second inequality relating to diagnostic reasoning problems but do not obey the first inequality relating to predictive reasoning problems. In the predictive reasoning scenarios, subjects show no significant difference between their probability judgments in full conditional and no-alternative conditional problems.

The two judgment phenomena described here can both be explained by the quantum inference model. Next, we describe the model in the framework of causal reasoning and demonstrate how it can account for the two findings.

2. THE QUANTUM INFERENCE MODEL OF CAUSAL REASONING

The quantum inference model has been adopted for causal reasoning problems because it provides a general way for updating probabilities about a hypothesis (e.g., the presence of an effect) given a set of information (e.g., different causes for the effect). The quantum model is not at odds with the causal model view set forth by Fernbach et al. (2011) which posits that individuals adopt a representation that approximates the structure of a system and probability judgments arise from this representation. Fernbach et al. (2011) formalize this idea using a causal Bayes net. While the quantum model provides a new way for calculating probabilities, quantum causal graphs can be constructed in a similar manner to causal Bayes nets and could potentially be used as a way to formalize the specific representation used by individuals.

For all of the applications discussed in this paper, the model assumes there is a single effect which can exist (e) or not exist ($\bar{e}$) and one or more causes which are either present (p) or absent (a). Based on this assumption there are four possible elementary events that could occur when considering a single effect and a single cause: the effect exists and the cause is present, the effect exists and the cause is absent, the effect does not exist and the cause is present, and the effect does not exist and the cause is absent. In quantum probability theory, the sample space used in classical probability theory is replaced by a Hilbert space (i.e., a complex number vector space). In our framework, the four elementary events are used to define an orthonormal basis for a four dimensional vector space V:

$$V = \text{span}\left\{\left|e \wedge p\right\rangle, \left|e \wedge a\right\rangle, \left|\bar{e} \wedge p\right\rangle, \left|\bar{e} \wedge a\right\rangle\right\}. \tag{3}$$

Quantum probability postulates the existence of a unit length state vector $|\psi\rangle \in V$ representing an individual's state of belief[1]. The belief state $|\psi\rangle$ can be expressed as a linear combination or superposition of the four basis states:

$$|\psi\rangle = \omega_{e,p} \cdot \left|e \wedge p\right\rangle + \omega_{e,a} \cdot \left|e \wedge a\right\rangle + \omega_{\bar{e},p} \cdot \left|\bar{e} \wedge p\right\rangle + \omega_{\bar{e},a} \cdot \left|\bar{e} \wedge a\right\rangle. \tag{4}$$

The weights such as $\omega_{e,p}$ are called probability amplitudes and determine the belief about a particular elementary event such as $e \wedge p$. The belief state vector can be represented by the four amplitudes when the basis for V is treated as the standard basis

[1]The use of Dirac, or Bra-ket, notation is in keeping with the standard notation used in quantum mechanics. For the purposes of this paper, $|\psi\rangle$ corresponds to a column vector whereas $\langle\psi|$ corresponds to a row vector. Following the convention in physics, we use ψ to denote amplitudes which are basis dependent and $|\psi\rangle$ to denote an abstract vector which is coordinate free.

for $\mathcal{C}^4$:

$$\psi = \begin{bmatrix} \omega_{e,p} \\ \omega_{e,a} \\ \omega_{\bar{e},p} \\ \omega_{\bar{e},a} \end{bmatrix} = \begin{bmatrix} \text{amplitude for effect exists and cause is present} \\ \text{amplitude for effect exists and cause is absent} \\ \text{amplitude for effect does not exist and cause is present} \\ \text{amplitude for effect does not exist and cause is absent} \end{bmatrix}. \quad (5)$$

Quantum events are defined geometrically as subspaces (e.g., a line or a plane) within this four dimensional space. For example, the event corresponding to the "effect exists" is defined as the subspace $L_e = \text{span}\{|e \wedge p\rangle, |e \wedge a\rangle\}$. Quantum probabilities are computed by projecting $|\psi\rangle$ onto subspaces representing events. Projectors for general events are defined in terms of the projectors for elementary events. For example, the projectors for the elementary events $e \wedge p$ and $e \wedge a$ are

$$P\left(e,p\right) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, P\left(e,a\right) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad (6)$$

and the projector for the event the "effect exists" corresponds to the sum of the two projectors: $P(e) = P(e,p) + P(e,a)$. To calculate the probability of this event, the state vector $|\psi\rangle$ is projected onto L_e by the projector P_e:

$$P\left(e\right)\psi = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \omega_{e,p} \\ \omega_{e,a} \\ \omega_{\bar{e},p} \\ \omega_{\bar{e},a} \end{bmatrix} = \begin{bmatrix} \omega_{e,p} \\ \omega_{e,a} \\ 0 \\ 0 \end{bmatrix}. \quad (7)$$

The probability of the event L_e is equal to the squared length of this projection:

$$p\left(L_e\right) = \left\|P\left(e\right)\psi\right\|^2 = \left\|\omega_{e,p}\right\|^2 + \left\|\omega_{e,a}\right\|^2. \quad (8)$$

One of the important differences between quantum probability theory and classical probability theory occurs when multiple events are considered. When multiple events are involved, quantum theory allows for these events to be incompatible. Intuitively, compatibility means that two events X and Y can be accessed simultaneously without interfering with each other. On the other hand, if X and Y are incompatible, they cannot be accessed simultaneously. From a cognitive standpoint, this implies that the two events are processed serially and one interferes with the other. Mathematically, at set of incompatible elementary events are represented by different bases for the same vector subspace. In the case of more general events, consider the event X represented by the subspace L_x with basis $|x_1\rangle, \ldots, |x_n\rangle$ and the event Y represented by the subspace L_y with basis $|y_1\rangle, \ldots, |y_n\rangle$. If the two events are incompatible, then the $|x_i\rangle$ basis is a unitary transformation of the $|y_i\rangle$ basis. If X and Y are compatible, then there is one basis representation for both events. In this case, quantum probability theory reduces to classic probability theory.

For the purposes of this paper, we assume that the effect is compatible with the causes and multiple causes are incompatible with each other. To formalize this notion, consider a single effect and a single cause (for clarity, call this cause "cause 1"). The basis defined in equation 3 can be used to represent beliefs about the effect and "cause 1." Now, suppose the same effect is considered in terms of a different cause (call this cause "cause 2"). Because "cause 1" and "cause 2" are incompatible, the four basis elements defined above for "cause 1" cannot be used to describe the relationship between the effect and "cause 2." This is because incompatible events are represented mathematically by different bases for the same vector space. Thus, a unitary transformation U is applied to the "cause 1" basis to "rotate" it to the "cause 2" basis. The transformation must be unitary to preserve the orthonormal nature of the basis elements. The result of the unitary transformation is a new set of basis elements for V that represents an individual's point of view associated with "cause 2":

$$V = \text{span}\left\{U\,|e \wedge p\rangle,\ U\,|e \wedge a\rangle, U\,|\bar{e} \wedge p\rangle, U\,|\bar{e} \wedge a\rangle\right\}. \quad (9)$$

As a point of comparison, a classical probability model for a single effect and two causes would use an eight dimensional sample space because there are two outcomes (e or $\bar{e}$) for the effect and two outcomes (p or a) for each cause. By allowing the causes to be incompatible, the eight dimensional space needed for the classical model is reduced to a four dimensional space in the quantum model. This reduction in dimension becomes even more dramatic when a single effect and n different causes are considered. In this case, the dimension of the sample space for the classical model would be 2^{n+1} whereas the quantum model with n incompatible causes continues to use only four dimensions. The vector space V of the quantum model remains four dimensional because the n different causes are accounted for by n different bases for V rather than an increase in the dimension of V[2]. Psychologically, the n different bases correspond to different points of view used when thinking about the existence of an effect and the presence or absence of a cause. Formally, there exists a set of unitary operators used to transform one set of basis vectors to another. This is analogous to rotating the axes in multidimensional scaling (Shepard, 1962; Carroll and Chang, 1970) or multivariate signal detection theory (Rotello et al., 2004; Lu and Dosher, 2008).

In the model, unitary transformations correspond to an individual's shifts in perspective and relate one point of view (i.e., basis) to another. So far, incompatible events have been described as defining different basis for V. An equivalent way of viewing incompatible events is to fix a basis for V such as the basis given in equation 3 and to transform the state vector $|\psi\rangle$ by a unitary operator whenever an incompatible event is being considered. In

[2]It should be noted, that the quantum inference model is not restricted to assuming that all causes are incompatible. If there are n causes for an effect, then it is possible to allow some of the causes to be compatible and others to be incompatible. In this case, the dimension of the vector space V would be between four and 2^{n+1}. Further, if there were multiple effects, the model could be extended to allow for incompatibility among effects. Because the current paper only considers a single effect and two possible causes, these modifications were not necessary.

other words, one can either "rotate" the vector space and leave the state vector fixed or one can "rotate" the state vector and leave the space fixed. In the applications below, the belief state is "rotated" and the basis is fixed.

2.1. CONSTRUCTION OF UNITARY MATRICES

Any unitary matrix can be constructed from the matrix exponential function $U = e^{-i\phi H}$ where H is a Hermitian matrix. (A Hermitian matrix is equal to its own conjugate transpose.) Thus, to construct the unitary operators for the quantum inference model, a Hermitian matrix H first needs to be defined. Following Trueblood and Busemeyer (2011), it is assumed that H is constructed from two components, $H = H_1 + H_2$.

We begin by describing the construction of a Hermitian matrix for a simple two dimensional problem and extend this to define H_1. Suppose that we have a vector space spanned by two basis vectors $|e \wedge p\rangle$ and $|e \wedge a\rangle$. In our original vector space, this is the subspace corresponding to the event the "effect exists." Also, we assume that this new space can be viewed from different perspectives and define the unitary matrix U_j to transform one perspective into another. This unitary matrix is constructed from a two dimensional Hermitian matrix W.

Any two dimensional Hermitian matrix can be described as a linear combination of the Pauli matrices:

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}; \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}; \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \tag{10}$$

We let W be defined as

$$W = \alpha_x \cdot \sigma_x + \alpha_y \cdot \sigma_y + \alpha_z \cdot \sigma_z. \tag{11}$$

Now, we can write the corresponding unitary matrix as

$$U_j = e^{-i\phi_j W} = e^{-i\phi_j(\alpha_x \cdot \sigma_x + \alpha_y \cdot \sigma_y + \alpha_z \cdot \sigma_z)}. \tag{12}$$

where we assume that $(\alpha_x^2 + \alpha_y^2 + \alpha_z^2)^{\frac{1}{2}} = 1$. By applying Euler's formula we can rewrite the unitary matrix as

$$U_j = \cos(\phi_j) \cdot I - i\sin(\phi_j) \cdot (\alpha_x \cdot \sigma_x + \alpha_y \cdot \sigma_y + \alpha_z \cdot \sigma_z) \tag{13}$$

where I is the 2×2 identity matrix. (Euler's formula states that $e^i\phi = \cos(\phi) + i\sin(\phi)$.) Equation 13 can be written as the matrix

$$U_j = \begin{bmatrix} \cos(\phi_j) - i \cdot \alpha_z \cdot \sin(\phi_j) & -(i \cdot \alpha_x + \alpha_y) \cdot \sin(\phi_j) \\ -(i \cdot \alpha_x - \alpha_y) \cdot \sin(\phi_j) & \cos(\phi_j) + i \cdot \alpha_z \cdot \sin(\phi_j) \end{bmatrix}. \tag{14}$$

From equation 14, the unitary matrix U_j produces a rotation of degree ϕ_j around the unit length vector $(\alpha_x, \alpha_y, \alpha_z,)$. (Please see Sakurai, 1994 for more details.) The Hermitian matrix W is said to be a generator for U_j because for small values of ϕ_j, the unitary matrix is approximately equal to $1 - i\phi_j W$. (Please see Nielsen and Chuang, 2000, Chapter 4 for more details.)

After applying the matrix U_j, the probability that $e \wedge p$ is true is periodic in the variable ϕ_j. If we want to ensure that p is favored throughout the presentation of the cause, then we must maintain a probability greater than 0.5 for p over a. In the model, the probability for p over a is maximized whenever $\alpha_y = 0$ and $\alpha_x = \alpha_z > 0$. By setting $\alpha_y = 0$, we avoid reversing the preference for p across time. The condition that $\alpha_x = \alpha_z > 0$ restricts probabilities for p to oscillate back and forth from 0.5 to 1.0 across time. Because the vector $(\alpha_x, \alpha_y, \alpha_z)$ has unit length, we must set $\alpha_x = \alpha_z = \frac{1}{\sqrt{2}}$. Now, define W as

$$W = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \tag{15}$$

and the 2×2 unitary matrix as

$$U_j = \exp\left\{-i\phi_j \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\right\}. \tag{16}$$

In the full four dimensional model, we specify the matrix H_1 in terms of the matrix W. Specifically, we assume that H_1 is the tensor product given by

$$H_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \otimes W = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \otimes \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 \end{bmatrix}. \tag{17}$$

A unitary matrix with H_1 as a generator transforms the amplitudes toward the presence of causes by rotating the probability amplitudes to favor events involving the "cause is present." In other words, the corresponding unitary matrix strengthens the amplitudes corresponding to p and weakens the amplitudes corresponding to a. Further, the unitary matrix corresponding to H_1 strengthens and weakens the amplitudes for causes to the greatest extent possible. This results from the fact that the matrix W was designed to maximize the probability of one type of information over another.

Next, we turn to the construction of the H_2 component of the Hermitian matrix H. As with H_1, we begin by defining a Hermitian matrix for a two dimensional space and then extend this to the four dimensional case. Consider the vector space spanned by the basis vectors $|e \wedge p\rangle$ and $|\bar{e} \wedge p\rangle$. This is the subspace of the full four dimensional vector space corresponding to the presence of a cause. We proceed exactly as before and define a Hermitian matrix V as a linear combination of Pauli matrices. Because we wish to maintain an overall probability greater than 0.5 for the existence or non-existence of an effect across time, we set $\alpha_y = 0$ and $\alpha_x = \alpha_z = \frac{1}{\sqrt{2}}$. Thus, we have $V = W$.

In order to easily write H_2 in terms of V, we rearrange the coordinate vector given in equation 5 so that

$$\psi = \begin{bmatrix} \omega_{e,p} \\ \omega_{\bar{e},p} \\ \omega_{\bar{e},a} \\ \omega_{e,a} \end{bmatrix}. \tag{18}$$

Now we have

$$H_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \otimes V = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \otimes \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$
$$= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 \end{bmatrix}. \quad (19)$$

Because we want to combine H_1 and H_2, we will need to use the same arrangement of coordinates for both matrices. To define H_2 in terms of the coordinates given in equation 5, we first switch row two with row three and column two with column three. Next, we switch row two with row four and column two with column four. The resulting matrix is

$$H_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}. \quad (20)$$

The Hermitian matrix H_2 evolves an individual's beliefs about an effect and the presence or absence of a cause. Specifically, it results in transforming amplitudes toward the event the "effect exists and cause is present" and toward the event the "effect does not exist and cause is absent." As in the case of H_1, the unitary matrix corresponding to H_2 evolves the amplitudes to the greatest extent possible.

Now, we define the Hermitian matrix H as

$$H = \frac{1}{\sqrt{2}} \left(\begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \right)$$
$$= \frac{1}{\sqrt{2}} \begin{bmatrix} 2 & 1 & 1 & 0 \\ 1 & -2 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}. \quad (21)$$

In the sum $H_1 + H_2$, the H_2 matrix affects the relation between causes and effects and the H_1 matrix biases the amplitudes toward the presence of causes. Both matrices are necessary components of H. The Hermitian matrix, H, was previously developed for psychological applications involving four dimensional vector spaces (Pothos and Busemeyer, 2009) and is identical to the one used in Trueblood and Busemeyer (2011). The parameter ϕ determines the degree of rotation and is used as a free parameter in the model. A different parameter value of ϕ is used for different causes. For more details about the derivation of the unitary operators, please see Pothos and Busemeyer (2009) and Trueblood and Busemeyer (2011).

3. MODELING THE PREDICTIVE AND DIAGNOSTIC PHENOMENA

Now that we have introduced the model, we illustrate how it can be applied to the two findings by Fernbach et al. (2010, 2011) concerning predictive and diagnostic judgments.

3.1. PREDICTIVE AND DIAGNOSTIC JUDGMENTS WITH STRONG AND WEAK ALTERNATIVE CAUSES

In experiment 1 conducted by Fernbach et al. (2011), 180 subjects provided probability judgments for predictive and diagnostic reasoning problems with strong and weak alternatives causes. In the experiment, twenty different question categories were used. These categories ranged from mothers and newborn babies to oxygen tanks and scuba divers. For each question category, there were two types of causes – one with strong alternatives and one with weak alternatives. In analyzing the data, Fernbach et al. (2011) averaged over the different categories. Because a large number of categories were used, any differences in the events themselves should average out.

To model data from this experiment, the quantum inference model assumes that equal weight is initially placed on the four elementary events defining the belief state in a manner similar to setting a uniform prior in a Bayesian model:

$$\psi_0 = \begin{bmatrix} \omega_{e,p} \\ \omega_{e,a} \\ \omega_{\bar{e},p} \\ \omega_{\bar{e},a} \end{bmatrix} = \begin{bmatrix} \sqrt{.25} \\ \sqrt{.25} \\ \sqrt{.25} \\ \sqrt{.25} \end{bmatrix}. \quad (22)$$

The original version of the quantum inference model (Trueblood and Busemeyer, 2011) was applied to inference problems involving a single hypothesis and two pieces of evidence. In this setting, it was assumed an individual adopted three different points of view throughout the inference problem: a point of view determined by the initial description of the problem, a point of view determined by the first piece of evidence, and a point of view determined by the second piece of evidence. A "rotation" of the belief state occurred whenever there was a shift in perspective. In an analogous manner, we assume here that there is a change in perspective (i.e., "rotation") between the initial point of view, the point of view associated with one of the causes, and the point of view associated with the other cause.

For predictive problems, the initial belief state is revised after an individual learns about the presence of a cause. Psychologically, the new information about the cause results in the individual shifting his or her perspective of the four elementary events. Mathematically, the initial belief state $|\psi_0\rangle$ changes to a new state by using a unitary operator to "rotate" the initial belief state: $U|\psi_0\rangle$. Because the individual learns the cause is present, the new state is then projected onto the "cause is present" subspace and is normalized to ensure that the length of the new belief state equals one:

$$|\psi_1\rangle = \frac{\left(P\left(e,p\right) + P\left(\bar{e},p\right)\right) U|\psi_0\rangle}{\left\| \left(P\left(e,p\right) + P\left(\bar{e},p\right)\right) U|\psi_0\rangle \right\|}. \quad (23)$$

The predictive probability is calculated by projecting the revised belief state onto the "effect exists" subspace and finding the squared length of the projection:

$$p\,(\text{Effect} \mid \text{Cause}) = \| P\,(e)\,|\psi_1\rangle \|^2. \quad (24)$$

For diagnostic problems, the initial belief state is revised after an individual learns the effect exists. In this case, the initial belief

state $|\psi_0\rangle$ does not need to be transformed by a unitary operator before it is projected onto the "effect exists" subspace. Because we are concerned with only a single effect, there is no need to change perspective between the initial belief state and the belief state associated with the knowledge that the effect is present[3]. In other words, the initial basis was chosen to describe a single effect being considered in the problem. Thus, the initial state is projected directly onto this subspace and is normalized resulting in a new belief state:

$$\psi_1 = \begin{bmatrix} \omega_{e,p} \\ \omega_{e,a} \\ \omega_{\overline{e},p} \\ \omega_{\overline{e},a} \end{bmatrix} = \begin{bmatrix} \sqrt{.5} \\ \sqrt{.5} \\ 0 \\ 0 \end{bmatrix}. \tag{25}$$

The diagnostic probability is calculated by projecting the revised belief state onto the "cause is present" subspace and finding the squared length of the projection. However, before projection, $|\psi_1\rangle$ is transformed by a unitary operator to account for the assumed incompatibility between the individual's current point of view and the point of view associated with the cause:

$$p\,(\text{Cause} \mid \text{Effect}) = \parallel P\,(c)\; U|\psi_1\rangle \parallel^2. \tag{26}$$

Because different causes are used in the weak and strong alternative causes problems (e.g. "A mother has a drug-addiction" versus "A mother has dark skin"), different parameter values of ϕ are used. Specifically, one value of ϕ is used to account for causes where the alternatives are weak and another value of ϕ is used to account for causes where the alternatives are strong. Equivalently, different causes are incompatible and thus different unitary operators are needed to "rotate" the state vector. All of the calculations presented here and for the other effects discussed below are also given in appendix B.

The important difference between predictive and diagnostic calculations is the ordering of projections and rotations. In the predictive case, the initial belief state is first rotated by the U matrix and then projected onto the "cause is present" subspace. In the diagnostic case, the initial state is first projected onto the "effect exists" subspace and then rotated by the U matrix. The model predicts that strong and weak alternative causes do not affect predictive judgments because the differences between these two situations, which are incorporated in the rotations, are wiped out by subsequent projections. This is not the case for diagnostic judgments because rotations occur after projections.

The model was fit to the mean judgments for the following four situations: predictive with weak alternative causes, predictive with strong alternative causes, diagnostic with weak alternative causes, and diagnostic with strong alternative causes. The model used two free parameters associated with the two different types of alternative causes (i.e., strong and weak) to model the four judgments. The model was fit by minimizing the sum of the squared error (SSE) between the experimental data and model predictions. The best fit parameters were $\phi_1 = -3.74$ for strong alternative causes and $\phi_2 = 0.48$ for the weak alternative causes. **Table 1** shows the experimental results and the best fitting model predictions. The mean squared error (MSE) for the model fit was less than 0.0005.

Next we show that the same quantum principles also account for the differences between predictive and diagnostic judgments in the more complex paradigm involving the no-alternative conditions.

3.2. PREDICTIVE AND DIAGNOSTIC JUDGMENTS WITH FULL AND NO-ALTERNATIVE CONDITIONALS

In experiment 1 conducted by Fernbach et al. (2010), 265 mental health practitioners provided probability judgments for predictive and diagnostic reasoning problems with full and no-alternative conditionals related to a scenario about a woman experiencing lethargy given she was diagnosed with depression.

To model this data, many of the same steps described above are used. Specifically, the probabilities for predictive and diagnostic reasoning problems with full conditionals are calculated in the exact same manner as above in equations 24 and 26 respectively.

To calculate the probabilities for predictive and diagnostic reasoning problems with no-alternative conditionals, it is assumed that an individual considers two causes when producing judgments. The first cause is the one explicitly given in the problem (i.e., the woman has been diagnosed with depression). The second cause is implicitly defined in the problem through the statement that there are no-alternative causes (i.e., the woman has not been diagnosed with any other medical or psychiatric disorders that cause lethargy). In keeping with the assumption that all causes are incompatible, these two causes are treated as such. Thus, two different unitary operators, U_1 and U_2 associated with the explicit present cause and the implicit absent cause respectively, are used when revising the belief state.

For predictive problems with the no-alternative conditional, the initial belief state given in equation 22 is first revised after an individual processes information about the presence of the explicit cause. The explicit cause is assumed to be processed first because it is more readily available. The initial state vector is updated according to equation 23 where U is defined as U_1. Next, the new state vector $|\psi_1\rangle$ is revised after the individual processes the information about the absence of the implicit cause. Because the two causes are incompatible, the current belief state $|\psi_1\rangle$ is changed to a new

Table 1 | Model fits for predictive and diagnostic judgments with strong and weak alternative causes.

Judgment type	Alternative strength			
	Weak		Strong	
	Data	Model	Data	Model
Diagnostic	0.817	0.803	0.585	0.561
Predictive	0.696	0.723	0.753	0.773

[3]If there was more than one effect under consideration, then it might be necessary to consider the effects as incompatible. In this case, there would be changes of perspective (i.e., "rotations") between the different effects and the initial belief state.

state by the unitary operator U_2. The new state is then projected onto the "cause is absent" subspace and normalized:

$$|\psi_2\rangle = \frac{(P(e,a) + P(\bar{e},a))\, U_2\, |\psi_1\rangle}{\| (P(e,a) + P(\bar{e},a))\, U_2\, |\psi_1\rangle \|}. \quad (27)$$

The predictive probability is then calculated by projecting the revised belief state onto to the "effect exists" subspace and finding the squared length of the projection:

$$p(\text{Effect} \mid \text{Cause, No Alternative Causes}) = \| P(e)\, |\psi_2\rangle \|^2. \quad (28)$$

For diagnostic problems with the no-alternative conditional, the initial belief state given in equation 22 is revised to the state given in equation 25 after an individual learns the effect exists. Next, the state undergoes revision when the individual considers information about the implicit cause being absent:

$$|\psi_2\rangle = \frac{(P(e,a) + P(\bar{e},a))\ \ U_2|\,\psi_1\rangle}{\| (P(e,a) + P(\bar{e},a))\ \ U_2|\,\psi_1\rangle \|}. \quad (29)$$

The diagnostic probability is calculated by projecting the belief state $|\psi_2\rangle$ onto to the "cause is present" subspace and finding the squared length of the projection. However, before projection, $|\psi_2\rangle$ is transformed by the unitary operator U_1 to account for the assumed incompatibility between the individual's current point of view and the point of view associated with the explicit cause:

$$p(\text{Cause} \mid \text{Effect, No Alternative Causes}) = \|P(c)\ U_1|\psi_2\rangle\|^2. \quad (30)$$

The model was fit to the mean judgments for the following four situations: predictive with full conditional, predictive with no-alternative conditional, diagnostic with full conditional, and diagnostic with no-alternative conditional. The model used two free parameters associated with the two unitary operators used for the two different types of causes (i.e., explicit and implicit). The model was fit by minimizing the sum of the squared error (SSE) between the experimental data and model predictions. The best fit parameters were $\phi_1 = -2.35$ for the explicit cause and $\phi_2 = -3.81$ for the implicit cause. **Table 2** shows the experimental results and the best fitting model predictions. The MSE for the model fit was less than 0.0003.

In summary, the quantum model uses the same principles to provide accurate fits to the results from both experiments. However, two parameters were used to fit four data points in each study. Obviously a stronger test of the assumptions underlying the quantum model is required before this account becomes very convincing.

Table 2 | Model fits for predictive and diagnostic judgments with full and no-alternative conditionals.

Judgment type	Conditional type			
	Full		No-alternative	
	Data	Model	Data	Model
Diagnostic	0.59	0.58	0.67	0.65
Predictive	0.69	0.67	0.68	0.69

4. ORDER EFFECTS IN CAUSAL REASONING

So far, the quantum inference model has been based on the assumption that causes are incompatible. This is the key assumption required to account for the findings. The current study was designed to gather experimental support for this assumption. If all events are compatible, then quantum probability theory reduces to classic probability theory. In particular, the events obey the commutative property of Boolean algebra. In a simple Bayesian inference model, the commutative nature of events implies order effects do not occur (Trueblood and Busemeyer, 2011). However, incompatible events do not have to obey the commutative property and can produce order effects. Thus, the quantum inference model with incompatible causes makes an *a priori* prediction that order effects exist in causal reasoning. The present study tests this prediction.

Subjects in the study were 113 undergraduate students at Indiana University who received experimental credit for introductory psychology courses. Each of the subjects completed a computer-controlled experiment where they read ten different randomized scenarios involving an effect and two causes with one the causes being present and the other cause being absent. For example, subjects might be asked about the likelihood that a high school cafeteria will serve healthier food next month (the effect) given the food budget remains the same (the absent cause) and a group of parents working to fight childhood obesity contacted the school about including healthier menu options (the present cause). All ten scenarios used in the experiment are given in appendix A.

The participants reported the likelihood of the effect on a 0–100 scale before reading either cause, after reading one of the causes, and again after reading the remaining cause. For a random half of the scenarios, subjects judged the present cause before the absent cause. For the remaining half of the scenarios, the subjects judged the absent cause before the present cause. The data was analyzed by collapsing across all ten scenarios. **Figure 1** shows the average probability judgments collapsed across the scenarios for the different orderings of the causes. A two sample t-test showed a significant recency effect ($t = 9.6408$, $df = 1128$, $p < 0.0001$). This implies that the second cause influenced subjects' beliefs more than the first cause. One might think that order effects are due to memory recall failures; however, memory recall is uncorrelated with order effects in sequential judgments (Hastie and Park, 1986).

The quantum model accounts for the order effect data in a manner similar to its account of predictive judgments with no-alternative conditionals. Specifically, there are two incompatible causes with one being present and the other being absent. Two different unitary operators, U_1 and U_2, are associated with the two causes respectively.

To start, the initial belief state is based on the probability judgments provided by subjects before either cause was presented. The mean probability of the effect given no causal information was

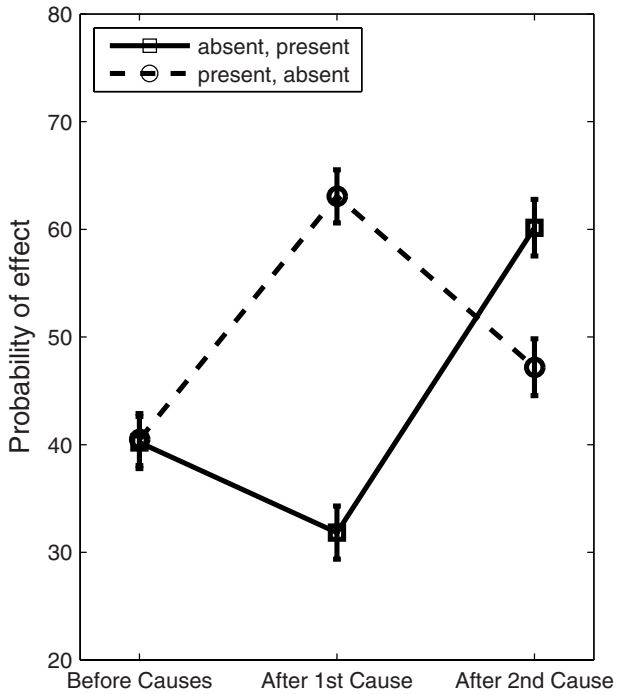


FIGURE 1 | Average probability judgments collapsed across 10 scenarios for two orderings of present and absent causes. The judgments exhibit a significant recency effect as illustrated by the crossing of the two curves on the graph. Error bars show the 95% confidence interval.

0.403. Thus, the initial state is defined as

$$\psi_0 = \begin{bmatrix} \omega_{e,p} \\ \omega_{e,a} \\ \omega_{\bar{e},p} \\ \omega_{\bar{e},a} \end{bmatrix} = \begin{bmatrix} \sqrt{\frac{0.403}{2}} \\ \sqrt{\frac{0.403}{2}} \\ \sqrt{\frac{0.597}{2}} \\ \sqrt{\frac{0.597}{2}} \end{bmatrix}. \quad (31)$$

When modeling judgments for the present cause followed by judgments for the absent cause, the initial belief state is first revised to accommodate the information about the present cause. Specifically, the initial state vector is updated according to equation 23 where U is defined as U_1. The probability of the effect given the present cause is calculated as in equation 24. Next, the new state vector $|\psi_1\rangle$ is revised after the individual processes the information about the absent cause. This updating occurs according to equation 27. The final probability of the effect given the present cause followed by the absent cause is calculated as in equation 28. To model the judgments for the reverse order of causes, absent cause followed by present cause, a similar set of steps are followed except that the roles of U_1 and U_2 were reversed (U_2 was applied first and U_1 second). It should be noted that the quantum model can produce both primacy and recency effects, and that these two effects oscillate across different values for the phi parameters.

The model was fit to the following four data points: p(Effect | Present Cause), p(Effect | Present Cause, Absent Cause), p(Effect | Absent Cause), p(Effect | Absent Cause, Present Cause). The model used two free parameters associated with the two unitary operators used for the two different types of causes (i.e., present and absent). The model was fit by minimizing the sum of the squared error (SSE) between the experimental data and model predictions. The best fit parameters were $\phi_1 = 3.67$ for the present cause and $\phi_2 = -1.57$ for the absent cause. **Table 3** shows the experimental results and the best fitting model predictions. The MSE for the model fit was less than 0.0002.

The existence of order effects provides support for the quantum inference model with incompatible causes. More importantly, the model has introduced a new direction for empirical study not considered before in causal reasoning.

5. ALTERNATIVE MODELS

Two other models of inference are worth mentioning. The first model is a causal Bayes net discussed in Fernbach et al. (2011). This model assumes that the relationship between causes and effects can be represented by a directed graph and probabilities are calculated from this structure. The second model is the belief-adjustment model developed by Hogarth and Einhorn (1992). This is an arithmetic model which assumes that beliefs are determined through an anchoring and adjustment process. While both models provide insights into the process of updating beliefs, neither model can provide an adequate account for all three causal reasoning phenomena.

5.1. CAUSAL BAYES NET

Fernbach et al. (2011) present a causal Bayes net as a possible account of predictive and diagnostic judgments with strong and weak alternative causes. In this model, the predictive probability of an effect given a cause is calculated using the noisy-or equation:

$$p\,(\text{Effect} \mid \text{Cause}) = W_c + W_a - W_c W_a \quad (32)$$

where $W_c = p$(Effect | Cause, No-Alternative Causes) is the causal power for the cause and $W_a = p$(Effect | No Causes) is the strength of alternative causes. The diagnostic probability of a cause given an effect is calculated by considering the complement:

$$p\,(\text{Cause} \mid \text{Effect}) = 1 - p\,(\text{No Cause}|\text{Effect}) \quad (33)$$

By applying Bayes' rule to the complement defined in equation 33, the diagnostic probability is given by

$$p\,(\text{Cause} \mid \text{Effect}) = 1 - (1 - P_c)\,\frac{W_a}{P_c W_c + W_a - P_c W_c W_a} \quad (34)$$

where $P_c = p$(Cause).

Table 3 | Model fits for order effects in predictive judgments.

Judgment order	After 1st judgment		After 2nd judgment	
	Data	Model	Data	Model
Present, absent	0.631	0.655	0.472	0.477
Absent, present	0.318	0.318	0.602	0.591

Fernbach et al. (2011) successfully applied this model to their data on predictive and diagnostic judgments with strong and weak alternative causes from experiment 1. However, they ultimately reject the model based on later experiments. Also, the model has not formally been applied to their findings with full and no-alternative conditionals (Fernbach et al., 2010). As such, it is unknown whether the model can provide a mathematical account of these data. Further, it is doubtful the model can account for order effects. Most Bayesian models have difficulty accounting for order effects due to the commutativity of events (Trueblood and Busemeyer, 2011). To model order effects, the model would need to introduce presentation order as another piece of information. In most experimental studies of order effects, order of presentation is randomly determined. Thus, order information is often irrelevant.

5.2. BELIEF-ADJUSTMENT MODEL

The second model worth noting is the belief-adjustment model originally developed to account for order effects (Hogarth and Einhorn, 1992). This model assumes that individuals update beliefs through a series of anchoring and adjustment steps. In the model, the degree of belief for an event B_k is a combination of the previous belief about the event and a weighting of the current information:

$$B_k = B_{k-1} + w_k \cdot (s(x_k) - R). \tag{35}$$

In the above equation, $s(x_k)$ is the strength of the current information, R is a reference point, and $0 < w_k < 1$ is an adjustment weight. By making assumptions about the encoding of information, the model can be reformulated as either an adding or averaging model (Hogarth and Einhorn, 1992). For the purposes of this paper, we will focus on the adding version of the model because we previously demonstrated that the adding model is superior to the averaging model in accounting for order effects (Trueblood and Busemeyer, 2011).

According to Hogarth and Einhorn (1992), the adding model arises when information is encoded in an absolute manner. It is assumed that $R = 0$ and $-1 \leq s(x_k) \leq 1$. Further, Hogarth and Einhorn (1992) made the assumption that the adjustment weight w_k depends on the state of the current belief and the sign of the difference $s(x_k) - R$. Specifically, if $s(x_k) \leq R$ then $w_k = B_{k-1}$, and if $s(x_k) > R$ then $w_k = 1 - B_{k-1}$. Using these constraints, the adding model is given by

$$B_k = \begin{cases} B_{k-1} + B_{k-1} \cdot s(x_k), & \text{if } s(x_k) \leq 0 \\ B_{k-1} + (1 - B_{k-1}) \cdot s(x_k), & \text{if } s(x_k) > 0 \end{cases} \tag{36}$$

Order effects arise from the model through the combination of the strength parameters and adjustment weights. The model requires as many strength parameters as pieces of information in the task. For example, the model would require two free parameters to fit the data from the order effects experiment discussed above. This is the same number of parameters used by the quantum inference model.

In previous work examining order effects (Trueblood and Busemeyer, 2011), the quantum model provided better fits to experimental data than the adding model. We also showed the quantum model more readily generalized across different response scales and populations through cross-validation. Further, the quantum model, unlike the adding model, made correct *a priori* predictions about probability judgments in jury decision-making tasks involving irrefutable evidence.

While the adding model can produce order effects, the model cannot provide an adequate account for predictive and diagnostic judgments with strong and weak alternative causes. According to the model, a predictive judgment is given by

$$p(\text{Effect} \mid \text{Cause}) = B_E = B_0 + (1 - B_0) \cdot s(\text{Cause}). \tag{37}$$

where it is assumed $s(\text{Cause}) > 0$ and B_0 is the prior belief in the effect. In order to account for the lack of a significant difference between predictive judgments involving strong and weak alternative causes, the model requires the strength of causes with strong alternatives to be equal to the strength of causes with weak alternatives. When considering causes such as "A mother has a drug-addiction" and "A mother has dark skin," this assumption seems unlikely.

Now consider the findings with full and no-alternative conditionals. According to the model, a predictive judgment with a no-alternative conditional is given by

$$\begin{aligned} &p(\text{Effect} \mid \text{Cause, No Alternative Causes}) \\ &\quad = B_E + B_E \cdot s(\text{Alternative Causes}) \end{aligned} \tag{38}$$

where B_E is given in equation 37 and $s(\text{Alternative Causes})$ is assumed to be negative because the alternative causes are absent. Thus to account for the experimental finding that predictive judgments with full and no-alternative conditionals are the same, the model requires

$$B_E = B_E + B_E \cdot s(\text{Alternative Causes}) \tag{39}$$

implying that $s(\text{Alternative Causes}) = 0$. It seems unlikely that information such as "[a patient] has not been diagnosed with any other medical or psychiatric disorder that would cause lethargy" would have a strength rating of zero. Thus, the adding model also fails to provide an adequate account of predictive judgments with full and no-alternative conditionals.

6. DISCUSSION

This paper illustrates that the quantum inference model can account for data from three different causal reasoning experiments. The quantum model is the first model that has been able to provide a unified account for all three effects. Previous models such as the causal Bayes net discussed in Fernbach et al. (2011) and the belief-adjustment model developed by Hogarth and Einhorn (1992) can only account for a subset of the findings. Further, the quantum model has previously been used to account for order effect data in a number of different inference tasks (Trueblood and Busemeyer, 2011) illustrating the generalizability of the model to a large range of phenomena.

The quantum inference model uses the concept of incompatibility to account for both the three causal reasoning phenomena presented in this paper and the order effect phenomena discussed in Trueblood and Busemeyer (2011). It might be the case that

humans can adopt either compatible or incompatible representations and are not constrained to use one or the other. In the case where individuals use a compatible representation, judgments should agree with the laws of classical probability theory. For common situations where circumstances are clear, it seems reasonable that individuals would adopt a compatible representation. For example, consider an electric kettle that only operates when it is plugged in (cause 1) and when it is switched on (cause 2). Because people have a great deal of experience with plugging in and switching on electronic appliances, they can form a compatible representation of these two causes.

However, for situations involving deeply uncertain events that have never before been experienced, perhaps incompatible representations are used. In this way, an incompatible representation is only adopted for causes that do not have the advantage of a wealth of past experience. For example, in the order effects experiment discussed in the previous section, it is doubtful that the subjects had prior experience considering a high school's food budget and an activist group fighting childhood obesity. Thus, these two causes are represented as incompatible because they cannot be accessed simultaneously without interfering with each other. In general, incompatibility offers an efficient and practical way for a cognitive system to deal with a large variety of information.

While the present paper does not want to conclude that the quantum inference model is true, the evidence presented here makes a convincing case for considering the quantum model to be a viable new candidate for modeling human causal reasoning. Using the same underlying principles, the model provided accurate fits to the data from experiments by Fernbach et al. (2010, 2011). More importantly, the model made an *a priori* prediction that order effects would occur in causal reasoning problems. The existence of order effects is a strong indicator that events should be treated as incompatible. As the key assumption of the model is the incompatibility of causes, the empirical finding of order effects is quite noteworthy. Future work will test the model with larger data sets, examine model complexity, and explore the model's predictions regarding the occurrence of primacy and recency effects.

REFERENCES

Aerts, D. (2009). Quantum structure in cognition. *J. Math. Psychol.* 53, 314–348.

Atmanspacher, H., Filk, T., and Romer, H. (2004). Quantum zero features of bistable perception. *Biol. Cybern.* 90, 33–40.

Bergus, G. R., Chapman, G. B., Levy, B. T., Ely, J. W., and Oppliger, R. A. (1998). Clinical diagnosis and order of information. *Med. Decis. Making* 18, 412–417.

Busemeyer, J. R., Pothos, E., Franco, R., and Trueblood, J. S. (2011). A quantum theoretical explanation for probability judgment errors. *Psychol. Rev.* 118, 193–218.

Carroll, J., and Chang, J.-J. (1970). Analysis of individual differences in multidimensional scaling via an n-way generalization of the "Eckhard young" decomposition. *Psychometrika* 35, 283–319.

Conte, E., Khrennikov, Y. A., Todarello, O., Federici, A., and Zbilut, J. P. (2009). Mental states follow quantum mechanics during perception and cognition of ambiguous figures. *Open Syst. Inform. Dyn.* 16, 1–17.

Fagin, R., and Halpern, J. Y. (1991). "A new approach to updating beliefs," in *Uncertainty in Artificial Intelligence*, Vol. vi, eds P. Bonissone, M. Henrion, L. Kanal, and J. Lemmer (Amsterdam: Elsevier Science Publishers), 347–374.

Fernbach, P. M., Darlow, A., and Sloman, S. A. (2010). Neglect of alternative causes in predictive but not diagnostic reasoning. *Psychol. Sci.* 21, 329–336.

Fernbach, P. M., Darlow, A., and Sloman, S. A. (2011). Asymmetries in predictive and diagnostic reasoning. *J. Exp. Psychol. Gen.* 140, 168–185.

Gopnik, A., Sobel, D. M., Schulz, L. E., and Glymour, C. (2001). Causal learning mechanisms in very young children: two-, three-, and four-year-olds infer causal relations from patterns of variation and covariation. *Dev. Psychol.* 37, 620–629.

Hastie, R., and Park, B. (1986). The relationship between memory and judgment depends on whether the judgment task is memory-based or on-line. *Psychol. Rev.* 93, 258–268.

Hogarth, R. M., and Einhorn, H. J. (1992). Order effects in belief updating: the belief-adjustment model. *Cogn. Psychol.* 24, 1–55.

Kushnir, T., Gopnik, A., Schulz, L., and Danks, D. (2003). "Inferring hidden causes," in *Proceedings of the 25th Annual Conference of the Cognitive Science Society*, eds R. Alterman and D. Kirsh (Boston, MA: Cognitive Science Society), 699–703.

Lu, Z.-L., and Dosher, B. (2008). Characterizing observers using external noise and observer models: assessing internal representations with external noise. *Psychol. Rev.* 115, 44–82.

Narens, L. (2003). A theory of belief. *J. Math. Psychol.* 47, 1–31.

Nielsen, M. A., and Chuang, I. L. (2000). *Quantum Computation and Quantum Information.* Cambridge: Cambridge University Press.

Pothos, E. M., and Busemeyer, J. R. (2009). A quantum probability explanation for violations of 'rational' decision theory. *Proc. R. Soc. Lond. B Biol. Sci.* 276, 2171–2178.

Rotello, C. M., Macmillan, N. A., and Reeder, J. A. (2004). Sum-difference theory of remembering and knowing: a two-dimensional signal detection model. *Psychol. Rev.* 111, 588–616.

Sakurai, J. J. (1994). *Modern Quantum Mechanics.* Addison-Wesley Publishing Company.

Shepard, R. N. (1962). The analysis of proximities: multidimensional scaling with an unknown distance function. *Psychometrika* 27, 125–140.

Trueblood, J. S., and Busemeyer, J. R. (2011). A quantum probability account of order effects in inference. *Cogn. Sci.* 35, 1518–1552.

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: 05 January 2012; accepted: 20 April 2012; published online: 14 May 2012.

Citation: Trueblood JS and Busemeyer JR (2012) A quantum probability model of causal reasoning. Front. Psychology ***3****:138. doi: 10.3389/fpsyg.2012.00138*

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

APPENDIX

STIMULI FOR ORDER EFFECTS EXPERIMENT

Scenario 1

- Initial Description: Mary is an average 33-year old American woman.
- Effect: How likely is it that Mary will weigh less in 1 month?
- Absent cause: Mary does not make any changes to her diet over the course of the month.
- Present cause: Mary recently began an exercise program where she works out for 4 h every week.

Scenario 2

- Initial Description: The Central High School football team won less than half of their games last season.
- Effect: How likely is it that the Central High School football team will have a winning season next year?
- Absent Cause: The football team uses the same plays this coming seasoning as they have in the past.
- Present Cause: The football team increases their weekly practice time.

Scenario 3

- Initial Description: Sara is a 40-year old American woman who has a generalized anxiety disorder.
- Effect: How likely is it that Sara will be less anxious within 3-months?
- Absent cause: Sara does not change her level of exercise over the 3-month period.
- Present cause: Sara meets with a psychologist every week.

Scenario 4

- Initial Description: Jane has two exams 1 week from today, one in her advanced physics course and one in her statistics course.
- Effect: How likely is it that Jane will do well on both exams next week?
- Absent cause: Jane does not make any changes to the amount of time she studies at home over the coming week.
- Present cause: Jane has been going to office hours for both classes for the last 3 weeks.

Scenario 5

- Initial Description: A soda company owns a popular caffeine free drink.
- Effect: How likely is it that sales of the caffeine free drink will increase next year?
- Absent cause: The advertising budget for the caffeine free drink for the coming year is the same as last year.
- Present cause: The soda company lowers the price of the caffeine free drink.

Scenario 6

- Initial Description: Paul is an average high school junior.
- Effect: How likely is it that Paul will be accepted into a top 50 college in 1 year?
- Absent cause: Paul does not make any changes to his extracurricular activities over the course of the year.
- Present cause: Paul improves his grades in all of his academic classes.

Scenario 7

- Initial Description: H. G. Industries is a manufacturing company.
- Effect: How likely is it that the output of H. G. Industries will increase over the course of a year?
- Absent Cause: H. G. Industries does not make any changes to their production line technology.
- Present Cause: H. G. Industries increases the number of employes working for the company.

Scenario 8

- Initial description: Liz is a 20-year old college sophomore who has a 3.0 GPA.
- Effect: How likely is it that Liz will earn an A in social psychology this semester?
- Absent Cause: Liz does not make any changes to her study habits this semester.
- Present Cause: Liz hopes to study social work in graduate school.

Scenario 9

- Initial description: L.Z. Inc. has a manufacturing plant that has been dumping waste in nearby Lake Lime for several years.
- Effect: How likely is it that L.Z. Inc. will start an initiative to clean up Lake Lime this year?
- Absent cause: L.Z. Inc. is using the same manufacturing process this year that it has in the past.
- Present cause: L.Z. Inc. has met with several environmental groups recently.

Scenario 10

- Initial description: A high school cafeteria serves lunch to students, and sets its upcoming menus at the beginning of each month.
- Effect: How likely is it that the cafeteria will serve healthier foods next month?
- Absent cause: The food budget for the coming month is the same as last month.
- Present cause: A group of parents are working to fight childhood obesity and have spoken to the school about including healthier options on their menus.

CALCULATIONS FOR THE QUANTUM MODEL

For all of the calculations, it is assumed that there is an initial belief state $|\psi_0\rangle$. To calculate probabilities for predictive judgments with a single present cause, the initial belief state is revised according to

$$|\psi_1\rangle = \frac{\left(P(e,p) + P(\bar{e},p)\right)\ U|\psi_0\rangle}{\left\| \left(P\left(e,p\right) + P\left(\bar{e},p\right)\right)\ U|\psi_0\rangle \right\|}.$$

This new state is then projected onto the "effect exists" subspace:

$$p\,(\text{Effect}|\text{Cause}) = \| \ P\,(e)\,|\psi_1\rangle \,\|^2.$$

If there is an additional absent cause, the $|\psi_1\rangle$ belief state is updated according to

$$|\psi_2\rangle = \frac{(P\,(e,a) + P\,(\bar{e},a))\ U_2\ |\psi_1\rangle}{\|(P\,(e,a) + P\,(\bar{e},a))\ U_2\ |\psi_1\rangle\|}$$

The probability is calculated by projecting this new state onto the "effect exists" space:

$$p\,(\text{Effect} \mid \text{Cause},\ \text{No Alternative Causes}) = \parallel P\,(e) \mid \psi_2\rangle \parallel^2.$$

For diagnostic judgments, the initial state is first revised by projecting it onto the "effect exists" subspace so that $|\psi_1\rangle = P(e)|\psi_0\rangle$. To calculate the probability of a present cause given the effect, the new belief state is revised and projected onto the "cause is present" subspace:

$$p\,(\text{Cause} \mid \text{Effect}) = \parallel P\,(c)\ U \mid \psi_1\rangle \parallel^2.$$

If there is an additional absent cause, the $|\psi_1\rangle$ belief state is updated according to

$$|\psi_2\rangle = \frac{(P\,(e,a) + P\,(\bar{e},a))\ U_2\ |\psi_1\rangle}{\parallel (P\,(e,a) + P\,(\bar{e},a))\ U_2\ |\psi_1\rangle \parallel}.$$

The probability is calculated by revising this new state and projecting onto the "cause is present" subspace:

$$p(\text{Cause} \mid \text{Effect},\ \text{No Alternative Causes}) = \parallel P\,(c)\ U_1\ |\psi_2\rangle \parallel^2.$$

frontiers in
PSYCHOLOGY

HYPOTHESIS AND THEORY ARTICLE
published: 02 April 2013
doi: 10.3389/fpsyg.2013.00150

A canonical theory of dynamic decision-making

***John Fox*[1,2]*, *Richard P. Cooper*[3] *and David W. Glasspool*[4]**

[1] *Oxford University, Oxford, UK*
[2] *University College London, London, UK*
[3] *Birkbeck, University of London, London, UK*
[4] *Deontics Research, Oxford, UK*

Edited by:
Erica Yu, University of Maryland, USA

Reviewed by:
Christian C. Luhmann, Stony Brook University, USA
Lael Schooler, Max-Planck Institute for Human Development, Germany

****Correspondence:***
John Fox, Department of Engineering Science, Parks Road, Oxford OX1 3PJ, UK.
e-mail: john.fox@eng.ox.ac.uk

Decision-making behavior is studied in many very different fields, from medicine and economics to psychology and neuroscience, with major contributions from mathematics and statistics, computer science, AI, and other technical disciplines. However the conceptualization of what decision-making is and methods for studying it vary greatly and this has resulted in fragmentation of the field. A theory that can accommodate various perspectives may facilitate interdisciplinary working. We present such a theory in which decision-making is articulated as a set of canonical functions that are sufficiently general to accommodate diverse viewpoints, yet sufficiently precise that they can be instantiated in different ways for specific theoretical or practical purposes. The canons cover the whole decision cycle, from the framing of a decision based on the goals, beliefs, and background knowledge of the decision-maker to the formulation of decision options, establishing preferences over them, and making commitments. Commitments can lead to the initiation of new decisions and any step in the cycle can incorporate reasoning about previous decisions and the rationales for them, and lead to revising or abandoning existing commitments. The theory situates decision-making with respect to other high-level cognitive capabilities like problem solving, planning, and collaborative decision-making. The canonical approach is assessed in three domains: cognitive and neuropsychology, artificial intelligence, and decision engineering.

Keywords: decision-making, autonomous agents, clinical decision-making, unified theories of cognition, cognitive systems

INTRODUCTION

The ability to respond flexibly to changing circumstances is fundamental to adaptive behavior in humans and other animals, and to artificial systems such as autonomous software agents and robots. Decision-making is a major source of theoretical questions (e.g., in economics, cognitive and social psychology, computer science, and AI) and practical challenges (e.g., in business, politics and conflict management, investments and insurance, voter and consumer behavior, law, and medicine). This vast range of interests has unfortunately led to great divergence of research methodologies (e.g., empirical observation, mathematical analysis, computational modeling, philosophical discourse) and fragmentation of decision research. There have of course been major attempts to develop domain independent perspectives, such as normative frameworks (e.g., Bayesian and expected utility models; game theory), behavioral decision models (e.g., heuristics and biases and prospect theory), and information processing approaches (e.g., neural networks and cognitive architectures). However, these attempts tend to take place from the viewpoint of one community and opportunities for sharing insights and theoretical unification are missed.

We offer a unified view of decision-making which addresses the following questions.

1. How can we understand the dynamic lifecycle of decision-making: from the situations and events that make a decision necessary to the influence of prior knowledge, beliefs, and goals which determine how a decision will be framed, preferences arrived at, and commitments to actions made (Fox and Das, 2000)?
2. What are the *general functions* that underpin and constrain the processes that implement such a lifecycle for any kind of cognitive agent, whether the agent is natural or artificial?
3. How does decision-making, conceived in this very general way, fit within cognitive science's strategic objective of a *unified theory of cognition* that can cut across psychology, computer science, AI, and neuroscience (e.g., Newell, 1990; Anderson, 2007; Shallice and Cooper, 2011)?
4. How can we apply this understanding to *decision engineering*, drawing on insights into how decisions are and/or ought to be made to inform the design of autonomous cognitive agents and decision support systems (e.g., Fox et al., 2003; Fox et al., 2010)?

The goal of a unified theory is ambitious, some will say hubristically so. However despite the long-term objective of unification the objective of this paper is more modest: to establish a framework and a language which can facilitate discussion between decision researchers in different communities, from theorists with distinct but complementary perspectives to practitioners such as doctors, engineers, and managers who wish to improve their decision-making.

We begin with a brief overview of classical approaches to decision theory, which we contrast with theories of dynamic decision-making (DDM). We then introduce some perspectives on DDM which we believe have been neglected. This paves the way for the introduction of the canonical framework, in which we seek to understand DDM in terms of a number of key capabilities which, we assert, a cognitive agent must have. These *canons* are abstracted from field-specific details; we acknowledge there are countless possible implementations and interpretations of the canons. In the final section we assess the canonical theory in three restricted settings: cognitive neuropsychology; artificial intelligence; and the design of practical decision support systems.

TRADITIONAL METHODOLOGIES AND THEORIES IN THE DECISION SCIENCES

Decision-making may be defined in very general terms as a process or set of processes that results in the selection of one item from a number of possible alternatives. Within this general definition, processes might be natural and conscious, as in deliberate choice amongst alternatives, but also unconscious (as in selecting the grip to use when grasping an object) or artificial (as in an expert system offering decision support). Moreover, decisions can be about *what to do* (*action*), but also about *what to believe* (*opinion*). We will later extend this definition to cover DDM, but for now it is sufficient to summarize three classical perspectives that are common in decision research.

PRESCRIPTIVE THEORIES OF RATIONAL CHOICE

Prescriptive decision theories have emerged from mathematics and mathematical economics where *rational choice* is taken to be central to understanding economic behavior and managing economic systems efficiently. The methodology focuses on establishing rational axioms for making decisions under uncertainty and consequences for systems of trade and commerce against defined valuations. The axioms typically express mathematical constraints which, if violated, can lead a decision-maker into suboptimal choices. Such prescriptive theories tend to be agnostic about the processes or algorithms that might implement or operationalize the mathematical constraints. Despite their theoretical importance the application of classical prescriptive decision models suffers from the practical problem that it is often difficult to estimate the quantitative parameters that they require (e.g., probabilities, utilities). Although they have informed research on human decision processes they provide limited insight into them and ignore key theoretical problems in DDM.

DESCRIPTIVE THEORIES OF NATURAL DECISION-MAKING

The goals of psychology are to explain human behavior and predict performance, irrespective of how performance compares with rational norms. Early psychological models of decision-making were influenced by rationalist theories as sources of theoretical concepts and normative standards against which to assess human decision-making, but there has been a trend away from this in recent decades. For example Simon's (1957) concept of "bounded rationality" emphasized human limited information processing capacity and strategies for accommodating this (e.g., satisficing). Kahneman and Tversky's heuristics and biases program also sought a more realistic account of cognitive processes in decision-making (Tversky and Kahneman, 1974) and Kahneman and Tversky (1979) developed a better description of how people evaluate potential losses and gains compared to mathematically prescribed norms. More recently Gigerenzer and Todd (2000) argue for the practical importance of simple heuristic strategies for fast decision-making.

DESIGN FRAMEWORKS FOR DECISION ENGINEERING

In contrast to the above perspectives, designers of *decision support systems* and other decision-making software view decision processes and applications in a way that is analogous to designing objects like bridges and buildings. Decision engineers therefore tend to be interdisciplinary in their approach, exploiting mathematical and normative theories, or being inspired by human decision-making as inartificial neural networks and "expert systems," or adopting a pragmatic mix of both. Decision engineers' primary concerns are with achieving specific objectives and they may adopt any methods that are effective in achieving this goal. Despite considerable practical success decision engineering risks use of *ad hoc* rather than principled design theories and, as a consequence, there can be considerable uncertainty about the performance of decision systems in practice.

DYNAMIC DECISION-MAKING

In all the above perspectives decision-making is typically viewed as a choice between a set of *predefined options*. This is unsatisfactory because a decision typically arises within a wider context, in which it is necessary to recognize when a situation or event requires a decision, determine the set of options, establish criteria for determining preferences, resolve conflicts, and so on. This is the focus of DDM.

Edwards (1962) characterized DDM in terms of the following features: (1) a series of decisions is required to achieve a goal; (2) decisions are not independent (decisions are constrained by earlier decisions); (3) the state of the problem changes, with changes in the environment, or as a consequence of the decision-maker's actions, and (4) decisions are made in real time. DDM is fundamental in practical domains, such as fire fighting, factory production, clinical decision-making, air traffic control, military command, and control, emergency management. In this section we briefly overview a few attempts to address the complexity of DDM.

NATURALISTIC DECISION MODELS

Writing for practitioners Drummond (1991) provides a "synoptic model" of a full decision cycle as follows:

1. Identify problem
2. Clarify and prioritize goals
3. Generate options
4. Evaluate options
5. Compare predicted consequences of each option with goals
6. Choose option with consequences most closely matching goals

She also identifies features of practical decision-making that are not so much to do with the dynamics of choice but are significant for a general account of decision-making, including Individual differences; collaboration, and joint decision-making; multiple

criteria and conflicting goals, and "problems within problems" (any step in a decision process can unexpectedly turn into a new problem as complex as the original one).

Klein's (2008) Naturalistic Decision-Making program seeks to understand how people make decisions in the real world and how they perform tasks in demanding situations. Important real world challenges include vagueness of goals, high stakes, significant levels of uncertainty about what is the case or what the consequences of actions might be, time pressure, team and organizational constraints, changing circumstances, and limited knowledge and experience. From extensive studies of experts such as fire fighters, he concluded that "by 1989, it was fairly clear how people didn't make decisions. They didn't generate alternative options and compare them on the same set of evaluation dimensions. They did not generate probability and utility estimates for different courses of action" (2008, p. 456). Klein wants to go beyond the simple choice paradigm of classical decision theory to ask how people maintain situational awareness, make sense of events, formulate goals, and construct plans to achieve them.

Brehmer (1992) also advocates a naturalistic approach, adopting computer simulations of real world situations such as fighting a forest fire as a research platform. Here it is possible to investigate features of human performance like the development and use of decision strategies, the importance of feedback, and how people learn to control complex and evolving situations, which find little place in classical decision theory or laboratory experiments. Since the contingencies of the simulation are under complete program control even complex decision-making can be studied in a systematic way.

In a well known discussion of human expertise Shanteau (1987) observed that expert decision-makers have many capabilities that cannot be accounted for with traditional theories. They know what to attend to in busy environments, what is *relevant* to decisions, and when to make exceptions to general rules, adapt to changing task conditions, and find novel solutions to problems. They also know a lot about what they know and can articulate the rationale for their decisions in terms of the relevant evidence and facts that support different options.

A general theory of DDM must address such capabilities, whether the focus is rational choice, human cognition, or engineering.

COGNITIVE ARCHITECTURES

As early as the 1950s Newell and Simon were exploring the value of computational concepts in understanding human cognition (e.g., Newell et al., 1958) which evolved into rule-based models (Newell and Simon, 1972) and finally the Soar project (Laird et al., 1987). Soar is relevant here because it was developed as a model of general intelligence that subsumed decision-making as a key component, and was seen as offering a unifying view of human and artificial intelligence[1].

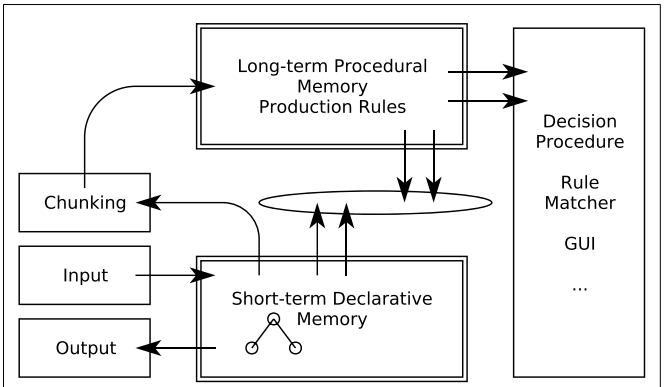


FIGURE 1 | The Soar information processing architecture.

The Soar architecture (**Figure 1**) showed how a relatively simple information processing mechanism could carry out a wide range of cognitive tasks. It extended Newell and Simon's established production rule approach by introducing some capabilities that a general theory of DDM needs to address, including dynamic generation of task goals, selection, and application of knowledge (rules) from long-term memory, and executing general problem solving strategies when no specific rules are available. In the latter case Soar "chunks" a new rule from the problem solving trace and adds it to long-term memory for use in future similar circumstances (Laird et al., 1987; Newell, 1990).

A central mechanism of the Soar architecture is a cyclical *decision procedure* (**Figure 2**). This reacts to and interprets new data ("elaboration") and makes a decision by comparing alternative cognitive operations based on the interpretation, selecting one and then applying it, resulting in a change in the state of short-term memory. This leads to a new cycle of processing.

Soar has been extensively used for modeling human performance on complex tasks and for designing and implementing expert systems, and is the foundation of Newell's (1990).

Our own work has also focused on computational architectures for modeling DDM (Fox, 1980; Cooper et al., 2003) and high-level cognition using systematic methods and tools (Cooper and Fox, 1998; Cooper et al., 2002). We have compared several cognitive processing models, including: rule-based models; Bayesian inference and connectionist classifiers and heuristic architectures. The models were successful in that we were able to simulate behavior on complex decision tasks in some detail but firm theoretical conclusions have proved elusive, because:

1. Seemingly distinct theoretical accounts have comparable abilities to simulate observed behavior at similar levels of detail because, we believe, many competing theories have comparable descriptive and explanatory power. They may in fact be indistinguishable in principle. Debating which particular class of theory best describes human decision-making may be unproductive.

[1] Soar was the first computationally well-specified cognitive architecture (see Newell, 1990). While Soar continues to be developed (e.g., Laird, 2012), discussions of cognitive architecture are now generally dominated by Anderson's (2007) ACT-R theory. Our focus on Soar is based on Soar's decision process and what Newell (1990) referred to as the *Problem Space Computational Model* (cf. **Figure 2**). This provides an intermediate level at which behavior may be described. This level of analysis is not considered within ACT-R.

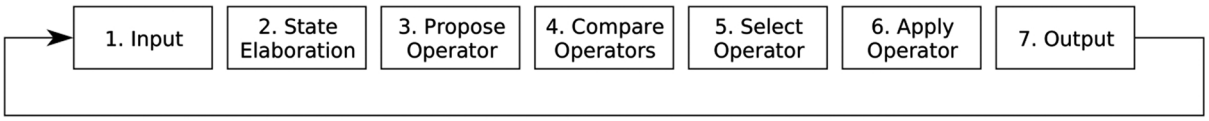


FIGURE 2 | The Soar dynamic decision cycle.

2. There was frequently more variance in individual subjects' behavior than between the models. Even if there is a fixed decision-making mechanism (a la Soar) a decision-maker's knowledge and what is learned on task are at least as important in explaining performance data. What a decision-maker "knows" will often have a greater impact on performance than any hypothetical rule-based, connectionist, Bayesian or other mechanism (cf. Newell, 1981).

ARTIFICIAL INTELLIGENCE AND AUTONOMOUS AGENTS

AI researchers have sought to design mechanisms for controlling robots and other automata in many DDM tasks, including situation monitoring and assessment, problem solving, scheduling, and planning through to cognitive vision and natural language understanding systems. Since the late 1980s there has been particular interest in the concept of *autonomous agents*, and in *multi-agent systems* in which agents cooperate to achieve shared goals.

In AI an autonomous agent is an entity (usually software) that inhabits some sort of environment and can react to situations and events and behave purposefully to achieve its goals. The environment may be physical (e.g., the agent is a robot or autonomous vehicle) or virtual (e.g., a simulation or the world wide web). **Table 1** summarizes the main capabilities that agent theorists have sought to automate, under three headings: interaction with the environment, cognitive capabilities, and cognitive control.

A few features of this table deserve comment. First, decision-making is an important cognitive function, but it is only one of a network of interrelated capabilities; reasoning; and problem solving can contribute to decision-making (in formulating decision options for example) while decision-making can contribute to problem solving and planning by assessing and selecting alternative problem solving strategies, plans, etc. Learning, in contrast, cuts across these capabilities in that any solution to a problem, plan or decision-making strategy that successfully achieves a goal may be worth remembering for future reuse. Second, autonomous decision-making can have multiple control regimes. Problem solving, planning, and even decision-making itself can be viewed as *deliberative* in that an agent reflects on its circumstances and goals to assemble one or more possible solutions to achieving its goals. On the other hand if the agent has learned from previous cases then it can operate *reactively* by retrieving candidate solutions from its knowledge base and making a decision by comparing the relative merits of the options.

A prominent computational theory of autonomous agent control and behavior draws on ideas from philosophy, psychology, and computer science in formalizing the concept of an agent. Following

Table 1 | Capabilities that are typical of agent systems described in the AI literature (Fox et al., 2003).

INTERACTIONS WITH ENVIRONMENT	
Perception	Observing and monitoring situations and events in the environment
Action	Executing actions that change or control the environment
Communication	Employing perception and action to interact with other agents
COGNITIVE CAPABILITIES	
Reasoning	Making inferences on the basis of environmental data, beliefs, goals, knowledge, etc.
Problem solving	Searching for explanations of observations, plans which will achieve goals etc.
Decision-making	Choosing between alternative hypotheses or actions
Scheduling	Sequencing actions and plans flexibly in response to circumstances
Planning	Constructing a set or sequence of actions to achieve a goal
Learning	Remembering solutions to newly encountered problems for future reuse
CONTROL CAPABILITIES	
Reactive behavior	Responding to situations and events in real time
Deliberative behavior	The application of cognitive capabilities in a purposive, coordinated way
Autonomy	Making plans, taking decisions, etc. without external programing or supervision

Bratman (1987) an agent is said to have *mental states* such as *beliefs*, *desires*, and *intentions* (**Table 2**). Such "BDI agents" have proved to be a practical basis for designing software agents (e.g., Rao and Georgeff, 1995). There are now many examples of agents which monitor their environments and maintain *beliefs* about them; generate goals (*desires*) with respect to the environment state, and if these are not consistent with their beliefs adopt plans (*intentions*) which will bring the environment into line with these goals.

Knowledge, beliefs, desires, and intentions are often held to be mere "folk psychology," of little scientific interest (e.g., Churchland, 1981). The fact that it has been possible to develop a formal interpretation of these and other cognitive states (e.g., Cohen and Levesque, 1990) and that they can be used to ground the design of practical software agents, suggests that such notions may have more theoretical power for understanding cognitive systems than is sometimes claimed. In the next section we discuss how

Table 2 | Some of the mental/cognitive states that have been studied in AI.

COGNITIVE STATES	
Beliefs	Specific information which an agent holds to be true at a particular moment in time
Desires	Specific goals which are currently influencing an agent's behavior
Intentions	Specific commitments to actions or plans which an agent has decided to carry out
Knowledge	General theories, rules, functions etc as distinct from situation-specific beliefs, desires, and intentions

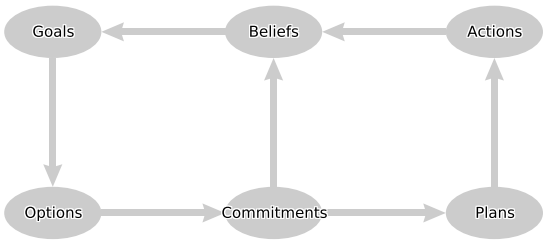


FIGURE 3 | The domino agent framework, an enhanced BDI agent model.

they theories of mental states can illuminate our understanding of DDM.

MENTAL STATES AND DYNAMIC DECISION-MAKING

Like Klein, Shanteau, Brehmer, and others our wish to understand high-level cognition has taken us out of the laboratory and into a world where decision-making is complex and indeed so difficult that even experienced practitioners, clinicians, make significant, and perhaps frequent errors[2]. We have studied decision-making in many routine medical settings including risk assessment; selection of tests and investigations; diagnosing the causes of a complaint; choosing treatments and prescribing drugs; implementing treatment plans; and team-based decision-making. This has led to a general framework for understanding clinical expertise and designing decision support tools that draws on some of the ideas discussed above. The "domino" model in **Figure 3** is a computational architecture in which cognitive states provide an expressive and productive basis for describing cognitive processes throughout the decision cycle.

Each node of the domino represents a cognitive state of a particular type and each arrow represents a process for updating these states: beliefs can lead to new goals; goals to options for decision-making, and options can lead to commitments (about what to believe or what to do) with an associated rationale. We first explain the model by means of a simple medical scenario and then outline how the processes that generate the states can be computationally realized.

> "Joan Smith has been rapidly losing weight, and there is no obvious reason for this." In a clinical setting this scenario would typically lead to intentions to decide the cause of the weight loss and, if necessary, decide on the best action. There are several possible physical or psychological causes, and hence multiple hypotheses for explaining the patient's complaint. Once the diagnostic options have been identified a decision-maker can determine what additional information to obtain (by asking questions, ordering investigations, etc.), and construct arguments for and against competing hypotheses based on the results. In due course the decision-maker can commit to a belief about the most convincing cause of the clinical problem.
>
> Suppose the decision-maker accepts a diagnosis of gastric ulcer. This leads to a new goal: decide the best treatment for the ulcer. Knowledge of gastrointestinal disease suggests a range of treatment options, and arguments can be constructed for and against the alternatives based on efficacy, side-effects, costs, drug interactions, and so on. A decision about the most preferred treatment plan is based on an assessment of all the arguments. The preferred treatment may be simple, like prescribing a drug, or a complex care plan of many steps. Plan steps may lead to new observations, leading in turn to new goals and changes to the plan, and sometimes reversal of earlier decisions.

MODELING WHAT A DECISION-MAKER KNOWS

Work in AI and cognitive modeling shows that an important challenge for decision theory concerns the representation and use of knowledge. Newell (1981) proposed that cognitive systems must be characterized at what he termed *the knowledge level* as well as the information processing level. This must describe the organization and semantics of knowledge, which enable and constrain cognitive processing. Since Newell's paper there has been a great deal of work on modeling knowledge as frames and other knowledge representation techniques developed in AI, the roots of which are in the semantic networks and memory models developed in cognitive psychology, and more recently formalized as *ontologies*. We now give a brief overview of how ontological concepts are being used in computer science and AI to represent knowledge; this is informal and simplified presentation of technical work in this area but is a necessary foundation for the rest of the paper.

An ontology can be thought of as a hierarchy of knowledge structures, in which each level of the hierarchy introduces a specific type of semantic information (**Figure 4**). For example the string "SCTID397825006" is a code for the medical term "gastric ulcer" in the SNOMED CT clinical coding system (Stearns et al., 2001). The code itself does not have any meaning; it is just a string of characters. A first level of meaning can be provided by an *ontological assignment*, linking the term "gastric ulcer" to a node in a *concept network* (as in "gastric ulcer" *is a kind of* "peptic ulcer"). This links the concept to more general categories through further assignments: "peptic ulcer" *is a kind of* "disease," "disease" *is a kind of* "abnormal state," and so on. The class structure facilitates

[2]The Institute of Medicine's famous report *To err is human* (Kohn et al., 2000) put patient safety at the top of the international health care agenda in 1999 and its ramifications are still being felt.

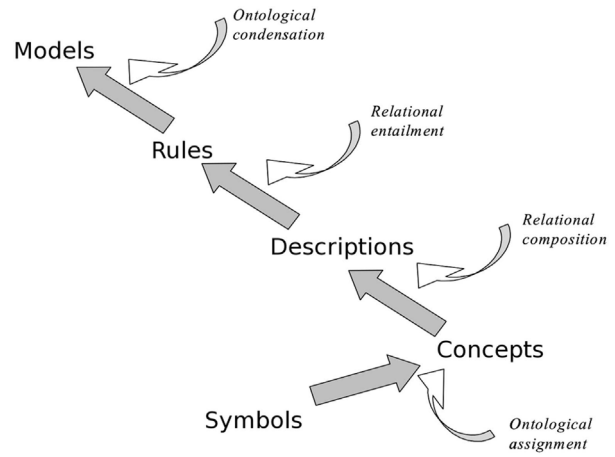


FIGURE 4 | The "ontological ladder," which formalizes knowledge as a hierarchy of increasingly complex and semantically rich conceptual structures.

an important form of reasoning called *inheritance*; the concept "gastric ulcer" can inherit properties from its super-class "peptic ulcer" and from the even more general class "disease" (e.g., every disease class and disease instance such as Joan Smith's gastric ulcer inherits the property *has_symptoms*). Concepts are indicated here using quotes, and properties and relationships using italics.

Concepts can participate in other relations as well as *is a kind of* relations, such as *causal* relations (e.g., "gastric ulcer" *causes* "hematemesis"). If a patient has a gastric ulcer we may look for symptoms which are specifically caused by this disease (e.g., "hematemesis") and, by inheritance, symptoms caused by more general kinds of peptic ulcer (e.g., "pain after meals").

The concept "disease" is also semantically linked to other broad concepts, such as "treatments" which may be linked to diseases through relations like *controls*, *eradicates*, and so forth. Treatments also have subclasses in the ontology (such as "drugs," "surgical treatments") which have class-specific properties (e.g., *side-effects*, *method of administration*) as well as properties inherited from the general class (e.g., *effective for*, *cost of*).

Descriptions can be combined (e.g., "indigestion" is "present" *and* "patient" is "elderly") and can participate in more complex structures like rules, as in

> "indigestion" is "present" *implies* "possible diagnosis" is "peptic ulcer" "indigestion" is "present" *and* "patient" is "elderly" *implies* "possible diagnosis" is "gastric ulcer"

Finally descriptions and rules can be condensed into models. Two important kinds of model that are common in medicine are the Scenario and the Task which are the foci of much clinical discussion and decision-making. For example:

> Scenario: "middle aged, overweight, male with hypertension not controlled by first-line therapy" Task: "eradicate tumor with surgery followed by 3 courses of adjuvant chemotherapy and annual follow-up for 3 yrs"

Models can participate in further ontological elaboration, forming elements of descriptions, rules, and higher-order models.

Ontologies are a major topic of research in knowledge representation, and currently offer the most sophisticated theoretical framework for understanding and applying knowledge in decision-making and other cognitive tasks.

A CANONICAL THEORY OF DYNAMIC DECISION PROCESSES

The domino model was devised as an integrated theory of reasoning, decision-making, planning, and other capabilities that an autonomous agent may possess. In previous work we have interpreted each arrow in the model formally, as a specialized logic with a distinct set of non-classical axioms and inference rules (Das et al., 1997; Fox and Das, 2000). Although the model has been successfully used in many clinical applications it is clear that the functions modeled by the domino scheme could be understood in many other ways, and that different research communities would likely adopt different interpretations. Our aim here is to re-describe the framework in a more general way in which the arrows are viewed as *canonical functions* that can be instantiated in different ways to suit the purposes of different disciplines and traditions. Each function is first summarized informally, and then presented using a notational device called a *signature*. Signatures are commonly used to describe properties of computer programs in terms of their input-output constraints without specifying the internal details of how the function is to be implemented (e.g., Spivey, 1989)[3]. The level of abstraction provided by such signatures corresponds to a first approximation to a formal characterization of Marr's computational level (Marr, 1982), specifying *what* is computed by the underlying process without specifying *how* it is computed (i.e., without specifying the algorithm that achieves the computation or the representations over which the algorithm operates).

CANON 1: BELIEF MAINTENANCE

Any agent (natural or artificial) should maintain a consistent set of beliefs and expectations with respect to its current circumstances, updating these as its environment is observed to change.

Belief maintenance is fundamental to practical decision-making and is fundamental to all the decision models discussed here[4]. Beliefs need to be revised in the light of new observations,

[3]We make two important caveats. First, we have adopted a particular vocabulary of cognitive states, such as "belief," "goal," "argument," "plan" etc. These terms are common parlance, but are often also used with a technical interpretation that differs from community to community. It would in fact be preferable to avoid these overloaded terms (perhaps calling the different types of state "beta," "gamma" "alpha," "pi," or some such) but we will persist with this more familiar terminology to make the presentation easier to understand. Readers should, however, be wary of interpreting the terms colloquially and keep the canonical meaning in mind. Second, we have adopted a language which distinguishes between different types of information (beliefs, ontologies, goals, etc.) in order to specify the canonical functions that we believe a decision-maker of the sophistication of a human agent needs to implement, but this does not imply that the internal implementation of the function must be symbolic or propositional. Moreover, use of a notation for cognitive states should not be taken to imply that we are speaking exclusively about agents that have explicit mental states. The use of signatures is agnostic about any underlying implementation provided that implementation is consistent with the constraints defined by the signatures.

[4]At the informal end of our field of interest are the humanities, subsuming philosophy and literature, social commentary and political theory, and so on. Even

new beliefs, or new knowledge. There are countless proposals for how belief maintenance can/should be implemented; well known ones include probabilistic updating; fuzzy inference; classical deduction (propositional and predicate calculus), and non-classical logics (e.g., modal, abductive, inductive, and non-monotonic logics).

Equation S1 is a general signature subsuming many kinds of belief maintenance. It expresses the idea that an agent arrives at and maintains its beliefs by applying general background knowledge (an ontology) to specific situation data.

$$\frac{\text{Observation} \times \text{Ontology}}{\text{Belief}}_{\text{BM}} \qquad \frac{\text{Belief} \times \text{Ontology}}{\text{Belief}}_{\text{BM}} \tag{S1}$$

Signatures are read as follows: a cognitive state of the type below the line (here a Belief) is functionally dependent on the cognitive state above the line under some set of axioms or algorithms BM. The × operator may be understood declaratively ("together with") or procedurally ("applied to") whichever is more intuitive. Note that the second variant of the signature is recursive, so beliefs can propagate forward if the ontology warrants this.

CANON 2: RAISING GOALS

An agent needs to ensure continuity of its intentions and actions over time, mediated by the concept of a goal state.

A goal is a cognitive state that serves to coordinate an agent's behavior even though circumstances may change and its decisions and plans need to be updated. Cognate concepts of goal include "desire" (as in intentional philosophy and in BDI theory); "drive" and "motivation" from classical psychology; "utility" from decision theory, "criteria" in multi-criteria decision models, and so on. The relationships between these terms are linguistically troublesome, but the concept has led to countless technical proposals in AI for representing and interpreting goal states in robots, planners, and other systems. The following signatures summarize the canon.

$$\frac{\text{Belief} \times \text{Ontology}}{\text{Goal}}_{\text{G}} \qquad \frac{\text{Goal} \times \text{Ontology}}{\text{Goal}}_{\text{G}} \tag{S2}$$

As before a cognitive state of the type below the line (here a Goal) is dependent on the agent's knowledge and current cognitive state. In the first of the two signatures the state that leads to a goal may be a specific scenario such as *patient presents with severe and chronic pain* leading to two goals: *decide the most plausible cause* and *decide the most preferred treatment*. The second signature covers a case common in AI planning: goals lead to sub-goals. The goal *put out the fire* may entail sub-goals to decide how to *get* to the fire, the *strategy of attack*, the *equipment* needed and so forth, any of which may lead to further sub-goals.

The next four canons are core capabilities in goal-based decision-making.

CANON 3: GENERATE OPTIONS

An agent should apply specific knowledge of previously effective solutions when it can and use general knowledge to solve problems when specific knowledge is not available.

An agent may be able to identify multiple possible solutions so we refer to these as "candidates" in subsequent signatures, using the term to subsume other common terms like "decision option," "problem solution," etc.

$$\frac{\text{Goal} \times \text{Belief} \times \text{Ontology}}{\text{Candidate}}_{\text{C}} \tag{S3}$$

Signature Eq. S3 abstracts across "strong" problem solving methods, which draw on specific domain knowledge, and domain-general but "weak" methods like means-ends analysis and constraint solving (Laird et al., 1987). It also includes intermediate methods, such as heuristic classification, which maps between ontologies, as in *symptom* → *diseases* and *diseases* → *treatments* (Clancey, 1985) and explanation-based decision-making that depends upon building causal models for choosing between actions (Pennington and Hastie, 1993).

CANON 4: CONSTRUCT REASONS

An agent should consider as many relevant lines of reasoning as is practical when establishing preferences over competing decision options.

There may be an indefinite number of reasons for accepting a hypothesis or selecting an action to achieve a goal. Reasons to believe in or to doubt hypotheses can be based on statistical evidence or logical justifications; reasons for choosing between alternative actions can be based on qualitative preferences or quantitative values. The following signature subsumes a range of strategies for constructing reasons for alternative options.

$$\frac{\text{Candidate} \times \text{Goal} \times \text{Belief} \times \text{Ontology}}{\text{Candidate} \times \text{Reason}}_{\text{R}} \tag{S4}$$

As Shanteau (1987) observed decision-making expertise is not only characterized by making good choices but also by meta-cognitive abilities like the ability to articulate the rationale for decisions. The signature provides for this meta-cognitive capability by making explicit the reasons for and against competing candidates.

CANON 5: AGGREGATE REASONS

When problem solving yields multiple candidate solutions an agent must establish an overall preference, taking account of all the reasons for each of the options.

Probabilistic updating is widely held to be the normatively correct way of establishing confidence in competing hypotheses, as is the expected utility extension for deciding about preferences over candidate actions. In many settings, however, it is impractical to estimate prior and conditional probabilities objectively, or

here some notion of belief maintenance is needed, though these areas are primarily served only by everyday language rather than technical frameworks. In fact natural languages are of course hugely expressive tools, and articulate many distinctions and subtle nuances around the idea of belief such as *possibility*, *plausibility*, *conviction*, *assumption*, *expectation*, *suspicion*, and *doubt*. It can be argued that such concepts also deserve a place in any discussion of decision-making (e.g., Fox, 2011) and deal fundamentally with the concept of belief maintenance as well. Such attitudes are frequently treated as merely part of our folk psychology and theoretically uninteresting. In our view they capture real distinctions which are important in all practical thinking and social interaction.

to model costs and benefits on a single dimension. Simpler functions for determining overall preferences are helpful here, such as the Bentham rule (add up the *reasons pro* and *reasons con* and take the difference)[5], or the equations of diffusion models (Ratcliff and McKoon, 2008). However there are many other aggregation functions which will deliver a preference ordering over the set of decision candidates. The signature below subsumes many specific functions that can establish the overall "merit" of a candidate.

$$\frac{\text{Goal} \times \text{Candidate} \times \text{Reason}}{\text{Candidate} \times \text{Merit}}_{\text{Agg}} \quad \text{(S5)}$$

One might assume that Merit must be a quantity, and aggregation a numerical algorithm. This is not necessarily the case. For example, we can describe preferences based on purely ordinal relations (A is preferred to B) and the preference determined on entirely logical grounds. Suppose, for example, that we have a reason R1 for preferring A, and a reason R2 for preferring B, but we also have some reason R3 which brings the veracity of R1 into doubt. All other things being equal we would prefer B to A. Informal preference rules, expected utility functions, argumentation systems and multi-criteria decision models are subsumed under this general signature.

CANON 6: COMMITMENT

If an agent can determine that its most preferred option will not change with further information then it should commit to that option. If there is missing information that, if known, would change the preference but the cost of acquiring that information is greater than the cost of taking the wrong decision then the agent can still be certain that its preference order will not change and a commitment can be made.

This canon of decision-making can be summarized by the following two variant signatures, covering the cases of *accepting a belief* (committing to one of a number of competing hypotheses) and *adopting a plan.*

$$\frac{\text{Candidate} \times \text{Merit} \times \text{Ontology}}{\text{Belief}}_{\text{Accept}}$$
$$\frac{\text{Goal} \times \text{Candidate} \times \text{Merit} \times \text{Ontology}}{\text{Goal} \times \text{Plan}}_{\text{Adopt}} \quad \text{(S6)}$$

Whether a belief is acceptable or not is independent of the agent's goals but the commitment to a plan is intimately bound up with the agent's (prior) goals. The Goal is also retained below the line to indicate that the commitment only holds as long as the goal is extant.

The remaining set of signatures cover capabilities which may form part of a wider theory of cognitive systems, but are less relevant to the decision-making focus of the discussion and so are dealt with more briefly.

CANON 7: PLAN ENACTMENT

If an agent is committed to a plan that is necessary to achieve one or more of its goals, then enactment of the plan should be optimized with respect to the agent's priorities and preferences.

[5]technically a linear function with equal weights

Enacting a plan can be a simple sequential execution of the plan's component steps, or involve flexible scheduling of the steps to accommodate changing circumstances. Enactment can be recursive, as execution of a step in the plan leads to new goals, which may require the current plan to be repaired, radically reconstructed, or abandoned. In all cases the effect of enactment is to update the current plan.

$$\frac{\text{Goal} \times \text{Belief} \times \text{Plan} \times \text{Ontology}}{\text{Goal} \times \text{Plan}}_{\text{Enact}} \quad \text{(S7)}$$

CANON 8: ACTION

An action is defined as a plan step that cannot be decomposed into smaller elements. If the preconditions of a planned action are satisfied then it may be executed without further decision-making.

$$\frac{\text{Plan} \times \text{Precondition}}{\text{Action}}_{\text{Execute}} \quad \text{(S8)}$$

Preconditions include logical preconditions (e.g., a situation holds) and material preconditions (e.g., a physical resource is available).

CANON 9: MONITORING

An agent should monitor the environment for important changes that may impact its goals and commitments and update its beliefs when necessary.

$$\frac{\text{Goal} \times \text{Observation} \times \text{Ontology}}{\text{Belief}}_{\text{Mon}} \quad \text{(S9)}$$

This can be viewed as a variant of the basic belief maintenance signature Eq. S1. If monitoring reveals that the preconditions of an intended action have ceased to hold (due to independent environmental changes or the effects of the agent's actions for example) these actions should be postponed or discarded. If the new situation invalidates a past decision the agent should reconsider its options.

CANON 10: LEARNING

An agent should update its knowledge as a result of experience.

There are many learning mechanisms, including associationist and statistical models in machine learning and cognitive neuroscience; case-based learning and rule induction in AI; reinforcement learning and "chunking" in cognitive psychology. Equation S10 represents learning as a generic process that updates its ontology by acquiring new scenario and task models.

$$\frac{\text{Belief} \times \text{Ontology}}{\text{Scenario}}_{\text{CL}} \qquad \frac{\text{Goal} \times \text{Belief} \times \text{Ontology}}{\text{Task}}_{\text{IL}} \quad \text{(S10)}$$

These signatures are only a starting point for describing learning in decision-making tasks; learning is a neglected topic in normative theory and behavioral studies of decision-making and we see this as a major challenge for general theories of DDM.

The 10 canons are offered as a first draft of a general framework for describing and discussing cognitive agents that can take decisions in the presence of uncertainty in dynamically changing

situations. We do not claim that the set is complete (we are sure it isn't) or that individual signatures cannot be improved (we know they can). At this point, however, we offer them as a basis for establishing an intuitive *lingua franca* for interdisciplinary discussions of decision-making theories, systems, and applications.

ASSESSING THE CANONICAL THEORY

In this section we consider the adequacy of the framework in three representative applications: (1) understanding the gross structure of the human cognitive architecture; (2) collaborative decision-making by autonomous agents, and (3) designing decision support systems.

ASSESSMENT 1: DYNAMIC DECISION-MAKING AND THE HUMAN COGNITIVE ARCHITECTURE

The canons discussed above are too general to make specific predictions about the detailed cognitive processes involved in human decision-making. They cannot therefore be seen as a theory of human decision-making as the lack of detail means they are not falsifiable in the Popperian sense (but see Lakatos, 1970; Cooper et al., 1996; Cooper, 2007). Our aims here are to show that the canons are nevertheless consistent with at least one large scale theory of the organization of cognitive processes, and that the framework provides a workable and useful approach to interpreting findings from cognitive neuropsychology and neuroscience.

The Soar cognitive architecture discussed in Section "Dynamic Decision Making" is one of several theories of the large scale organization of the human cognitive system that have been proposed and refined over the last half century (see also ACT-R: Anderson, 2007). These theories typically propose a set of functional *components* of the putative cognitive system, together with *interfaces* or mechanisms for interaction between those components, with the aim being to specify a system that is capable of supporting all aspects of human cognitive processing. The scope of these theories is therefore much broader than that of many theories in cognitive psychology, which typically focus on a single domain such as memory, attention, language comprehension, or choice.

While aspects of the canonical framework map on to processes or mechanisms within Soar, there is also a promising mapping onto the functional components of another cognitive architecture, namely the Contention Scheduling/Supervisory System (CS/SS) model of Norman and Shallice, 1986; see also Shallice and Burgess, 1996). The CS/SS model (**Figure 5**) has not been as well developed computationally as other cognitive architectures such as Soar or ACT-R. However, Glasspool (2005) proposed a mapping of the components of the CS/SS model to the domino architecture of **Figure 3**, and many of the model's components can be understood as performing functions that correspond to the canons.

The CS/SS model draws a basic distinction between routine behavior (held to be controlled by learned schemas within the CS

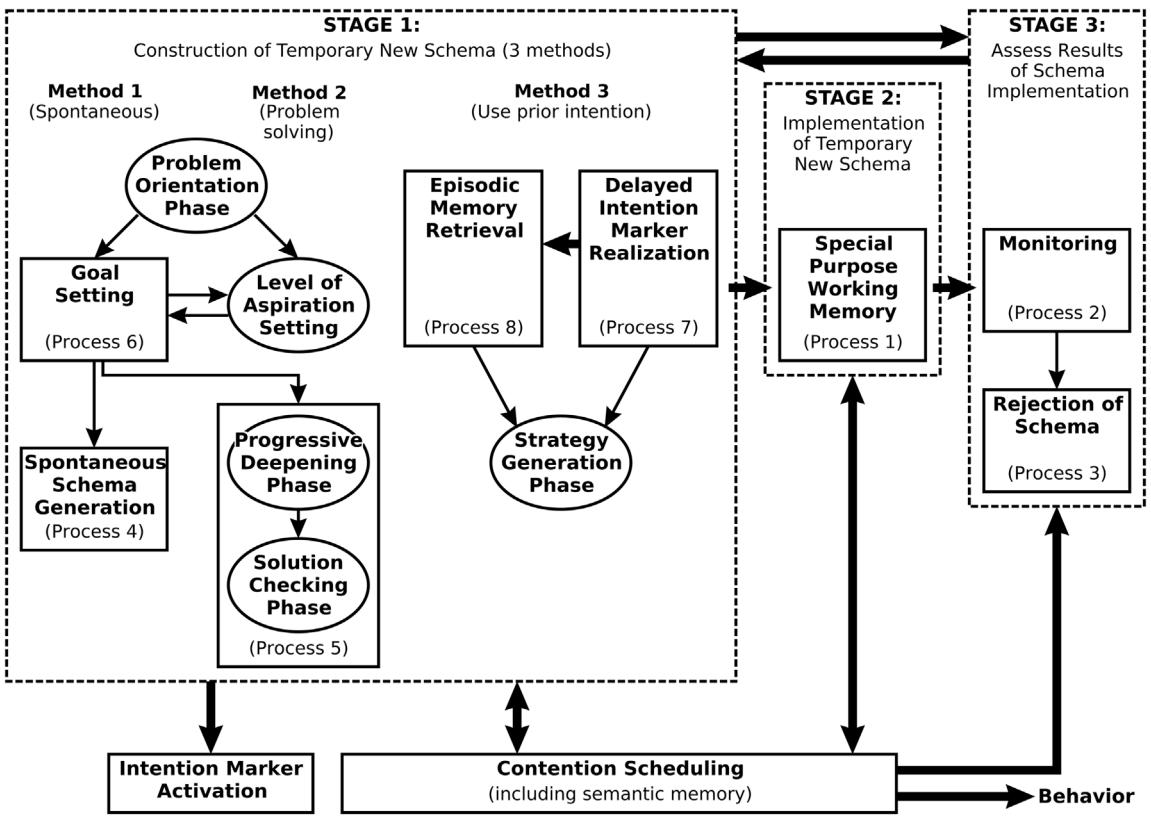


FIGURE 5 | The Contention Scheduling/Supervisory System model of Norman and Shallice (1986), as elaborated by Shallice and Burgess (1996).

system) and non-routine behavior (held to be controlled through the construction and maintenance of temporary schemas by the SS). CS operationalizes canons 7 and 8: at the lowest level it maps plans to individual actions subject to their preconditions (Eq. S8), but CS is hierarchical, and at higher levels (even with routine behavior) it may map plans to sub-plans (Eq. S7).

In the initial version of the CS/SS model described by Norman and Shallice (1986) the output of the SS – temporary schemas for the control of behavior via CS – was specified but few details were given of the mechanisms by which such schemas might be generated. In response to concerns that it was homuncular, Shallice and Burgess (1996) elaborated the SS (see **Figure 5**), specifying eight processes, several of which were held to operate in different phases. Comparison of these processes with the canonical framework reveals many parallels, as well as possible limitations of both approaches.

Consider first belief maintenance (canon 1). The SS model does not explicitly include processes related to perceptual input or maintenance of declarative knowledge. There is thus nothing akin to canon 1, but an elaboration of the SS would clearly require such processes. Indeed, a more recent description of the SS model (see Shallice and Cooper, 2011, figure 12.27) includes such processes within Method 2 of the construction of temporary new schemas (cf. **Figure 5**).

Raising goals and problem solving are addressed explicitly within the SS model. Canon 2 is implemented via the Problem Orientation Phase, and specifically by process 6 (Goal Setting). Canon 3, by contrast, subsumes three routes within the SS model, corresponding to the three methods by which temporary schemas are generated. Temporary schemas can be related to "options" in the domino model but in this case the CS/SS model provides an account of decision-making that elaborates upon the canonical framework. This should not be surprising given that the canonical framework is intended as an abstraction over theories of decision-making.

Canons 4 and 5 (constructing and aggregating reasons) are incorporated in processes 4 and 5. In particular the Solution Checking Phase of the SS must evaluate candidate solutions and reject them if they have insufficient merit. However, the SS model lacks a detailed description of how this evaluation might occur. Canons 4 and 5 provide an abstract specification of the necessary processes.

The output of stage 1 and input to stage 2 is a commitment to a plan (canon 6).

Monitoring, as described by canon 9, corresponds directly to process 2 of stage 3 of the SS model. However, within the SS model the product of monitoring is not any belief, but a specific belief, namely that the current schema is not achieving the intended goal. This signals rejection of the current schema (process 3 of **Figure 5**), a meta-cognitive process that triggers another round of processing.

At least two forms of learning are supported by the CS/SS model: accretion of episodic memories, and proceduralization of frequently generated temporary schemas. The two variants of signature Eq. S10 are applicable, on the assumption that episodic memories are equated with scenarios and proceduralized schemas are equated with tasks. These assumptions are consistent with the usage of these terms and brings to the fore the potential value of describing a cognitive architecture within the canonical framework: it forces us to be explicit about cognitive constructs (schemas, episodic memories, etc.) and the relationships between them.

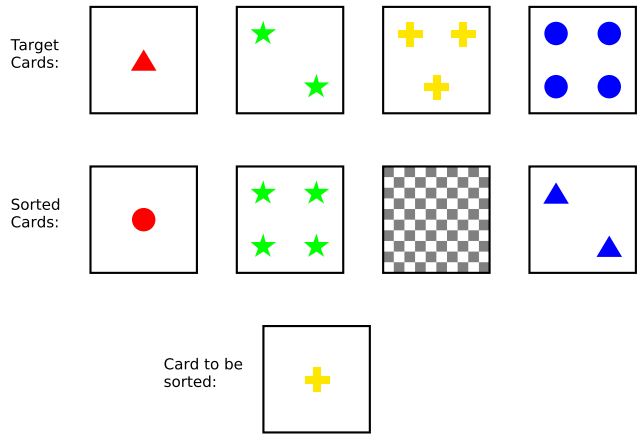


FIGURE 6 | The Wisconsin Card Sorting Test. The four target cards are shown across the top row and four piles for sorted cards in the second row, the third of which is currently empty. The card to-be-sorted is at bottom. If the subject is sorting according to color or form, this card should be placed under the third target card, but if he/she is sorting according to number then it should be placed under the first target card.

EXAMPLE: THE WISCONSIN CARD SORTING TEST

We now consider the CS/SS model and assess whether details of *dynamic* cognitive processing on a laboratory task might be consistent with the canonical theory as well as the static cognitive architecture. In the Wisconsin Card Sorting Test (WCST) subjects sort a series of colored cards using feedback provided by the experimenter to deduce the appropriate sorting criterion. Each card shows one to four shapes (triangles, stars, crosses, or circles) printed in red, green, yellow, or blue. Four "target" cards are positioned at the top of the work surface (see **Figure 6**). A series of cards is then presented and the subject is required to place each one under one of the targets, after which the experimenter indicates whether the choice was correct. For example, the experimenter might first select color as the sorting criterion, giving positive feedback if the subject's choice matches the color feature, but negative feedback otherwise. A complication is that after a series of correct choices the experimenter changes the sorting criterion without warning. The subject must therefore adapt his/her strategy throughout the task.

The WCST is not normally described as a decision-making task, though each trial requires a decision about where to place each successive card and the task as a whole has many features of DDM. In the standard presentation of the task, the cards to-be-sorted are in an ordered sequence and the first card shows one green triangle. Behavior on the task may be described as follows[6]:

[6]This decomposition of WCST processes is based closely on a computational model of WCST performance described by Glasspool and Cooper (2002).

Out the outset presenting a card to the subject places the subject in a situation for which no routine behavior exists. The SS is therefore invoked, raising the goal of sorting the card (Process 6 in **Figure 5**, or signatures Eqs S1 and S2 in the canonical theory). Problem solving (Eq. S3) can take at least two forms, depending on the subject's strategy.

1. The first strategy simply yields four decision options corresponding to the locations under the four target cards. Suppose the to-be-sorted card shows one green diamond. A choice can be made on three lines of reasoning: (1) place the to-be-sorted card under target 1 because the shapes match; (2) place it under target 1 because the numbers match; (3) place it under target 2 because colors match. Aggregating the reasons (Eq. S5) there are now two arguments for placing the card under target 1 and one for placing under target 2, yielding a preference for target 1 so the subject places it under card 1 (Eq. S6).
2. A more sophisticated application of Eq. S3 yields an alternative strategy. Here the different sorting rules (sort by shape, sort by number, sort by color) can themselves be considered as decision options, with Eq. S4 being instantiated by arguments for/against each rule. Initially there will be no specific arguments for or against any rule so adoption of a specific rule (Eq. S5) would be random. Once a sorting rule has been selected, a further processing cycle of processing would be required to apply the selected sorting rule to the current card.

Are there any principled grounds for choosing between the simplistic and the sophisticated approach? Within the CS/SS model this will depend on the *aspiration setting*: if a specific solution is required (as may occur on trial 1) then the simplistic approach will suffice, but if a general solution is required (as subsequent trials demand) then the more sophisticated approach will be necessary.

After the subject adopts a plan to place the card (Eq. S7) and executes the placement (Eq. S8) the experimenter provides feedback. This feedback may trigger another round of processing. Effective use of feedback is a substantial source of individual variability on the task. More able, "attentive" or "energized" subjects may detect a learning opportunity (Eq. S1) and raise a goal to translate the feedback into another strategy (Eq. S2), with the options being the three sorting rules (Eq. S3: shape, number, or color); cycling through Eqs S4–S6 will yield beliefs relating to the veracity of each of these rules. In this way the canonical framework is able to describe meta-cognition in processing feedback as well as basic decision-making.

Processing is slightly different on the second trial because the arguments for placing a card under a particular target will have different merit. If feedback on the previous trial was positive (and is correctly assimilated) arguments for sorting by the rule(s) that matched on that trial will be stronger, while arguments for sorting by the rule(s) that did not match will be weaker. This potentially allows the cognitive system to be more discriminating about the options and select the correct target card, though additional trials and feedback cycles may be necessary to eliminate all but one sorting rule. However, once the subject has narrowed down the possible sorting rules to one it is possible to anticipate positive feedback. That is, in applying Eqs S7 and S8, the subject also establishes a belief that subsequent feedback from the experimenter will be positive. This makes monitoring (Eq. S9) a computationally simple process: if the observed feedback differs from the expected feedback then the subject may infer that the sorting rule applied on the current trial is incorrect. This will count as an argument against that sorting rule on the following trial. Failure to set up this expectation or to take account of violations of the expectation will result in the subject continuing to sort cards by a previously appropriate rule in the face of continued negative feedback (i.e., so-called perseverative errors).

Once the correct sorting rule has been determined it is necessary to maintain a record of this rule (presumably in working memory) across trials so that it may be used to support the argument for placing each to-be-sorted card under the matching target card. This is an instance of belief maintenance (Eq. S1), which as discussed above is not explicitly included within the CS/SS model of Shallice and Burgess (1996).

THE CONTENTION SCHEDULING/SUPERVISORY SYSTEM MODEL AS AN INSTANCE OF THE CANONICAL THEORY

The CS/SS model of the human cognitive architecture is not as well developed as some theoretical accounts of the human cognitive architecture, but it is unique within cognitive psychology in being grounded in a domain-general view of cognitive function and being supported by findings from neuropsychological studies[7]. For example, with respect to the WCST, subjects occasionally make "set loss" errors, where their behavior suggests that, after correctly inferring the sorting rule (as evidenced by a run of correctly sorted cards), they spontaneously forget the rule. Such errors are particularly common in neurological patients with lesions affecting the inferior medial prefrontal cortex (Stuss et al., 2000), a region Shallice et al. (2008) associate with "attentiveness." Within the canonical framework, the subject's difficulty here may be understood as concerning a particular aspect of belief maintenance (Eq. S1), but one that relates to retaining existing beliefs, rather than to deriving new beliefs, e.g., interpreting observations.

Perseverative errors, in which subjects fail to switch sorting rules in the presence of sustained negative feedback, are common in the behavior of patients with prefrontal lesions. Stuss et al. (2000) suggest that the perseverative errors of patients with lesions in right dorsolateral prefrontal cortex are due to monitoring failure. However avoiding perseverative errors also involves switching away from a previously reinforced rule, a process discussed in the psychological literature under the rubric of "set shifting" or "task setting" which is frequently held to involve left dorsolateral prefrontal cortex. This process implements a form of Eq. S7 but the perseverative errors of different patient groups suggests there may be multiple underlying causes (Stuss et al., 2000), and as noted above, failures in monitoring (Eq. S9) may also result in such errors.

[7]Other cognitive architectures such as Soar (Newell, 1990) or ACT-R (Anderson, 2007) are grounded in specific areas of cognition (e.g., problem solving or associative memory) and view cognitive processing as equivalent to the operation of a production system.

The functional components of the CS/SS model are supported by a great deal of evidence from cognitive psychology, cognitive neuropsychology, and cognitive neuroscience. Shallice (2006), for example, reviews a number of neuropsychological studies in which patient behavior may be interpreted as a specific impairment in "the production of one or more procedures for attaining a goal" (i.e., Eq. S3, the generation of candidate options). Other studies, also reviewed by Shallice (2006), imply that processes related to checking that on-going processing or behavior is working toward ones current goals may also be selectively impaired (see also Shallice and Cooper, 2011). **Table 3** summarizes some of this evidence relating the processes of the CS/SS model to each of the signatures, drawing on further widely accepted views of the human cognitive architecture and its function in problem solving and decision-making.

ASSESSMENT 2: JOINT DECISION-MAKING BY AUTONOMOUS AGENTS

Understanding the foundations of autonomous operation of intelligent systems in complex, unpredictable environments is at the heart of AI. As discussed above it is a major focus of current research on software agents (e.g., Wooldridge, 2000; Fox et al., 2003; Poole and Mackworth, 2010). A major subfield of agent research is on *multi-agent* systems, in which autonomous agents interact to achieve common goals (Wooldridge, 2009). This field looks at models of how tasks can be shared between collaborating but individually autonomous agents, what forms of communication need to take place to achieve common objectives (such as informing, requesting, negotiating, persuading, and joint decision-making), and other cognitive functions.

We have carried out an initial assessment of the canonical theory by means of a computer simulation of a multi-agent decision-making task. The model has been built using the COGENT modeling tool, which is used to visualize the cognitive architectures of individual agents using an extended box-and-arrow notation, and implemented using rule-based and logic programing techniques. **Figure 7** shows two views of an agent network in which three agents interact with each other in order to make a simple medical decision. They share information about a hypothetical patient with chest pain and two of them jointly arrive at a treatment decision. In this diagram ellipses represent various kinds of data repository and rectangles are "compound" modules that can contain lower-level information processing components (see Cooper and Fox, 1998; Cooper et al., 2002 for more detail). Arrows indicate flow of information between modules.

The left panel shows a network of agents in which the three outer rectangles represent agents which communicate with each other through a "switchboard." A "patient records" agent provides information about patients with a specific medical problem (chest pain in this simulation). Agents C and S are the main actors in decision-making; agent C has cardiology knowledge and leads decision-making about the treatment of each patient, and agent S has specialist knowledge about safe drug use and can advise where there might be doubt.

The right panel shows the internal structure of agent C, which implements decision-making based on a version of the domino model extended with specialized stores for data and knowledge and processing modules which implement inter-agent communication and learning. Agent C, for example, has a set of data repositories which can be accessed by all processing modules: a *working memory* containing current cognitive states (beliefs, goals, plans etc); a *knowledge base* of domain facts, rules, and functions that are common to all agents (e.g., decision schemas, communication conventions), and specialist knowledge that is unique to each agent. Lastly, there is a knowledge base which contains a record of *past cases* and *learned knowledge* that can inform future decision-making. When this model is in operation working memory is constantly monitored by all the processing modules to determine whether any rules are applicable and, if so, the relevant cognitive state data are updated. These updates may lead to the conditions of other rules becoming instantiated, either within the same module (e.g., a new belief state propagates to further belief states) or another module (e.g., an updated belief state leads to an updated goal state).

EXAMPLE

This illustration[8] has been selected purely to illustrate the operation of the decision model and is not intended to be medically realistic. Some operational detail is omitted for clarity.

Phase 1

Agent C receives information from the patient records agent saying that Mrs. Smith is an elderly patient who has complained of chest pain. From its knowledge about such problems agent C infers the possibility of a heart attack and from this the more specific possibility of a myocardial infarction (MI). Agent C's knowledge indicates the need to "manage" the MI and a goal to achieve this is raised. Managing MI has in fact a number of aspects, including preventing blood clotting and pain and these are raised as subgoals. Agent C consults its knowledge base and finds two drugs that might satisfy these goals: clopidogrel and aspirin. Both are effective for analgesia and prevention of blood clotting and easily available, and aspirin is also modestly priced. These arguments are weighed up leading to a simple conclusion that aspirin is preferred. However MI is a dangerous condition so the decision to prescribe aspirin is qualified as *provisional*, meaning that the agent will not act on the basis of this preference but will carry out further investigation and, depending on the results of this investigation, it may abandon the tentative decision in favor of an alternative.

Phase 2

The presence of a provisional decision triggers a rule in the goal processing module of agent C to raise a goal to consult another agent, S, that has specialist knowledge about the safety of drugs. C's problem solver then retrieves a suitable interaction plan from its knowledge base. Communications are modeled using standard interaction "performatives" such as "inform," "explain," "query," "request," "instruct," etc. **Figure 8** summarizes the dialog that follows.

Agent C first sends a *request* to agent S to enter into an *inform* dialog, meaning that the two agents should follow a particular protocol for providing information. The request is accepted by S, so C sends a *request* for confirmation that aspirin is appropriate for

[8]The illustration is based on an example developed by Sanjay Mogul.

Table 3 | Some relationships between the canonical functions and selected evidence from cognitive psychology and cognitive neuroscience.

Signature	Summary
S1 (belief maintenance)	Beliefs may be supported by the environment (i.e., inferred from perceptual input) or inferred from long-term knowledge and other beliefs. Both must be actively maintained in working memory (e.g., by rehearsal)
S2 (raising goals)	Much behavior, with the possible exception of habitual behavior, can be understood as being purposive or goal-directed. In experimental psychology, high-level task goals are set by the experimenter, with subjects deriving lower-level goals for individual trials. Findings from experimental psychology and more generally indicate that goals provide local coherence of behavior
S3 (problem solving)	A variety of problem solving strategies or heuristics may be recruited to generate solutions for a given goal. This includes so-called "weak" methods which are general, knowledge-lean, heuristics such as hill-climbing and means-ends analysis, as well as knowledge-rich, task-specific strategies, acquired through experience
S4 (reasons for decisions)	Evolutionary arguments (e.g., Mercier and Sperber, 2011) suggest that argumentation is central to human decision-making. According to this view, generating arguments for or against propositions is an essential step in persuading others
S5 (aggregation)	One neuropsychological hypothesis is that aggregation of the merit of arguments is based on somatic markers – emotionally biased valences associated with decision options acquired through positive and negative experience (Damasio, 1994). Damasio relates the association of somatic markers with candidates to the amygdala and ventromedial prefrontal cortex
S6 (commitment)	Commitment to a single decision candidate is required by theories such as Damasio's somatic marker hypothesis. In the specific context of selecting one word from a set, commitment has been related to the inferior frontal gyrus (Shallice and Cooper, 2011, Section 9.13)
S7 (plan enactment)	Plan enactment is most closely related to the function of task setting, held by many to be a function of left lateral prefrontal cortex (e.g., Shallice et al., 2008)
S8 (action)	The contention scheduling system provides an account of how intentions are mapped to actions, subject to available resources
S9 (monitoring)	A substantial body of evidence suggests that many cognitive processes create expectations that under normal operation are continuously monitored. Perceptual processes may also monitor the external environment for deviations from expected perceptual input. Shallice et al. (2008) relate monitoring to right dorsolateral prefrontal cortex, though an alternative view is that anterior cingulate cortex compares expectations with observations, generating an error signal when there is a mismatch
S10 (learning)	There are many forms of learning. One is learning to associate consequences with cognitive and motor actions. These consequences then become expectations which are used by monitoring. A second critical form is reinforcement learning, where positive or negative reward can increase or decrease the merit of a candidate in the context of a goal

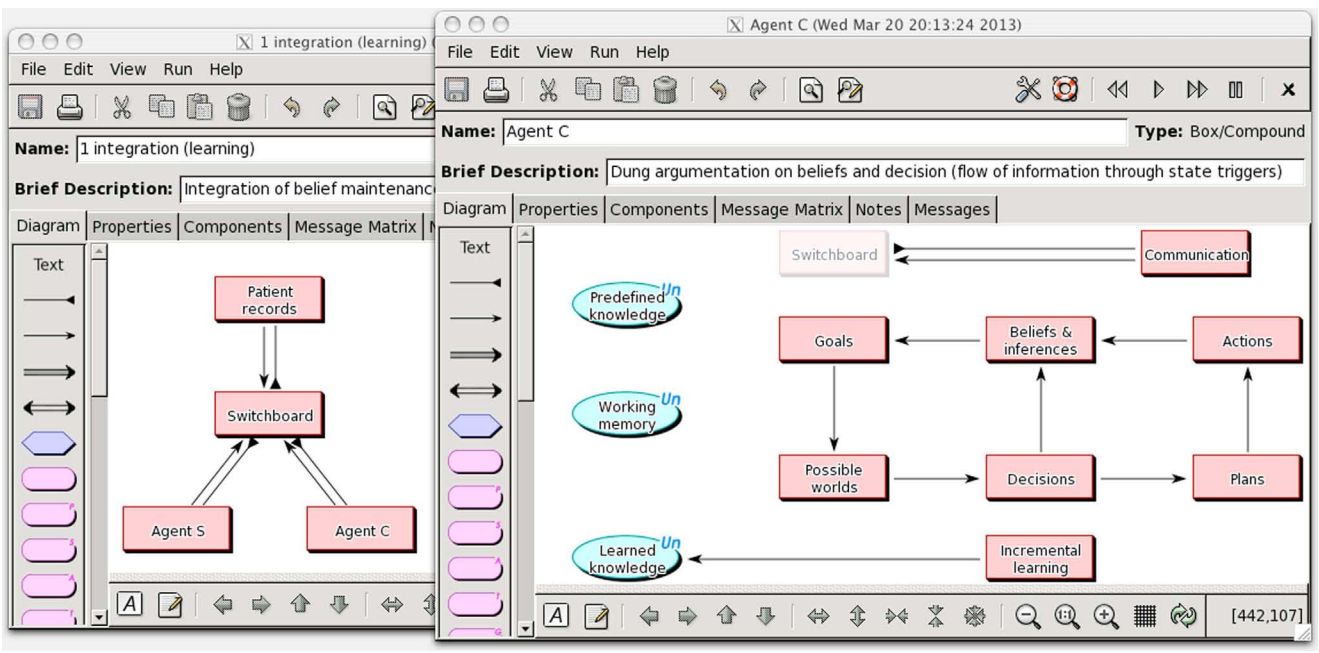


FIGURE 7 | Multi-agent network for cooperative decision-making (left), and information processing architecture for an autonomous agent (right).

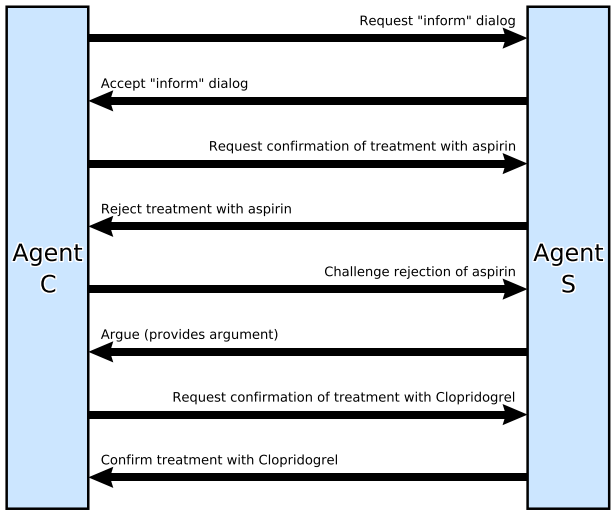


FIGURE 8 | Sequence diagram for some of the interactions between agent C and agent S during the joint decision-making simulation.

Mrs. Smith. S then goes through its own decision process: raising a goal to decide on the best prevention of pain and clotting and generating a set of treatment options for Mrs. Smith. Its specialist knowledge of treatments for MI indicates aspirin, clopidogrel, and a further option: proton-pump inhibitors. S now proceeds to construct arguments for and against all three options. During this process it applies an argument schema that "If a treatment is proposed and is known to exacerbate a condition, and the patient has that condition, then this constitutes an argument against the proposed treatment." Agent S has domain knowledge that aspirin exacerbates gastritis, and (after a *request* for information from the patient record agent) it finds that Mrs. Smith has gastritis so this yields an argument against prescribing aspirin. S *informs* C that clopidogrel is therefore preferred.

Phase 3

Agent C raises a goal to understand the rationale for the Agent A's advice. One way of doing this is to "challenge" the advice in order to elicit the reasons for the recommendation. Agent S reflects on its rationale and provides an explanation as a set of arguments in an *inform* message. This leads to Agent C adding another argument against aspirin to working memory and clopidogrel is now the preferred option for Agent C as well.

Phase 4

As a final part of this experiment we implemented a simple learning mechanism for acquiring new knowledge which can be used in future decisions based on a record of the whole episode. An episodic record of the decision includes the goal that was active and the set of beliefs that held when a decision was committed. Another learning mode was also demonstrated by recording how frequently particular decisions are taken in particular contexts, which can be used to weight options in future decisions.

THE AGENT ARCHITECTURE AS AN INSTANCE OF THE CANONICAL THEORY

The agent architecture owes much to the domino framework so there is naturally a good mapping between this implementation and the canonical framework (**Table 4**). Phase 1 traverses Eqs S1–S6, and phase 2 traverses Eqs S2–S8. Communication is modeled by means of goal-based dialogs represented as plans (Eq. S6), and individual communication acts are instances of the action signature Eq. S8.

The signatures specify general constraints on the inputs and outputs of each information processing component of the agent architecture. At this abstract level each component is a black box and may be implemented in any number of ways: as a conventional algorithm, a set of production rules, a logic program, in hardware, or in some other way. In the COGENT implementation the signatures are translated into eight sets of specialized production rules, one set associated with each component.

ASSESSMENT 3: DESIGNING DECISION SYSTEMS

Our research was originally motivated by a wish to understand cognitive processes that underpin human judgment and to apply this understanding by developing a technology for designing decision-making systems for use in dynamic, complex, and safety critical settings (Fox and Das, 2000). Taking medicine as a concrete focus we established three engineering requirements.

SCOPE

The technology should be able to model any type of decision, represented by the corpus of decisions common in clinical practice (e.g., hazard detection, risk assessment, test selection, referral, diagnosis and treatment decisions, and many others). The framework should also cover the full lifecycle of decision-making from the point where the need for a decision is established to the point where a choice can be made. A general *symbolic decision procedure*[9] was developed for this purpose (Fox et al., 1990; Fox and Krause, 1991; Huang et al., 1993) which later evolved into the domino model.

REALISM

Applications must cope with the constant change and high-levels of uncertainty typical of real world environments, where quantitative decision models are impractical (e.g., due to lack of data) and which naturalistic decision models seek to address. A key proposal was a preference and choice model based on logical argumentation (Fox et al., 1993; Krause et al., 1995) which has features in common with reason-based choice in cognitive psychology (e.g., Shafir et al., 1993), formal argumentation for decision-making in AI (e.g., Amgoud and Prade, 2009) and theories of argumentation in social and evolutionary psychology (e.g., Mercier and Sperber, 2011).

[9]"A symbolic decision procedure can be characterized as an explicit representation of the knowledge required to define, organize and make a decision, and . . . a logical abstraction from the qualitative and quantitative knowledge that is required for any specific application. A SDP may include a specification of when and how the procedure is to be executed" (Fox and Krause, 1991, p. 106).

Table 4 | The relation between the canon signatures and functions which are implemented in the multi-agent decision-making scenario.

Signature	Summary
S1 (belief maintenance)	Any rule in the agent model can make inferences by applying knowledge to the current working memory state and add, delete or replace information in the working memory. Every item of data in working memory is tagged with the grounds for believing it (e.g., the goal and assumptions which justify it). It uses this to maintain a consistent overall belief state
S2 (raising goals)	Goals are a form of belief which are used to determine which knowledge and rules are potentially in play at any moment
S3 (problem solving)	Any kind of problem solving technique can be implemented in the COGENT programing system, with the solution then added to working memory
S4 (reasons for decisions)	A form of argumentation based on defeasible logic is used to generate and maintain arguments for competing solutions as the working memory belief state changes
S5 (aggregation)	In the multi-agent decision-making scenario a simple improper linear aggregation function is implemented (adding up pros and cons) though other aggregation functions can be implemented
S6 (commitment)	The multi-agent scenario includes two kinds of commitment, provisional (reversible), and firm (irreversible)
S7 (plan enactment)	Dialog plans are simple lists of communication actions that are executed in sequence but can be interrupted if a communication is received from another agent
S8 (action)	The main kinds of actions that are included in this demonstration are standard communication performatives from speech act theory and agent communication languages
S9 (monitoring)	The whole domino system is a kind of "monitor" in that every computational component can respond to any update to the working memory state at any time
S10 (learning)	Two simple learning mechanisms have been implemented. These monitor the working memory and when a decision process terminates these mechanisms (1) add rules to the agent's episodic knowledge and (2) update frequency counters which can be used to update the agent's confidence in competing decision options

IMPLEMENTABILITY

The practical development of decision support services requires an expressive implementation language for modeling and implementing decisions and other tasks. The symbolic decision procedure and argumentation model proved to be an effective foundation for a practical decision modeling language (Das et al., 1997; Fox and Das, 2000), the most developed version of which is a published standard (Sutton and Fox, 2003). *PROforma*[10] has proved to be capable of modeling a wide range of decision processes in a way that clinicians find natural to understand and use. It has been used to deploy many clinical applications which are in routine use (e.g., the NHS Direct triaging service in the UK[11]; support for decision-making by multidisciplinary teams, Patkar et al., 2012).

PRO*forma* reifies the logical processes of the domino into a *task model*. It is a knowledge representation language (for modeling expertise) and a programing language (for implementing decision support systems and autonomous agents). A typical PRO*forma* model is a network of decisions, plans and other tasks which can be enacted in a predefined sequence, or concurrently, or in response to circumstances. Two simple example networks are shown in **Figure 9**.

The first example starts with an "enquiry" (any data acquisition process, shown as a diamond). This may acquire data from many sources (e.g., database, a sensor or other device or querying a human user or another agent). The decision (circle) that follows the enquiry can be taken only when the enquiry has completed, which is specified by the connecting arrow. The decision applies relevant knowledge to interpret the data that has been acquired up to this point by constructing and assessing arguments for and against the various decision options. One option here is a simple action (e.g., send the patient home) while the other is a plan (e.g., a course of treatment).

The second example captures a decision-making pattern called SOAP (for "Subjective"; "Objective"; "Assessment"; "Plan") which is a mnemonic familiar to clinicians that refers to the routine process of taking a patient history, deciding what to do, and then doing it. There are two enquiries in this example; one acquires information about the patient's subjective complaint and experience while the other captures objective data such as the patient's height, weight, and blood pressure. It does not matter in what order the data are acquired but the *Assess* decision may not be taken until both enquiry tasks have been done.

PROforma AS AN INSTANCE OF THE CANONICAL THEORY

As shown in **Table 5** the *PROforma* language instantiates signatures Eqs S1–S8 in the canonical framework, though different interpreters for the language implement some details differently.

SUMMARY AND DISCUSSION

We have briefly examined traditional perspectives on decision-making, including decision theory (prescriptive models grounded

[10] A contraction of *process* and *formalization*.
[11] http://www.nhsdirect.nhs.uk/CheckSymptoms.aspx

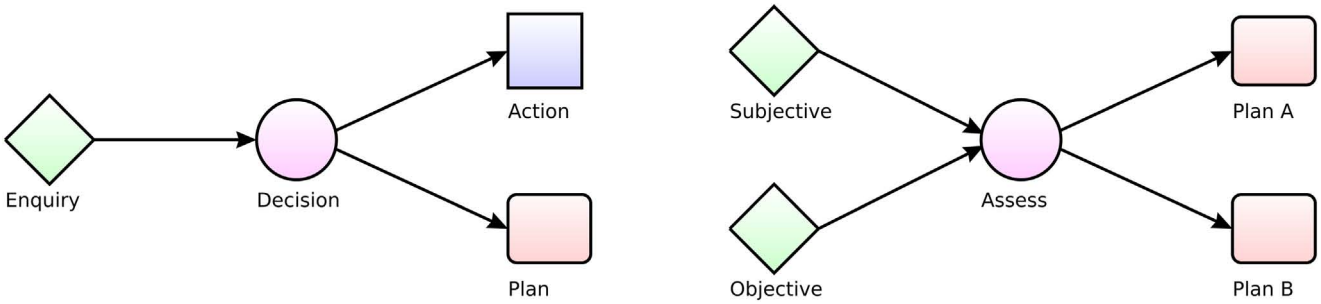


FIGURE 9 | Simple *PROforma* task networks (The Tallis decision support development software can be downloaded from www.cossac.org/tallis for research use).

Table 5 | The relation between the canon signatures and task representations in the *PROforma* modeling language.

Signature	Summary
S1 (belief maintenance)	Beliefs in *PROforma* are data derived from the external environment (e.g., "age = 54") or indirectly by inference or decision-making ("diagnosis = ulcer") and also "meta-data" (e.g., "diagnosis = completed"). If the data change then beliefs can be automatically updated, e.g., order of preference over a set of options in a decision that is currently in progress
S2 (raising goals)	A *PROforma* goal is a logical combination (and/or/not) of situation descriptions which do not currently hold. A goal can be raised by any kind of task; the task will be terminated if the goal descriptions become true
S3 (problem solving)	Current *PROforma* interpreters are limited to retrieving a set of options from a knowledge base
S4 (reasons for decisions)	Reasons in *PROforma* are *logical arguments* that represent *evidential arguments* in deciding between hypotheses and *preference arguments* when deciding between actions
S5 (aggregation)	A decision assesses all the argument for and against each option to determine their net overall force, and establish an order of preference over the options. The prior confidence and strength of arguments can be taken into account in the aggregation process
S6 (commitment)	Each option in a *PROforma* decision may include a rule which defines the conditions in which the option can be "recommended" when the application is supporting a third party decision, or automatically "committed" if the system is configured to operate autonomously
S7 (plan enactment)	A *PROforma* plan is a network of tasks, in which the scheduling of decisions and other tasks can be predefined or determined dynamically
S8 (action)	When a *PROforma* action is scheduled for execution it first checks any preconditions (such as beliefs or goals being true, resources being available)
S9 (monitoring)	There is no specific support for monitoring. However continuous monitoring can be implemented using general language features
S10 (learning)	The *PROforma* language standard does not currently support learning

in rational axioms); decision science (descriptive, empirically grounded theories of decision-making), and decision engineering (techniques for supporting human decision-makers and developing autonomous decision agents). The central motivation for developing a canonical theory is to provide a *lingua franca* for discussions between researchers, and with practitioners, based on a common set of intuitive but well-defined concepts and processes. This project will be successful if members of community A find the framework sufficiently versatile and clear for describing their view of decision-making to members of community B and vice versa, and if productive conversations ensue.

The canonical framework developed here does not fit squarely into any one of the traditional paradigms. The canons are neither normative nor descriptive; we have used the term "requisite" elsewhere (Fox, 1981)[12]. The canons say that any general decision procedure must address certain functional requirements in some way (maintaining beliefs, raising goals, making commitments, and so forth). However we are not imposing any particular way in which the canons must be implemented. That is, the canons are framed at the highest of Marr's levels (Marr, 1982): they specify what function should be computed but make no commitment to the algorithms involved or the representation of information over

[12]The term derives from Ashby's Law of Requisite Variety (Ashby, 1958), which concerns the regulation of the behavior of a system, R. The law states that "R's capacity as a regulator cannot exceed its capacity as a channel for variety" (Ashby, 1958, p. 86). The law is "requisite" in the sense that it states a requirement on the variety possible within the system R.

which they operate. This is in contrast with a normative theory like expected utility theory that commits the implementer to update belief in a way that is constrained by the probability axioms and measures of value must satisfy "rational" axioms like transitivity and so forth. Our only claim is that any proposal for a specific theory of DDM must implement some or all of the canons in some way.

A clear limitation of our program, consequently, is that the canonical theory does not address the particular concerns of each decision research community in detail. Psychological mechanisms are insufficiently specified in the canonical signatures to make predictions about how human decision-makers actually behave or how human performance differs from prescriptive norms. Nor can the canons be claimed to be axioms of rational inference, such as those offered by statistical decision theory or mathematical logic. Lastly the canonical form does not offer tools for designing practical applications. Nevertheless a canonical framework may have benefits for specialist researchers in all three traditions in that the general canons can be instantiated for particular purposes by specific procedures or mechanisms. We close with a short discussion of some of the benefits that the framework may offer to theorists, scientists, and engineers.

CONTRIBUTIONS TO DECISION THEORY

1. General canons of cognition help to promote discussion between communities with different theoretical commitments. The belief maintenance canon Eq. S1 for example can be instantiated by probabilistic inference, fuzzy logic, default reasoning, and so on. These are often seen as competitors but in our view it would be more helpful to see them as alternative ways of modeling uncertain reasoning to address different requirements and constraints.
2. The canons offer a broader context within which to investigate formal theories of decision-making than is usual. Classical expected utility theory, for example, "can give no scientific advice" about when a decision is needed or what is relevant to framing it (Lindley, 1985). This seriously limits the scope of current theory and the canonical framework suggests a number of ways in which normative models could be extended.

CONTRIBUTIONS TO DECISION SCIENCE

1. Canonical forms can be used to provide functional explanations of behavioral, clinical, or neurological data obtained in decision-making tasks and to map between neuro-anatomical organization and cognitive-level functions (Shallice and Cooper, 2011).
2. The theory bridges "folk" psychology, common sense reasoning, agent theories in AI and philosophical theories of mind, and potentially engages with the vocabulary of the humanities and everyday discourse.
3. The canonical theory provides a framework that may help to resolve debates which arise from different assumptions about methodology. A current dispute in cognitive science, for example, concerns whether psychological theories are constrained by neuroscience data and vice versa. Coltheart (2006) has argued that a *true psychological theory* exists at such a level of abstraction that data about the neurological implementation of cognition cannot in principle confirm or refute the theory. The abstract cognitive theory that Coltheart seeks is at the canonical level – it can inform psychological theory at a functional level but need not confront implementation details which are particular to human cognition.

CONTRIBUTIONS TO DECISION ENGINEERING

1. The abstract signatures are insufficiently specified to be directly computable but there are clearly many specific algorithms that will take, as input, data of the types specified "above the line" and generate, as output, data of the types specified "below the line"[13]. Rather than just making the theory too vague to be useful this has the practical implication that we could design and implement decision-making systems by assembling them out of standard components which comply with the signatures without knowledge of internal mechanisms.
2. Responsibility for practical decision-making is often distributed across professional teams and in the future such teams will be increasingly supported by automated services. In an emergency management system for example distinct subsystems will be responsible for capturing data while others will be required to integrate data from multiple sources and "judge its sense, meaning, relevance, and reliability; decide what the options for action are and make effective decisions" (Carver and Turoff, 2007, p. 34). Faced with this complexity designers will wish to engineer systems using standard modules for decision-making, planning, communication etc. and a canonical model offers a way of specifying and linking such modules.

CONCLUSION

In this paper we have considered the whole cycle of DDM: recognizing and framing a problem in light of current beliefs; clarifying and prioritizing goals; generating options that would achieve current goals; evaluating preferences over the options; and aggregating preferences to select the best. We have not sought to develop any new theory related to the specifics of any one of these capabilities. Rather, our approach has been to develop an over-arching framework that subsumes specific theories of these individual subprocesses and understand how they are related to each other.

In our view theoretical understanding of the processes involved in human DDM is advancing, but in a somewhat chaotic way in which many competing research traditions, theoretical concepts, and engineering techniques vie for pre-eminence. We hope that our discussion has demonstrated that there is potential for establishing a cross-disciplinary framework that promotes constructive discussion between communities, leading to collaboration, and even synergy rather than competition.

[13] Each signature can be thought of as a pair of colored sockets into which standard cognitive components can be plugged if and only if colors of the plugs match colors of the sockets.

REFERENCES

Amgoud, L., and Prade, H. (2009). Using arguments for making and explaining decisions. *Artif. Intell.* 173, 413–436.

Anderson, J. R. (2007). *How Can the Human Mind Exist in the Physical Universe?*. New York: Oxford University Press.

Ashby, W. R. (1958). Requisite variety and its implications for the control of complex systems. *Cybernetica* 1, 83–99.

Bratman, M. E. (1987). *Intentions, Plans and Practical Reason*. Cambridge, MA: Harvard University Press.

Brehmer, B. (1992). Dynamic decision making: human control of complex systems. *Acta Psychol. (Amst.)* 81, 211–241.

Carver, L., and Turoff, M. (2007). Human-computer interaction: the human and computer as a team in emergency management information systems. *Communications of the ACM – Emergency Response Information Systems: Emerging Trends and Technologies* 50, 33–38.

Churchland, P. (1981). Eliminative materialism and the propositional attitudes. *J. Philos.* 78, 67–90.

Clancey, W. J. (1985). Heuristic classification. *Artif. Intell.* 27, 289–350.

Cohen, P. R., and Levesque, H. (1990). Intention is choice with commitment. *Artif. Intell.* 42, 213–261.

Coltheart, M. (2006). What has functional neuroimaging told us about the mind (so far)? *Cortex* 42, 323–331.

Cooper, R. P. (2007). The role of falsification in the development of cognitive architectures: insights from a Lakatosian analysis. *Cogn. Sci.* 31, 509–533.

Cooper, R. P., and Fox, J. (1998). COGENT: a visual design environment for cognitive modelling. *Behav. Res. Methods, Instrum. Comput.* 30, 553–564.

Cooper, R. P., Fox, J., Farringdon, J., and Shallice, T. (1996). A systematic methodology for cognitive modelling. *Artif. Intell.* 85, 3–44.

Cooper, R. P., Yule, P., and Fox, J. (2003). Cue selection and category learning: a systematic comparison of three theories. *Cogn. Sci. Quarterly* 3, 143–182.

Cooper, R. P., Yule, P. G., Fox, J., and Glasspool, D. W. (2002). *Modelling High-Level Cognitive Processes*. Mahwah: Lawrence Erlbaum Associates.

Damasio, A. R. (1994). *Descarte' Error: Emotion, Reason, and the Human Brain*. New York: Grosset/Putnam.

Das, S. K., Fox, J., Elsdon, D., and Hammond, P. (1997). A flexible architecture for autonomous agents. *J. Exp. Theor. Artif. Intell.* 9, 407–440.

Drummond, H. (1991). *Effective Decision Making: A Practical Guide for Management*. London: Kogan Page.

Edwards, W. (1962). Dynamic decision theory and probabilistic information processing. *Hum. Factors* 4, 59–73.

Fox, J. (1980). Making decisions under the influence of memory. *Psychol. Rev.* 87, 190–211.

Fox, J. (1981). Towards a reconciliation of fuzzy logic and standard logic. *Int. J. Man. Mach. Stud.* 15, 213–220.

Fox, J. (2011). "Arguing about the evidence," in *Evidence, Inference and Enquiry*, eds P. Dawid, W. Twining and M. Vasilaki (Oxford: Oxford University Press), 151–182.

Fox, J., Beveridge, M., and Glasspool, D. W. (2003). Understanding intelligent agents: analysis and synthesis. *AI Commun.* 16, 139–152.

Fox, J., Clark, D. A., Glowinski, A. J., and O'Neil, M. J. (1990). Using predicate logic to integrate qualitative reasoning and classical decision theory. *IEEE Trans. Syst. Man Cybern.* 20, 347–357.

Fox, J., and Das, S. (2000). *Safe and Sound: Artificial Intelligence in Hazardous Applications*. Cambridge: American Association for Artificial Intelligence and The MIT Press.

Fox, J., Glasspool, D. W., Patkar, V., Austin, M., Black, L., South, M., et al. (2010). Delivering clinical decision support services: there is nothing as practical as a good theory. *J. Biomed. Inform.* 43, 831–843.

Fox, J., Krause, P., and Elvang-Gøransson, M. (1993). "Argumentation as a general framework for uncertain reasoning," in *Proceedings of the Ninth Conference Annual Conference on Uncertainty in Artificial Intelligence (UAI-93)*. San Francisco: Morgan Kaufmann.

Fox, J., and Krause, P. J. (1991) "Symbolic decision theory and autonomous systems" in *UAI'91 Proceedings of the Seventh Conference on Uncertainty in AI*. San Francisco: Morgan Kaufman, 103–111.

Gigerenzer, G., and Todd, P. M. (2000). *Simple Heuristics that Make us Smart*. New York: Oxford University Press.

Glasspool, D. W. (2005). "The integration and control of behaviour: insights from neuroscience and AI," in *Visions of mind: Architectures for Cognition and Affect*, ed. D. N. Davis (Hershey, PA: IDEA Group), 208–234.

Glasspool, D. W., and Cooper, R. P. (2002). "Executive processes," in *Modelling High-Level Cognitive Processes*, Chap. 8, ed. R. P. Cooper (Mahwah: Lawrence Erlbaum Associates).

Huang, J., Fox, J., Gordon, C., and Jackson-Smale, A. (1993). Symbolic decision support in medical care. *Artif. Intell. Med.* 5, 415–430.

Kahneman, D., and Tversky, A. (1979). Prospect theory: an analysis of decision under risk. *Econometrica* XLVII, 263–291.

Klein, G. (2008). Naturalistic decision making. *J. Hum. Fact. Ergon. Soc.* 50, 456–460.

Kohn, L. T., Corrigan, J. M., and Donaldson, M. S. (eds). (2000). *To Err Is Human: Building a Safer Health System. Institute of Medicine*. Washington, DC: National Academy Press.

Krause, P., Ambler, S., Evlang-Goransson, M., and Fox, J. (1995). A logic of argumentation for reasoning under uncertainty. *Comput. Intell.* 11, 113–131.

Laird, J. E. (2012). *The Soar Cognitive Architecture*. Cambridge: MIT Press.

Laird, J. E., Newell, A., and Rosenbloom, P. S. (1987). SOAR: an architecture for general intelligence. *Artif. Intell.* 33, 1–64.

Lakatos, I. (1970). "Falsification and the methodology of scientific research programmes," in *Criticism and the Growth of Knowledge*, eds I. Lakatos and A. Musgrave (Cambridge: Cambridge University Press), 91–195.

Lindley, D. V. (1985). *Making Decisions*. London: John Wiley.

Marr, D. (1982). *Vision*. San Francisco, CA: Freeman.

Mercier, H., and Sperber, D. (2011). Why do humans reason? Arguments for an argumentative theory. *Behav. Brain Sci.* 34, 57–111.

Newell, A. (1981). The knowledge level. (Presidential address). *AI Mag.* 2, 1.

Newell, A. (1990). *Unified Theories of Cognition*. Cambridge: Harvard University Press.

Newell, A., Shaw, J. C., and Simon, H. A. (1958). Elements of a theory of human problem solving. *Psychol. Rev.* 65, 151–166.

Newell, A., and Simon, H. A. (1972). *Human Problem Solving*. Englewood Cliffs, NJ: Prentice-Hall.

Norman, D. A., and Shallice, T. (1986). "Attention to action: willed and automatic control of behaviour," in *Consciousness and Self Regulation*, Vol. 4. eds R. Davidson, G. Schwartz, and D. Shapiro (New York: Plenum), 1–18.

Patkar, V., Acosta, D., Davidson, T., Jones, A., Fox, J., and Keshtgar, M. (2012). Using computerised decision support to improve compliance of cancer multidisciplinary meetings with evidence-based guidance. *BMJ Open* 2, pii: e000439.

Pennington, N., and Hastie, R. (1993). Reasoning in explanation-based decision making. *Cognition* 49, 123–163.

Poole, D. L., and Mackworth, A. K. (2010). *Artificial Intelligence: Foundations of Computational Agents*. Cambridge: Cambridge University Press.

Rao, S., and Georgeff, M. P. (1995). "BDI-agents: from theory to practice," in *Proceedings of the First International Conference on Multi-agent Systems (ICMAS'95)*, ed. V. Lesser (Menlo Park: The MIT Press), 312–319.

Ratcliff, R., and McKoon, G. (2008). The diffusion decision model: theory and data for two-choice decision tasks. *Neural Comput.* 20, 873–922.

Shafir, E., Simonson, A., and Tversky, A. (1993). Reason-based choice. *Cognition* 49, 11–36.

Shallice, T. (2006). "Contrasting domains in the control of action: the routine and the non-routine," in *Attention and Performance XXI: Processes of Change in Brain and Cognitive Development*, eds Y. Munakata and M. Johnson (Oxford: Oxford University Press), 3–29.

Shallice, T., and Burgess, P. W. (1996). The domain of supervisory processes and temporal organisation of behaviour. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 351, 1405–1412.

Shallice, T., and Cooper, R. P. (2011). *The Organisation of Mind*. Oxford: Oxford University Press.

Shallice, T., Stuss, D. T., Picton, T. W., Alexander, M. P., and Gillingham, S. (2008). Mapping task switching in frontal cortex through neuropsychological group studies. *Front. Neurosci.* 2, 79–85. doi:10.3389/neuro.01.013.2008

Shanteau, J. (1987). "Psychological characteristics of expert decision makers," in *Expert Judgment and Expert Systems*, eds J. L. Mumpower, O. Renn, L. D. Phillips and V. R. R. Uppuluri (Berlin: Springer-Verlag), 289–304.

Simon, H. (1957). *Administrative Behavior*, 2nd Edn. New York: MacMillan.

Spivey, J. M. (1989). *The Z Notation: A Reference manual*. Hemel Hempstead: Prentice Hall International Series in Computer Science.

Stearns, M. Q., Price, C., Spackman, K. A., and Wang, A. Y. (2001). "SNOMED clinical terms: overview of the development process and project status," in *Proceedings of the AMIA Symposium, American Medical Informatics Association*, ed. S. Bakken (Philadelphia: Hanley and Belfus), 662–666.

Stuss, D. T., Levine, B., Alexander, M. P., Hong, J., Palumbo, C., Hamer, L., et al. (2000). Wisconsin card sorting test performance in patients with focal frontal and posterior brain damage: effects of lesion location and test structure on separable cognitive processes. *Neuropsychologia* 38, 388–402.

Sutton, D., and Fox, J. (2003). The syntax and semantics of the PRO*forma* guideline modelling language. *J. Am. Med. Inform. Assoc.* 10, 433–443.

Tversky, A., and Kahneman, D. (1974). Judgment under uncertainty: heuristics and biases. *Science* 185, 1124–1131.

Wooldridge, M. (2000). *Reasoning about Rational Agents*. Cambridge: The MIT Press.

Wooldridge, M. (2009). *An Introduction to Multi Agent Systems*, 2nd Edn. Chichester: John Wiley and Sons.

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: 21 June 2012; paper pending published: 12 July 2012; accepted: 08 March 2013; published online: 02 April 2013.

Citation: Fox J, Cooper RP and Glasspool DW (2013) A canonical theory of dynamic decision-making. Front. Psychol. ***4****:150. doi: 10.3389/fpsyg.2013.00150*

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