frontiers
RESEARCH TOPICS

MENTAL IMAGERY

Topic Editors
Joel Pearson and Stephen M. Kosslyn

frontiers in
PSYCHOLOGY

frontiers in
HUMAN NEUROSCIENCE

Cover image provided by Ibbl sarl,
Lausanne CH

ISSN 1664-8714
ISBN 978-2-88919-149-9
DOI 10.3389/978-2-88919-149-9

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.

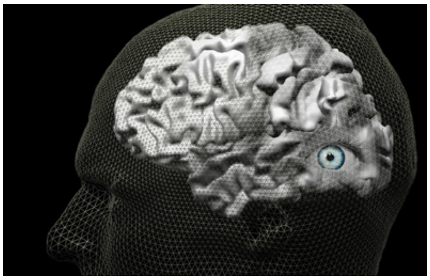
WHAT ARE FRONTIERS RESEARCH TOPICS?

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

MENTAL IMAGERY

Topic Editors:
Joel Pearson, The University of New South Wales, Australia
Stephen M. Kosslyn, Stanford University**,** USA



Our ability to be conscious of the world around us is often discussed as one of the most amazing yet enigmatic processes under scientific investigation today. However, our ability to imagine the world around us in the absence of stimulation from that world is perhaps even more amazing. This capacity to experience objects or scenarios through imagination, that do not necessarily exist in the world, is perhaps one of the fundamental abilities that allows us successfully to think about, plan, run a dress rehearsal of future events, re-analyze past events and even simulate or fantasize abstract events that may never happen.

Empirical research into mental imagery has seen a recent surge, due partly to the development of new neuroscientifc methods and their clever application, but also due to the increasing discovery and application of more objective methods to investigate this inherently internal and private process.

As this topic is being cross-hosted in both Frontiers in Human Neuroscience and Frontiers in Perception Science, we invite researchers from different fields to submit opinionated but balanced reviews, new empirical, theoretical, philosophical or technical papers covering any aspect of mental imagery. In particular, we encourage submissions focusing on different sensory modalities, such as olfaction, audition somatosensory etc. Similarly, we support submissions focusing on the relationship between mental imagery and other neural and cognitive functions or disorders such as visual working memory, visual search or disorders of anxiety.

Together, we hope that collecting a group of papers on this research topic will help to unify theory while providing an overview of the state of the field, where it is heading, and how mental imagery relates to other cognitive and sensory functions.

Table of Contents

05 ***Mental Imagery***
Joel Pearson and Stephen M. Kosslyn

06 ***Hemispheric Differences with in the Fronto-Parietal Network Dynamics Underlying Spatial Imagery***
Alexander T. Sack and Teresa Schuhmann

16 ***Unmasking the Perky Effect: Spatial Extent of Image Interference on Visual Acuity***
Adam Reeves and Catherine Craver-Lemley

23 ***Training Visual Imagery: Improvements of Metacognition, But not Imagery Strength***
Rosanne L. Rademaker and Joel Pearson

34 ***New Percepts via Mental Imagery?***
Fred W. Mast, Elisa M. Tartaglia and Michael H. Herzog

39 ***An Emerging Paradigm: A Strength-Based Approach to Exploring Mental Imagery***
Tadhg E. MacIntyre, Aidan P. Moran, Christian Collet and Aymeric Guillot

51 ***Vividness of Visual Imagery and Incidental Recall of Verbal Cues, When Phenomenological Availability Reflects Long-Term Memory Accessibility***
Amedeo D'Angiulli, Matthew Runge, Andrew Faulkner, Jila Zakizadeh, Aldrich Chan and Selvana Morcos

69 ***The Effects of Visual Imagery on Face Identification: An Erp Study***
Jianhui Wu, Hongxia Duan, Xing Tian, Peipei Wang and Kan Zhang

77 ***Electrophysiological Potentials Reveal Cortical Mechanisms for Mental Imagery, Mental Simulation, and Grounded (Embodied) Cognition***
Haline E. Schendan and Giorgio Ganis

99 ***A Cross-Modal Perspective on the Relationships Between Imagery and Working Memory***
Lora T. Likova

113 ***Verbal to Visual Code Switching Improves Working Memory in Older Adults: An fMRI Study***
Mariko Osaka, Yuki Otsuka and Naoyuki Osaka

121 ***Mental Imagery for Musical Changes in Loudness***
Freya Bailes, Laura Bishop, Catherine J. Stevens and Roger T. Dean

130 ***Imagining is Not Doing but Involves Specific Motor Commands: A Review of Experimental Data Related to Motor Inhibition***
Aymeric Guillot, Franck Di Rienzo, Tadhg MacIntyre, Aidan Moran and Christian Collet

152 ***Mental Imagery of Speech: Linking Motor and Perceptual Systems Through Internal Simulation and Estimation***
Xing Tian and David Poeppel

163 ***Effect of Biomechanical Constraints in the Hand Laterality Judgment Task: Where Does It Come from?***
Gilles Vannuscorps, Agnesa Pillon and Michael Andres

172 ***Understanding Immersivity: Image Generation and Transformation Processes in 3D Immersive Environments***
Maria Kozhevnikov and Rupali P. Dhond

182 ***Abacus in the Brain: A Longitudinal Functional MRI Study of a Skilled Abacus User with a Right Hemispheric Lesion***
Satoshi Tanaka, Keiko Seki, Takashi Hanakawa, Madoka Harada, Sho K. Sugawara, Norihiro Sadato, Katsumi Watanabe and Manabu Honda

frontiers in
PSYCHOLOGY

EDITORIAL
published: 23 April 2013
doi: 10.3389/fpsyg.2013.00198

Mental imagery

Joel Pearson*[1]* and Stephen M. Kosslyn*[2]

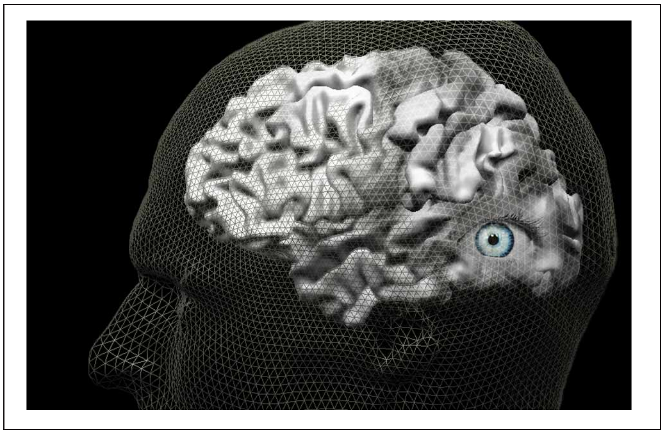
[1] *School of Psychology, The University of New South Wales, Sydney, NSW, Australia*
[2] *Minerva University, San Francisco, CA, USA*
**Correspondence: jpearson@unsw.edu.au*

Edited by:
Philippe G. Schyns, University of Glasgow, UK

Reviewed by:
Philippe G. Schyns, University of Glasgow, UK

Our ability to be conscious of the world around us is often discussed as one of the most amazing yet enigmatic processes under scientific investigation today. However, our ability to imagine the world around us in the absence of stimulation from that world is perhaps even more amazing.

Our capacity to re-experience objects or scenarios that we've encountered before, and to notice new things about those experiences, is itself remarkable. But perhaps more remarkable still is our ability to experience objects or events that do not exist in the world, through our imagination. This is perhaps one of the fundamental abilities that allow us successfully to plan, run dress rehearsals of future events, re-analyze the past—and even simulate or fantasize events that may never happen. In short, it could be argued that this ability is one of the main factors that have allowed us as a species to dominate our planet so profoundly.



Empirical research into mental imagery has seen a recent surge, which is partly a result of new neuroscientific methods and their clever application—but is also due to the discovery and application of additional sorts of objective methods to investigate this inherently internal and private process.

Here we introduce an inspiringly broad range of work that focuses on mental imagery. This ebook contains the work from a broad range of researchers in different fields, both empirical work and reviews. Chapters range from the role of imagery in music, biomechanics, and mathematics to the functions of the cerebral hemispheres in imagery and imagery's effects on sensory perception.

This collection provides a cohesive and broad-spectrum addition to the rapidly growing field of mental imagery. This set of articles provides theoretical insights and an overview of the state of empirical understanding, where it is heading, and how mental imagery relates to other cognitive and sensory functions.

Received: 24 March 2013; accepted: 02 April 2013; published online: 23 April 2013.
Citation: Pearson J and Kosslyn SM (2013) Mental imagery. Front. Psychol. ***4****:198. doi: 10.3389/fpsyg.2013.00198*
This article was submitted to Frontiers in Perception Science, a specialty of Frontiers in Psychology.

frontiers in
PSYCHOLOGY

REVIEW ARTICLE
published: 28 June 2012
doi: 10.3389/fpsyg.2012.00214

Hemispheric differences within the fronto-parietal network dynamics underlying spatial imagery

Alexander T. Sack* and Teresa Schuhmann

Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, Netherlands

Edited by:
Joel Pearson, The University of New South Wales, Australia

Reviewed by:
Raymond Van Ee, University Utrecht, Netherlands
Angelika Lingnau, University of Trento, Italy

****Correspondence:***
Alexander T. Sack, Faculty of Psychology and Neuroscience, Maastricht University, Universiteitssingel 40, 6200 MD Maastricht, Netherlands.
e-mail: a.sack@maastrichtuniversity.nl

Spatial imagery refers to the inspection and evaluation of spatial features (e.g., distance, relative position, configuration) and/or the spatial manipulation (e.g., rotation, shifting, reorienting) of mentally generated visual images. In the past few decades, psychophysical as well as functional brain imaging studies have indicated that any such processing of spatially coded information and/or manipulation based on mental images (i) is subject to similar behavioral demands and limitations as in the case of spatial processing based on real visual images, and (ii) consistently activates several nodes of widely distributed cortical networks in the brain. These nodes include areas within both, the dorsal fronto-parietal as well as ventral occipito-temporal visual processing pathway, representing the "what" versus "where" aspects of spatial imagery. We here describe evidence from functional brain imaging and brain interference studies indicating systematic hemispheric differences within the dorsal fronto-parietal networks during the execution of spatial imagery. Importantly, such hemispheric differences and functional lateralization principles are also found in the effective brain network connectivity within and across these networks, with a direction of information flow from anterior frontal/premotor regions to posterior parietal cortices. In an attempt to integrate these findings of hemispheric lateralization and fronto-to-parietal interactions, we argue that spatial imagery constitutes a multifaceted cognitive construct that can be segregated in several distinct mental sub processes, each associated with activity within specific lateralized fronto-parietal (sub) networks, forming the basis of the here proposed dynamic network model of spatial imagery.

Keywords: spatial imagery, object imagery, brain imaging, imagery and parietal cortex, imagery and premotor cortex, imagery and frontal cortex, spatial attention, spatial working memory

SPATIAL IMAGERY – A MULTIFACETED COGNITIVE-PSYCHOLOGICAL CONSTRUCT

Humans are capable of performing a variety of higher order cognitive abilities such as problem solving, reasoning, contemplating, but also language comprehension, object recognition, spatial orientation, or the vivid re-experience of previously perceived or processed information stored in memory. All of these cognitive functions require, and are to a large extent based on, our ability to generate, inspect, and manipulate inner mental representations of objects, events, and scenes that are not physically present. This ability of mental imagery thus describes a multi-facetted set of cognitive processes that are at the heart of most forms of abstract reasoning or contemplating (Kosslyn et al., 1995; Cohen et al., 1997; Kanwisher and Wojciulik, 2000; Riesenhuber and Poggio, 2000).

While mental imagery by itself is a multifaceted psychological construct that shows conceptual and neurobiological overlap with related cognitive processes such as attention and memory, it is useful to also subdivide mental imagery according to the sensory modality based on which the mental representation is generated. In this sense, the processes that are involved in generating, inspecting, and manipulating visual images in the absence of visual input are referred to as visual mental imagery (Finke, 1989). Objects in visual imagery can be manipulated much like actual objects. Hence, a mentally generated inner image can easily also be mentally transformed, distorted, or rotated in our mind. This can help to reason about the consequences of a potential corresponding physical manipulation (Kosslyn et al., 1998). (Visuo)Spatial imagery particularly refers to the inspection and evaluation of spatial features (e.g., distance, relative position, configuration) and/or the spatial manipulation (e.g., rotation, shifting, reorienting) of mentally generated visual images. When we speak about spatial imagery in the remainder of this article, we thus refer to the mental representation of visual objects, events, or scenes which are either mainly defined by spatial characteristics (e.g., the visual imagination of a spatial configuration) and/or which require in addition to the mere generation of the mental representation, a spatial analysis or manipulation to be mentally performed upon this mental visual image.

Spatial imagery, just like all forms of imagery, is by definition a subjective, private experience that cannot be measured directly, but has to be empirically inferred by indirect measures. These measures vary from subjective self reports on the vividness or size of the mental image, to more objective tasks such as mentally rotating a visually presented object to assess whether it matches, or is mirrored to, a second visual object (Shepard and Metzler, 1971).

It is believed that the completion of such mental rotations rely, at least partly, on spatial mental imagery. In line with this rational, some studies have shown that mental rotation tasks are indeed performed by mentally rotating an object as if it were moving through the intermediate positions along a trajectory, as would occur if the object was physically rotated (Kosslyn et al., 1998; Carpenter et al., 1999; Richter et al., 2000). Since accuracy and response latency of these mental spatial rotations can be objectively measured and compared with other experimental conditions of, e.g., real manual rotation (Sack et al., 2007), such psychophysical experiments offer a means for assessing spatial imagery performance in a behaviorally more controlled manner.

Although it appeared that the question of which exact brain areas are activated during spatial imagery largely depends on the specific features of the imagery task being investigated, e.g., which spatial operation has to be performed based on which mental object, the emerging picture of brain imaging studies is that our capability to mentally visualize, inspect, and manipulate objects is subserved by distributed cortical networks that include regions that are similarly activated when performing comparable perceptual operations (Thompson et al., 2009; Cichy et al., 2011; but see also Lee et al., 2011, nicely showing that although imagery and perception have similar neural substrates, they may involve different network dynamics; Seurinck et al., 2011). Another important and converging finding of these previous imaging studies is that both conceptually and in terms of underlying neural mechanisms it seems important to distinguish cortical regions and neural mechanisms involved in tasks that require participants to mentally represent specific object categories (Ishai et al., 2000) or specific features of objects (e.g., color, size, shape), from those cortical regions and neural mechanisms involved in tasks that explicitly require processing of spatially coded information or spatial manipulation (Trojano et al., 2000). This distinction of object versus spatial imagery can be regarded as analogs to the dichotomy between ventral (what) versus dorsal (where) information processing during visual perception (Mishkin and Ungerleider, 1982; Mishkin et al., 1983;Haxby et al., 1991, 1994).

THE NEUROBIOLOGICAL SEGREGATION OF *WHAT* AND *WHERE* DURING SPATIAL IMAGERY

The neurobiological mechanisms underlying spatial imagery are characterized by widely distributed cortical networks with a multitude of nodes and interaction-patterns in the brain. Numerous neuropsychological (Levine et al., 1985; Farah et al., 1988) and neuroimaging studies (Cohen et al., 1996;Mellet et al., 1996, 1998; D'Esposito et al., 1997; Richter et al., 1997; Knauff et al., 2000; Trojano et al., 2000) have aimed at unraveling the neural foundations of mental imagery using a wide variety of imagery tasks (for a review see Kosslyn et al., 2001). These imaging studies have consistently revealed that the pure imagination and mental representation of a specific mental object results in neural activity within category-specific occipital-temporal regions of the ventral visual processing pathway (Ishai et al., 2000, 2002; O'Craven and Kanwisher, 2000), including superior occipital areas (Mellet et al., 1995, 1996; D'Esposito et al., 1997; de Borst et al., 2011), inferior temporal regions (Carpenter et al., 1999; Mechelli et al., 2004; de Borst et al., 2012), parahippocampal cortex (de Borst et al., 2012), and in some tasks early visual cortex (EVC; Stokes et al., 2011) and/or even primary visual cortex (Kosslyn et al., 1999; Slotnick et al., 2005; de Borst et al., 2012). Likewise, brain regions within the dorsal visual processing pathway are recruited during the spatial processing or manipulation of these mental representations (Kawashima et al., 1995;Mellet et al., 1995, 1996; Cohen et al., 1996; Tagaris et al., 1997; Kosslyn et al., 1998;Sack et al., 2002, 2005, 2008). These cortical regions within the dorsal pathway that in this sense are maybe more strictly related to the spatial aspect of spatial imagery are the bilateral inferior and superior parietal lobule (SPL; Richter et al., 1997; Knauff et al., 2000;Trojano et al., 2000, 2002;Sack et al., 2002, 2005, 2008), bilateral intraparietal sulcus (IPS); precuneus; (Mellet et al., 1996; Trojano et al., 2000;Sack et al., 2002, 2005, 2008), middle forntal gyrus (MFG), supplementary motor area (SMA), frontal eye fields (FEF), and premotor cortex (PMC; Kawashima et al., 1995;Mellet et al., 1995, 1996; Cohen et al., 1996; Tagaris et al., 1997; Kosslyn et al., 1998; Richter et al., 2000; Trojano et al., 2000; Lamm et al., 2001;Sack et al., 2002, 2005, 2008; Sack, 2009; de Borst et al., 2012). Regarding this spatial aspect of spatial imagery, Thompson et al. (2009) suggested differentiating between visualizing spatial locations versus mentally transforming locations, both relying on distinct neural sub networks within the dorsal pathway. Concretely, whereas the visualization of spatial locations recruited mainly areas within occipito-parietal sulcus, medial posterior cingulate, and precuneus, mental spatial transformations were correlated more with activation in superior portions of the parietal lobe and in the postcentral gyrus. Still, since any spatial inspection or manipulation during spatial imagery requires some sort of mental (object) representation upon which the spatial operation can be based and performed on, the core neural network of spatial imagery typically includes brain areas of both the dorsal fronto-parietal as well as ventral occipito-temporal visual processing pathway. **Figure 1** depicts this core network of spatial imagery in the brain, segregated and color-coded in order to distinguish the spatial dorsal (red-colored) from the content ventral (rose-colored) network in the brain activated during spatial imagery.

HEMISPHERIC LATERALIZATION DURING SPATIAL IMAGERY

As can be seen in **Figure 1**, most functional imaging studies show bilateral fronto-parietal networks to be activated during the execution of spatial imagery. The fact that both, left and right posterior parietal cortex (PPC; mostly SPL and IPS) is recruited during spatial imagery, is, at first glance, in contrast to most neuropsychological studies on patients with focal brain lesions which generally propose a dominant role of the left hemisphere in visual imagery (Farah et al., 1985; D'Esposito et al., 1997). In a critical clinical review on visual mental imagery, Trojano and Grossi (1994) presented a number of single cases as well as group studies which demonstrate a dominant role of left posterior parietal areas for mental imagery. Nonetheless, the authors also reported evidence of the role of the right hemisphere in visuospatial imagery as well as in perceptual visuospatial processing. Right brain damaged patients with neglect also show neglect symptoms in imagery tasks, and non-neglect right hemisphere patients show visuospatial deficits in perceptual visuospatial processing and during imagery tasks. While the left hemisphere seems to have a specific role for

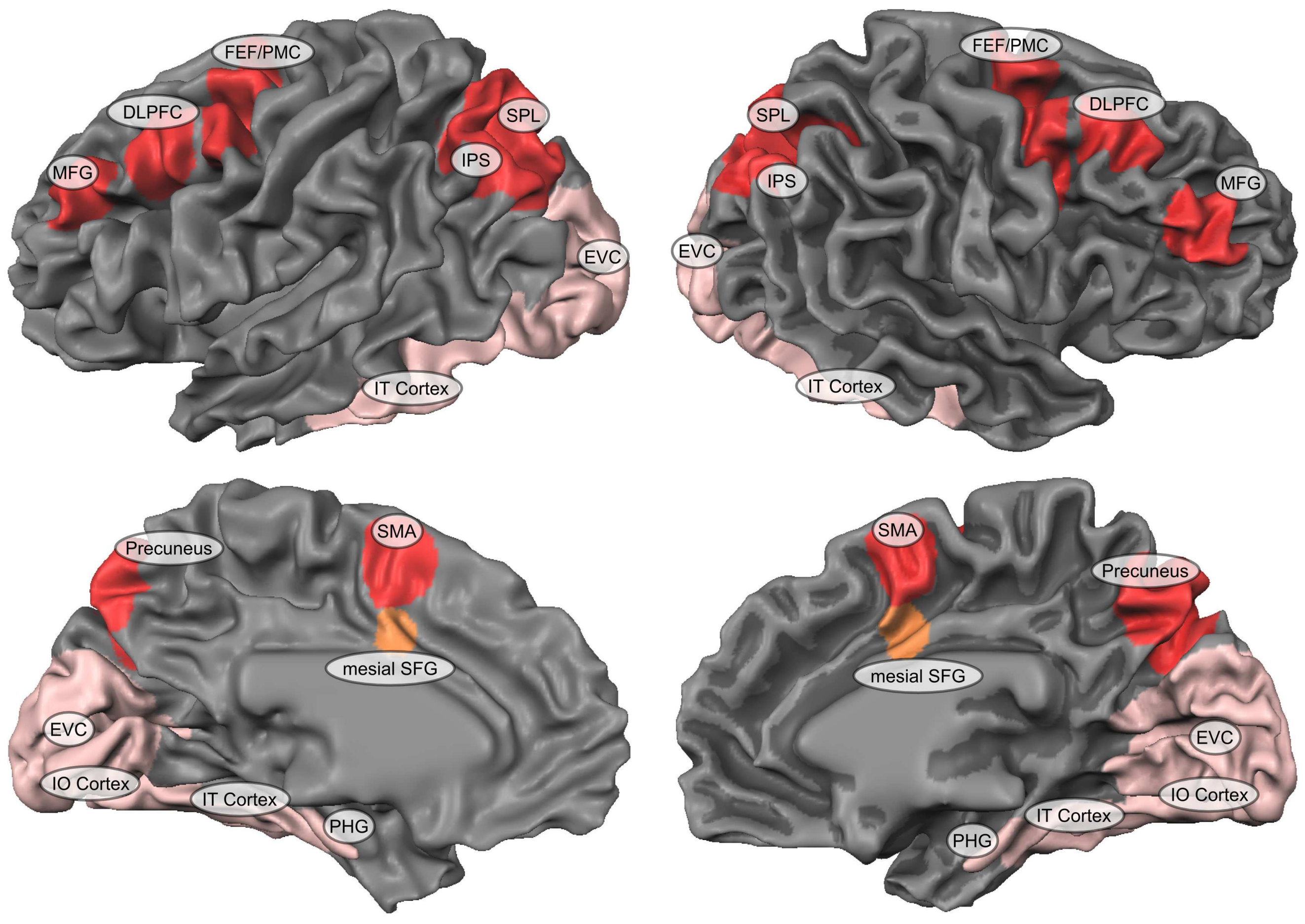


FIGURE 1 | The core neural network of spatial imagery. This figure depicts the neural network of spatial imagery including brain areas of both the dorsal fronto-parietal (red-colored) as well as ventral occipito-temporal (rose-colored) visual processing pathway. It summarizes in one figure the different regions identified in various imagery studies as described in the Section "The neurobiological segregation of *what* and *where* during spatial imagery" of the current manuscript. Most prominent regions within the dorsal fronto-parietal network include bilateral SPL, superior parietal lobe; IPS, Intraparietal sulcus; MFG, middle forntal gyrus; DLPFC, dorsolateral prefrontal cortex; FEF, frontal eye fields; PMC, premotor cortex; Precuneus; and SMA, supplementary motor area. Most prominent regions within the ventral occipito-temporal network include bilateral EVC, early visual cortex; IT, inferior temporal cortex; IO, inferior occipital cortex; PHG, parahippocampal gyrus. The mesial superior frontal gyrus (mSFG) plays a special integrative role in the context of spatial imagery and is therefore color-coded separately.

mental imagery, the right hemisphere seems to be of a more general relevance for visuospatial functions (Trojano and Grossi, 1994).

Our group has contributed to the question of hemispheric lateralization within bilateral parietal cortex (PC) during spatial imagery by using conventional functional magnetic resonance imaging (Trojano et al., 2000), fMRI mental chronometry (Formisano et al., 2002), repetitive (Sack et al., 2002), and time-resolved (Sack et al., 2005) transcranial magnetic stimulation (TMS) experiments. By using a spatial imagery task that involves the generation as well as spatial comparison of mental images, we demonstrated, using event-related fMRI, that the bilateral parietal activity associated with this task includes a temporal activation sequence from left to right PC. When relating and modeling different features of the fMRI responses to the behavioral measures, we found that the *duration* of activation of the early left parietal activation and the *onset* of the (late) right parietal activation correlate with reaction time during spatial imagery performance. These results support the involvement of both parietal lobes in mental imagery, but suggest that each parietal lobe might have a distinct functional role at different moments in time. The sequential activation from left to right suggests that the early left and late right parietal activation during spatial imagery support different components of the cognitive process, for example the generation and subsequent analysis of the visual image. We therefore concluded that within the bilateral PPC activity during spatial imagery, the left PPC underlies the generation of mental images, while the right PPC subserves the spatial processing upon these images.

Such modular models of spatial imagery that propose a division of labor between hemispheres in which the generation of mental representation from memory rely primarily on structures in the posterior left hemisphere, while spatial operations upon these mental representation rely primarily on structures in the posterior right hemisphere, also provide a solution to the aforementioned apparent discrepancy between lesion and imaging studies with regard to the hemispheric lateralization of visual imagery, and are in this sense in agreement with both neuropsychological lesion as well as brain imaging findings of spatial imagery (for review see Sack, 2009). However, from fMRI measurements alone one cannot assess the exact functional necessity or behavioral contribution of a given brain area for a specific mental sub process, such as mental image generation versus spatial analysis. We therefore used non-invasive functional brain stimulation to focally and transiently disrupt neural processing in either left or right PC during spatial imagery, and assessed the respective behavioral effect of this unilateral functional lesion within PPC on spatial imagery performance (Sack et al., 2002). This brain interference study revealed a hemisphere-specific effect of parietal stimulation with only the right parietal disruption leading to spatial imagery impairments. These results contribute new constraints to the modular model of bilateral activation in spatial imagery and are at first glance not in accordance with the aforementioned hemispheric lateralization and division of labor between hemispheres during spatial imagery. Indeed, if left PPC underlies the generation of mental representations and right PPC reflects the spatial operations upon these mental representations, a suppression of either of these brain regions should result in impaired spatial imagery performance. Fortunately, based on a combined fMRI and transcranial magnetic stimulation study (Sack et al., 2002), we were able to further fractionate specialized processing components in the right PPC and revealed the existence of highly dynamic compensatory mechanisms between the left and right hemisphere during the execution of spatial imagery (Sack et al., 2005). This study suggested that although the left PPC is predominantly specialized in mental image generation and the right PPC in spatial comparisons of imagined content, the right hemisphere is also able to immediately compensate for (virtual) lesions of the left hemisphere by taking over its specific function, but not vice versa. Hence, in case of left parietal functional lesion, the right PC will now subserve both functions, mental image generation, and spatial analysis of the mental image. Discrepancies across studies concerning the hemispheric lateralization during mental imagery likely arise because different aspects of imagery are carried out by different parts of a bi-hemispheric neural network. The fact that an isolated deficit of the ability to generate inner visual images following unilateral lesion is clinically hardly reported could also be explained on the basis of the compensatory processes revealed in our study (Sack, 2009, 2010). Interestingly, such hemispheric asymmetries between left and right PPC apply to both, the direct functional relevance (only right parietal disruption leads to behavioral impairments) as well as ability of inter-hemispheric compensation (right PC can compensate for left PC, but not nice versa).

DYNAMIC ANTERIOR-TO-POSTERIOR BRAIN NETWORK CONNECTIVITY DURING SPATIAL IMAGERY

The execution of various spatial imagery paradigms consistently activates core areas of the dorsal fronto-parietal visual pathway, including bilateral parietal, prefrontal, and premotor areas (Kawashima et al., 1995;Mellet et al., 1995, 1996; Cohen et al., 1996; Richter et al., 1997; Tagaris et al., 1997; Kosslyn et al., 1998; Carpenter et al., 1999; Knauff et al., 2000;Trojano et al., 2000, 2002; Lamm et al., 2001;Sack et al., 2002, 2005, 2008; de Borst et al., 2012). With regard to the spatial processing component of spatial imagery, a strong focus has been put on the bilateral PPC activation due to the prominent role of PPC within the dorsal spatial processing stream. Carpenter et al. (1999) correlated the increment of reaction time during mental rotation of cubes with changes in regional cerebral activation. With higher angular disparity, activation increased in the parietal lobes bilaterally, but not in the temporal lobe. This discrepancy indicates that the PC plays a central role in the visuospatial transformations of mental rotation (Goebel et al., 1998; Formisano et al., 2002; Sack et al., 2002) whereas the ventral (temporal) pathway, which is essential for identifying a figure, does not specifically support this operation. This again seems to strengthen and justify the prominent role of PPC in spatial imagery research. However, importantly, this study also revealed that the activity in the motor areas of the frontal lobe was significantly higher during the mental rotation paradigm as compared to a motor control condition, suggesting that "the so-called motor areas are not simply involved in motor planning and execution" (Carpenter et al., 1999), but play a crucial role in the computation of imagined motion of objects as well. In a similar vein, Richter et al. (2000) used fMRI to investigate the participation of the neocortical motor areas in the Shepard and Metzler's (1971) mental rotation task. Seven regions of interest (ROIs) were analyzed separately: Left and right SPL, SMA, and left and right premotor areas. The results showed that the observed activation in premotor areas was likely related to the very execution of the mental rotation task (Richter et al., 2000).

These studies thus indicate the potential functional contribution of prefrontal and premotor brain areas during spatial imagery. The question remained, however, whether these prefrontal and premotor activities during mental rotation tasks are more related to the potential involvement of visual working memory rather than being critical neural structures for the visual imagery process *per se*. Or in other words, what would happen to the here identified premotor and prefrontal activations in case of pure visual imagery, i.e., when generating mental representations of objects that have never been perceived before? In such cases, the generation of visual images does not result from the reactivation of previously stored memories but does result from an online construction of images based on the processing of, e.g., verbal instructions and their encoding in a visual format. Mellet et al. (1996, 1998, 2000) used PET to monitor regional cerebral blood flow variations while participants were constructing mental images of objects made of three-dimensional cube assemblies from acoustically presented instructions. Compared to a control condition, the mental construction task specifically activated a bilateral occipito-parietal-frontal network, including the superior occipital cortex, the inferior PC, and again also the PMC. These studies thus suggest that in addition to the well-established functional role of posterior parietal cortices during spatial imagery, also the prefrontal and premotor activations revealed during imagery task are of direct functional relevance for the imagery performance and likely also sub serve specific cognitive sub functions within the multifaceted cognitive-psychological construct of imagery. However, as described above, while some of the previous functional imaging studies on spatial imagery tried to subscribe different cognitive (sub) functions to, e.g., left versus right PPC (Formisano et al., 2002;Sack et al., 2002, 2005), a systematic investigation of the specific functional contribution and/or the exact spatio-temporal interactions with the always co-activated prefrontal and premotor brain regions during spatial imagery is missing. All studies discussed so far either focused exclusively on the PPC and neglected all additionally activated brain regions (Trojano et al., 2000; Formisano et al., 2002;Sack et al., 2002, 2005) or simply descriptively reported the co-activation of anterior premotor and prefrontal brain regions within large fronto-parietal networks (Mellet et al., 1998, 2000; Carpenter et al., 1999; Richter et al., 2000), without analyzing the network dynamics between these anterior premotor/prefrontal and posterior parietal activations during spatial imagery. These shortcomings were mainly due to respective boundaries in spatial and temporal resolution of the functional imaging techniques being used, and due to the unavailability of more advanced analyses tools for functionally segregating the acquired brain imaging data into separate networks of effective brain connectivity.

In our recent work, our group aimed to address this shortcoming (Sack et al., 2008). Participants were asked to mentally construct an inner image of either an entirely new object or a new configuration of objects, simply based on either visually or verbally presented instructions. Importantly, participants were required to construct and spatially rotate these abstract mental images generated from sequentially presented instructions, meaning that only pieces of the final mental object were provided in successive steps, sequentially building up the final mental object that had never been perceptually encountered as a whole. This behaviorally controlled spatial imagery paradigm was investigated using time-resolved event-related functional magnetic resonance imaging and analyzed based on data-driven and multivariate fMRI analysis tools. By accounting for the full spatial pattern of brain activity measured simultaneously at many locations, we functionally segregated an early from a late premotor-parietal-occipito-temporal-cortex (OTC) and late premotor-prefrontal activation network. We revealed, using effective brain connectivity analyses, that the information coming from sensory brain regions was first sent to bilateral PMC and then to bilateral medial dorsal PC. The early left PC received additional input from bilateral occipito-parietal regions. One might thus speculate that this early bilateral premotor-(medial dorsal) parietal activation network underlies the online processing of the sequentially presented modality-independent spatial instructions (Corbetta et al., 1993; Jonides et al., 1993; Haxby et al., 1994; Courtney et al., 1996; Prabhakaran et al., 2000). In contrast, the identified late premotor-parietal network showed a clear hemispheric difference with only the late left PMC projecting back to bilateral parietal regions while at the same time sending neural signals to bilateral OTC. This specific effective connectivity network might thus represent the juxtaposing of the sequentially presented stimuli and thus the successive construction of the slowly emerging final mental object representation during imagery (Roland and Gulyas, 1995;Mellet et al., 1996, 1998; D'Esposito et al., 1997). During this late premotor-parietal activity network, neural input was also sent to bilateral prefrontal cortex (PFC). This late bilateral premotor-prefrontal activation network might sub serve the necessary maintenance of the spatially rotated visual mental object in spatial short-term working memory (Tulving et al., 1994; Buckner et al., 1995; Moscovitch et al., 1995; Roland and Gulyas, 1995; Smith et al., 1995; Courtney et al., 1996; Cohen et al., 1997). These findings suggest that the activation flow underlying the construction and spatial transformation of visual mental images first recruits premotor regions which then project to, or receive information on demand from, parietal regions in a top-down manner, putting prefrontal, and PMC into a new and central focus also during higher cognitive functions (Sack et al., 2008). Such functionally coupled activations of the parietal and premotor cortices have also been described for other cognitive functions (Abe and Hanakawa, 2009) and visuospatial tasks, such as spatial localization (Haxby et al., 1994) or shifting of spatial attention (Corbetta et al., 1993), and in situations explicitly involving the spatial working memory (Jonides et al., 1993; Courtney et al., 1996). The exchange of information between the premotor regions and the dorsal route thus appears to be a general feature during spatial processing, whatever the nature of the initial input. It is thus likely that the parietal "perceptual" pole and the frontal "motor" pole systematically exchange spatial information, whether a motor action is envisioned or not, thus executing the encoding of a spatial environment in its descriptive and behavioral aspects.

INTEGRATION BETWEEN DORSAL AND VENTRAL PATHWAYS DURING SPATIAL IMAGERY

In the previous sections, we have described and shown that within the dorsal fronto-parietal network activated during spatial imagery, a direction of information flow seems to exist from anterior frontal/premotor to posterior parietal cortices. However, in addition to the dorsal fronto-parietal network, all imagery paradigms also activate several nodes of the ventral occipito-temporal visual processing pathway (Ishai et al., 2000, 2002; O'Craven and Kanwisher, 2000), including superior occipital areas (Mellet et al., 1995, 1996; D'Esposito et al., 1997; de Borst et al., 2012), inferior temporal regions (Carpenter et al., 1999; Mechelli et al., 2004; de Borst et al., 2012), parahippocampal cortex (de Borst et al., 2012), and in some tasks even primary visual cortex (Kosslyn et al., 1999; de Borst et al., 2012). This seems to make perfect sense because any spatial operation or spatial processing in spatial imagery requires some sort of mental (object) representation upon which the spatial inspection or manipulation can be based and performed on. Subsequently, the core neural network of spatial imagery includes brain areas of both the dorsal fronto-parietal as well as ventral occipito-temporal processing pathway, representing the "what" and "where" aspects of spatial imagery (see **Figure 1**).

An open question, however, then is how the necessary dynamic interaction between the "what" and "where" aspects of spatial imagery is integrated in the brain. de Borst et al. (2011) directly addressed this question by investigating the functional role of the frontal regions and their interaction with the "what," "where" and early visual regions during complex visuospatial scene imagery. Participants were required to perform detailed scene imagery that captured both object and spatial mental imagery aspects, while measuring all imagery-relevant network nodes in the brain and their relative temporal onset of activation using event-related fMRI. In addition, electroencephalography (EEG) was recorded to validate the fMRI latency results and to derive more information on the underlying functional roles from the involved frequency bands. This study could nicely demonstrate that the "what" and "where" aspects of spatial imagery are integrated into one visually imagined scene by the frontal regions, including PMC, right MFG, and mesial SFG (mSFG). The mSFG seemed to be most crucial for this integration process because it was activated earliest and predicted later imagery performance. The early and behaviorally relevant involvement of the frontal regions suggest that these frontal regions indeed "orchestrate" the ventral occipital-temporal "what" and dorsal parietal "where" regions during spatial imagery. Hence, the frontal regions and in particular the mSFG seem to integrate those areas encoding the detailed mental representation with those areas encoding the spatial layout, in order to form and maintain the subjective experience of one coherent mental picture. In line with this interpretation, several working memory studies have also suggested that mSFG plays a role in the integration of visual and spatial features during visual short-term memory (VSTM; Mitchell et al., 2000; Prabhakaran et al., 2000; Munk et al., 2002). In accordance with the already described anterior-to-posterior information flow within the dorsal fronto-parietal network during spatial imagery (Sack et al., 2008) this study also showed a rather late involvement of bilateral PPC, suggesting again that these PPC activations seem to be drawn on by the frontal regions during the later stages of image construction, rather than being a dorsal starting point of spatial imagery. In line with this integrative role of the anterior premotor and prefrontal regions in spatial imagery, Abe and Hanakawa (2009) also suggested that the functional interplay through the prefrontal-premotor connections may mediate the integration of specific sub-operations for multi-step cognitive manipulation.

DYNAMIC NETWORK MODEL OF SPATIAL IMAGERY

In an attempt to summarize, converge, and integrate the brain imaging findings of our own and other groups on spatial imagery over the past 10 years, **Figure 2** depicts the here proposed new dynamic network model of spatial imagery.

Spatial imagery consistently activates several nodes within widely distributed cortical networks in the brain. Importantly, these nodes include areas within both, the dorsal fronto-parietal as well as ventral occipito-temporal visual processing pathway (all silver-shaded areas in **Figure 2**). Most prominent regions within the dorsal pathway during spatial imagery include bilateral PC, PMC, and PFC. Likewise, most prominent regions of the ventral pathway activated during spatial imagery are located along the OTC, and include inferior temporal regions and parahippocampal cortex, but also superior occipital areas and in some conditions even primary visual cortex (see **Figure 2**, silver-shaded areas).

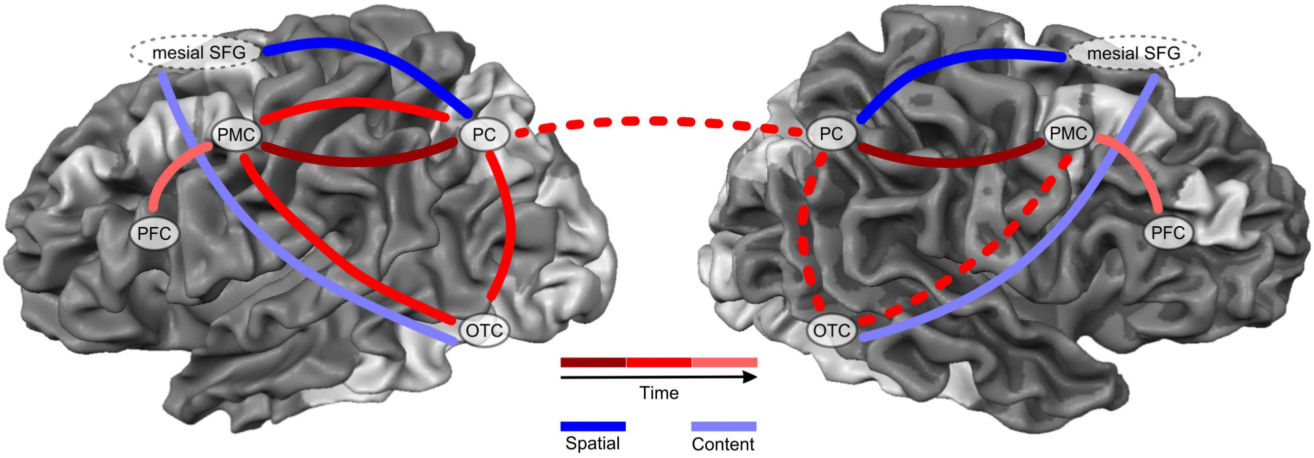


FIGURE 2 | Dynamic network model of spatial imagery. This figure depicts the here proposed dynamic network model of spatial imagery. Spatial imagery consistently activates several nodes within both, the dorsal fronto-parietal as well as ventral occipito-temporal visual processing pathway (silver-shaded areas). Most prominent regions within the dorsal pathway during spatial imagery include bilateral PC, Parietal Cortex; PMC, Premotor Cortex; and PFC, Prefrontal Cortex. Likewise, most prominent regions of the ventral pathway activated during spatial imagery are located along the Occipito-Temporal-Cortex (OTC), and include inferior temporal regions and parahippocampal cortex, but also superior occipital areas and in some conditions even primary visual cortex (silver-shaded areas). During spatial imagery, these two pathways can be labeled as representing the CONTENT (curve color-coded in light blue) versus SPATIAL (curve color-coded in dark blue) aspects of spatial imagery. The areas within the fronto-parietal dorsal network dynamically exchange information during spatial imagery with a direction of information flow from anterior frontal/premotor regions to posterior parietal cortices. An early bilateral PMC-PC (solid curve color-coded in dark red) can be segregated from a later left-lateralized PMC-PC-OTC activation network (solid curve color-coded in red). Spatial imagery thus first recruits bilateral anterior premotor cortices, which then send neural information to, or receive on demand neural information from, bilateral parietal cortices. This model thus proposes that the well-established PPC activation during spatial imagery seems to be drawn on by the frontal regions at later stages in the course of imagery, rather than being a dorsal starting point of spatial imagery, as previously suggested. Moreover, in order to form one coherent mental visuospatial picture, all these content and spatial aspects and segregated processing stages of spatial imagery need to be integrated at a brain system level. This integration process is done by the mesial SFG. Mesial SFG orchestrates remote ventral occipital-temporal "what" regions and dorsal parietal "where" regions in order to integrate areas encoding the detailed mental visual representation with those areas encoding the spatial layout or manipulation. Finally, the late neural dynamic information flow between bilateral PMC and PFC (curve color-coded in light red) represents the necessary maintenance of the now spatially processed or manipulated and thus integrated mental object in (spatial) visual short-term working memory.

The parallel processing within dorsal and ventral networks makes perfect sense in case of spatial imagery. It shows that indeed the imagination and mental representation of any specific mental object or scene always results in neural activity within category-specific occipital-temporal regions of the ventral visual processing pathway. This is true for all forms and aspects of visual imagery, whether or not it includes a spatial processing component. In case of spatial imagery, however, additional brain regions within the dorsal visual processing pathway are recruited because of the here required processing of spatially coded information and/or spatial manipulation of these mental representations. In this sense, the two pathways can be labeled as representing the CONTENT (**Figure 2**; curve color-coded in light blue) versus SPATIAL (**Figure 2**; curve color-coded in dark blue) aspects of spatial imagery (**Figure 2**).

Importantly, the areas within the fronto-parietal dorsal network dynamically exchange information during spatial imagery with a direction of information flow from anterior frontal/premotor regions to posterior parietal cortices. A spatio-temporal segregation has been suggested, dissociating an early bilateral PMC-PC (**Figure 2**; curve color-coded in dark red) from a later left-lateralized PMC-PC-OTC activation network (**Figure 2**; curve color-coded in red). Importantly, spatial imagery thus first recruits bilateral anterior premotor cortices, which then send neural information to, or receive on demand neural information from, bilateral parietal cortices. This model thus proposes that the well-established PPC activation during spatial imagery seems to be drawn on by the frontal regions at later stages in the course of imagery, rather than being a dorsal starting point of spatial imagery, as previously suggested.

So far, the model describes the parallel processing within the ventral CONTENT and dorsal SPATIAL network of spatial imagery, as well as the dynamic anterior-to-posterior information flow within the dorsal fronto-parietal activation network. However, in order to form one coherent mental visuospatial picture, all these aspects and segregated processing stages need to be integrated at a brain system level. We could recently suggest that this integration process is likely done by the co-activated frontal regions, in particular the mSFG. mSFG seems to literally orchestrate remote ventral occipital-temporal "what" regions and dorsal parietal "where" regions in order to integrate areas encoding the detailed mental visual representation with those areas encoding the spatial layout or manipulation. In line with this idea, neural dynamic information flow during spatial imagery was revealed also late between bilateral PMC and PFC, likely representing the necessary maintenance of the spatially processed or manipulated and thus integrated mental object in (spatial) visual short-term working memory (**Figure 2**; curve color-coded in light red).

QUO VADIS SPATIAL IMAGERY?

Besides the specific questions on the neurobiology, lateralization, and/or spatio-temporal dynamics within and across the described fronto-parietal and occipito-temporal network activity during spatial imagery, a striking conceptual uncertainty occurs on how to differentiate the concept of spatial imagery to related phenomena such as spatial attention or spatial visual working memory. While these related processes are often treated as separate cognitive-psychological constructs, it is undisputed that the ability to maintain visual information online in working memory largely depends on mental imagery, and likewise, performing spatial manipulations upon mentally generated visual images shares large conceptual overlap with the concept of (covert) visuospatial attention and memory. In an attempt to combine aspects of the so-called analog (James, 1890; Paivio, 1971; Kosslyn, 1980; Kosslyn and Ochsner, 1994) versus propositional (Anderson and Bower, 1973; Pylyshyn, 1973) theory of mental imagery, Kosslyn (1980) proposed a computational hybrid imagery model consisting of three basic components: a visual buffer, long-term stored representations, and image-processing operations. Visual images are generated by retrieving information from long-term visual memory (LTM) and constructing them in a spatial format in the visual buffer. Although the LTM contains both analog and propositional components, the final visual image is an analog representation. Images in the buffer can be mentally manipulated and transformed (e.g., by rotation or scanning) and then be inspected for new information (Palmer, 1999); a process likely largely mediated by spatial attention mechanisms. In line with this depictive view of visual mental imagery, Slotnick et al. (2005) demonstrated that visual mental imagery can evoke cortical activity with precise visual field topography, i.e., imagery-induced retinotopic maps that are similar to the perception maps.

As evident from the attempts to segregate and subscribe different cognitive sub functions to the revealed sub networks during spatial imagery using functional imaging, these psychological labels or cognitive-psychological constructs of the revealed brain activation pattern often have to fall back on other cognitive-psychological constructs such as attention and memory. In fact, attention and memory processes are consistently used to explain and describe the various mental sub processes and their underlying neural activations during the execution of (spatial) imagery. This also holds true the other way around: attentional modulation occurs in the absence of any phenomenal experience, e.g., in form of expectations. Kastner et al. (1999) described attention-modulated activity in fronto-parietal areas and in primary visual cortex during a visuospatial attention task without visual stimulation. Regions in parietal and frontal cortex responded when observers covertly pay attention to a peripheral location in expectation of the stimulus occurrence. Therefore, attention is not necessarily associated with visual perception. However, although several neuroimaging studies have shown that attention can modulate extrastriate cortical regions within the dorsal and ventral processing streams (Kanwisher and Wojciulik, 2000; Martinez et al., 2001; Yantis et al., 2002), the role of the primary visual cortex remains controversial (Kastner et al., 1999; Posner and Gilbert, 1999; Sengpiel and Hubener, 1999). But what psychological and neural mechanisms are responsible for the maintenance of spatial attention in the absence of visual stimuli? And what are the conceptual differences to spatial imagery and spatial working memory processes? Mental imagery is defined as a perceptual experience in the absence of an appropriate physical stimulus (Finke, 1989). Conceptually, attention and imagery can both be characterized by specific top-down processes that modulate extrastriate cortex activity in the absence of visual stimulation. On a neuronal level both, spatial imagery and spatial attention show immense functional overlapping: mental imagery is generally, like attention, accompanied by an activation of a fronto-parietal network (Trojano et al., 2000; Formisano et al., 2002; Sack et al., 2002). Several imagery studies have shown that mental imagery, like attention, can modulate extrastriate cortices (Le Bihan et al., 1993; Roland and Gulyas, 1995; D'Esposito et al., 1997). Finally, some imagery studies have even, like attention, revealed primary visual cortex activity during visual imagery (Kosslyn et al., 1999; de Borst et al., 2011; Slotnick et al., 2005). However, the involvement of occipital areas during imagery is, like in attention, still a matter of debate (Kosslyn and Ochsner, 1994; Roland and Gulyas, 1995; Mellet et al., 2000). The ability to maintain spatial information online in memory depends on spatial attention and might be mediated by spatial imagery. Imagery might in general aid any cognitive functions in the respective sensory domain. It has, for example, been claimed that visual imagery ability correlates with visuospatial memory span (Kail, 1997). Keogh and Pearson (2011) showed that performance in visual working memory – but not iconic visual memory – can be predicted by the strength of mental imagery.

This raises the question whether spatial imagery, spatial attention, and memory in the end represent very similar or even identical neuronal processes that can only be differentiated on a psychological level as servants for a thinking framework among neuroscientists. On a neuronal level, these two or three so far rather independently investigated constructs might prove to recruit similar or even identical neural structures. Hence, what might be a very crucial and important semantic and conceptual difference for psychology and neuroscience might for our brain in the end simply require the recruitment of identical networks of in this sense identical processes.

However, while this might in the end indeed show to be the case, such a conclusion would at this point be premature. One problem of most paradigms investigated so far in the context of visual imagery and/or VSTM is the fact that they are usually conceptually confounded by not representing pure operationalizations of the cognitive-psychological construct under investigation. Most cognitive and neuroscientific studies of mental imagery required participants to learn and memorize visual images, or recall images from past autobiographical experiences. Such "imagery" paradigms are by definition confounded by processes involving visual working or episodic memory.

Hence, it might very well be that despite all the conceptual and apparent neurobiological overlap between imagery, attention, and memory, clear differences and neurobiological segregations can be found and described on the basis of, e.g., cleaner experimental designs directly comparing the different processes, and/or more elaborated multivariate brain imaging analysis tools. Another elegant approach might be to use functional brain interference techniques such as transcranial magnetic stimulation (TMS) to chart the time point at which the identified similar neural structures activated by imagery and memory are functionally relevant for either or both of these processes. Using this approach, Cattaneo et al. (2009), e.g., compared the functional contribution of EVC in short-term memory retention and visual mental imagery at different moments in time using chronometric TMS. They could show that short pulses of functional interference applied to EVC at the start of the retention interval increased reaction times for the memory trials, but not for the imagery trials; while later TMS pulses over EVC affected both processes, memory and imagery. In other words, while both visual imagery and VSTM recruited identical neural structures within EVC, the time point of functional recruitment was significantly different between these two processes. This enabled the authors to segregate between imagery and memory in the temporal rather than spatial neural domain. More studies like these, directly comparing and segregating the neurobiology underlying imagery, attention, and memory within one experimental session are needed to clearly separate and differentiate the different mental sub processes involved in spatial imagery from those involved in attention and memory. In the end, we should be able to explain the individual cognitive, phenomenological, psychological, and behavioral differences within spatial imagery (see, e.g., Borst and Kosslyn, 2010), as well as the differences between imagery, attention, and memory, by linking each of these processes to different underlying neurobiological processes and spatio-temporal network dynamics in the brain.

ACKNOWLEDGMENTS

The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC Grant agreement no. (263472).

REFERENCES

Abe, M., and Hanakawa, T. (2009). Functional coupling underlying motor and cognitive functions of the dorsal premotor cortex. *Behav. Brain Res.* 198, 13–23.

Anderson, J. R., and Bower, G. H. (1973). *Human Associative Memory*. New York: Winston.

Borst, G., and Kosslyn, S. M. (2010). Individual differences in spatial mental imagery. *Q. J. Exp. Psychol. (Hove)* 63, 2031–2050.

Buckner, R. L., Petersen, S. E., Ojemann, J. G., Miezin, F. M., Squire, L. R., and Raichle, M. E. (1995). Functional anatomical studies of explicit and implicit memory retrieval tasks. *J. Neurosci.* 15, 12–29.

Carpenter, P. A., Just, M. A., Keller, T. A., Eddy, W., and Thulborn, K. (1999). Graded functional activation in the visuospatial system with the amount of task demand. *J. Cogn. Neurosci.* 11, 9–24.

Cattaneo, Z., Vecchi, T., Pascual-Leone, A., and Silvanto, J. (2009). Contrasting early visual cortical activation states causally involved in visual imagery and short-term memory. *Eur. J. Neurosci.* 30, 1393–1400.

Cichy, R. M., Heinzle, J., and Haynes, J. D. (2011). Imagery and perception share cortical representations of content and location. *Cereb. Cortex* 22, 372–380.

Cohen, J. D., Perlstein, W. M., Braver, T. S., Nystrom, L. E., Noll, D. C., Jonides, J., and Smith, E. E. (1997). Temporal dynamics of brain activation during a working memory task. *Nature* 386, 604–608.

Cohen, M. S., Kosslyn, S. M., Breiter, H. C., Digirolamo, G. J., Thompson, W. L., Anderson, A. K., Brookheimer, S. Y., Rosen, B. R., and Belliveau, J. W. (1996). Changes in cortical activity during mental rotation. A mapping study using functional MRI. *Brain* 119(Pt 1), 89–100.

Corbetta, M., Miezin, F. M., Shulman, G. L., and Petersen, S. E. (1993). A PET study of visuospatial attention. *J. Neurosci.* 13, 1202–1226.

Courtney, S. M., Ungerleider, L. G., Keil, K., and Haxby, J. V. (1996). Object and spatial visual working memory activate separate neural systems in human cortex. *Cereb. Cortex* 6, 39–49.

de Borst, A. W., Sack, A. T., Jansma, B. M., Esposito, F., De Martino, F., Valente, G., Roebroeck, A., Di Salle, F., Goebel, R., and Formisano, E. (2011). Integration of "what" and "where" in frontal cortex during visual imagery of scenes. *Neuroimage* 60, 47–58.

de Borst, A. W., Sack, A. T., Jansma, B. M., Esposito, F., De Martino, F., Valente, G., Roebroeck, A., Di Salle, F., Goebel, R., and Formisano, E. (2012). Integration of "what" and "where" in frontal cortex during visual imagery of scenes. *Neuroimage* 60, 47–58.

D'Esposito, M., Detre, J. A., Aguirre, G. K., Stallcup, M., Alsop, D. C., Tippet, L. J., and Farah, M. J. (1997). A functional MRI study of mental image generation. *Neuropsychologia* 35, 725–730.

Farah, M. J., Gazzaniga, M. S., Holtzman, J. D., and Kosslyn, S. M. (1985). A left hemisphere basis for visual mental imagery? *Neuropsychologia* 23, 115–118.

Farah, M. J., Hammond, K. M., Levine, D. N., and Calvanio, R. (1988). Visual and spatial mental imagery: dissociable systems of representation. *Cogn. Psychol.* 20, 439–462.

Finke, R. A. (1989). *Principles of Mental Imagery*. Cambridge, MA: MIT Press.

Formisano, E., Linden, D. E., Di Salle, F., Trojano, L., Esposito, F., Sack, A. T., Grossi, D., Zanella, F. E., and Goebel, R. (2002). Tracking the mind's image in the brain I: time-resolved fMRI during visuospatial mental imagery. *Neuron* 35, 185–194.

Goebel, R., Khorram-Sefat, D., Muckli, L., Hacker, H., and Singer, W. (1998). The constructive nature of vision: direct evidence from functional magnetic resonance imaging studies of apparent motion and motion imagery. *Eur. J. Neurosci.* 10, 1563–1573.

Haxby, J. V., Grady, C. L., Horwitz, B., Ungerleider, L. G., Mishkin, M., Carson, R. E., Herscovitch, P., Schapiro, M. B., and Rapoport, S. I. (1991). Dissociation of object and spatial visual processing pathways in human extrastriate cortex. *Proc. Natl. Acad. Sci. U.S.A.* 88, 1621–1625.

Haxby, J. V., Horwitz, B., Ungerleider, L. G., Maisog, J. M., Pietrini, P., and Grady, C. L. (1994). The functional organization of human extrastriate cortex: a PET-rCBF study of selective attention to faces and locations. *J. Neurosci.* 14, 6336–6353.

Ishai, A., Haxby, J. V., and Ungerleider, L. G. (2002). Visual imagery of famous faces: effects of memory and attention revealed by fMRI. *Neuroimage* 17, 1729–1741.

Ishai, A., Ungerleider, L. G., and Haxby, J. V. (2000). Distributed neural systems for the generation of visual images. *Neuron* 28, 979–990.

James, W. (1890). *The Principles of Psychology*, Vol. 1 and 2. New York: Henry Holt and Co, Inc.

Jonides, J., Smith, E. E., Koeppe, R. A., Awh, E., Minoshima, S., and Mintun, M. A. (1993). Spatial working memory in humans as revealed by PET. *Nature* 363, 623–625.

Kail, R. (1997). Processing time, imagery, and spatial memory. *J. Exp. Child Psychol.* 64, 67–78.

Kanwisher, N., and Wojciulik, E. (2000). Visual attention: insights from brain imaging. *Nat. Rev. Neurosci.* 1, 91–100.

Kastner, S., Pinsk, M. A., De Weerd, P., Desimone, R., and Ungerleider, L. G. (1999). Increased activity in human visual cortex during directed attention in the absence of visual stimulation. *Neuron* 22, 751–761.

Kawashima, R., Roland, P. E., and O'Sullivan, B. T. (1995). Functional anatomy of reaching and visuomotor learning: a positron emission tomography study. *Cereb. Cortex* 5, 111–122.

Keogh, R., and Pearson, J. (2011). Mental imagery and visual working memory. *PLoS ONE* 6, e29221. doi:10.1371/journal.pone.0029221

Knauff, M., Kassubek, J., Mulack, T., and Greenlee, M. W. (2000). Cortical activation evoked by visual mental imagery as measured by fMRI. *Neuroreport* 11, 3957–3962.

Kosslyn, S. M. (1980). *Image and Mind*. Cambridge, MA: Harvard University Press.

Kosslyn, S. M., Behrmann, M., and Jeannerod, M. (1995). The cognitive neuroscience of mental imagery. *Neuropsychologia* 33, 1335–1344.

Kosslyn, S. M., Digirolamo, G. J., Thompson, W. L., and Alpert, N. M. (1998). Mental rotation of objects versus hands: neural mechanisms revealed by positron emission tomography. *Psychophysiology* 35, 151–161.

Kosslyn, S. M., Ganis, G., and Thompson, W. L. (2001). Neural foundations of imagery. *Nat. Rev. Neurosci.* 2, 635–642.

Kosslyn, S. M., and Ochsner, K. N. (1994). In search of occipital activation during visual mental imagery. *Trends Neurosci.* 17, 290–292.

Kosslyn, S. M., Pascual-Leone, A., Felician, O., Camposano, S., Keenan, J. P., Thompson, W. L., Ganis, G., Sukel, K. E., and Alpert, N. M. (1999). The role of area 17 in visual imagery: convergent evidence from PET and rTMS. *Science* 284, 167–170.

Lamm, C., Windischberger, C., Leodolter, U., Moser, E., and Bauer, H. (2001). Evidence for premotor cortex activity during dynamic visuospatial imagery from single-trial functional magnetic resonance imaging and event-related slow cortical potentials. *Neuroimage* 14, 268–283.

Le Bihan, D., Turner, R., Zeffiro, T. A., Cuenod, C. A., Jezzard, P., and Bonnerot, V. (1993). Activation of human primary visual cortex during visual recall: a magnetic resonance imaging study. *Proc. Natl. Acad. Sci. U.S.A.* 90, 11802–11805.

Lee, S. H., Kravitz, D. J., and Baker, C. I. (2011). Disentangling visual imagery and perception of real-world objects. *Neuroimage* 59, 4064–4073.

Levine, D. N., Warach, J., and Farah, M. (1985). Two visual systems in mental imagery: dissociation of "what" and "where" in imagery disorders due to bilateral posterior cerebral lesions. *Neurology* 35, 1010–1018.

Martinez, A., Di Russo, F., Anllo-Vento, L., Sereno, M. I., Buxton, R. B., and Hillyard, S. A. (2001). Putting spatial attention on the map: timing and localization of stimulus selection processes in striate and extrastriate visual areas. *Vision Res.* 41, 1437–1457.

Mechelli, A., Price, C. J., Friston, K. J., and Ishai, A. (2004). Where bottom-up meets top-down: neuronal interactions during perception and imagery. *Cereb. Cortex* 14, 1256–1265.

Mellet, E., Briscogne, S., Tzourio-Mazoyer, N., Ghaem, O., Petit, L., Zago, L., Etard, O., Berthoz, A., Mazoyer, B., and Denis, M. (2000). Neural correlates of topographic mental exploration: the impact of route versus survey perspective learning. *Neuroimage* 12, 588–600.

Mellet, E., Petit, L., Mazoyer, B., Denis, M., and Tzourio, N. (1998). Reopening the mental imagery debate: lessons from functional anatomy. *Neuroimage* 8, 129–139.

Mellet, E., Tzourio, N., Crivello, F., Joliot, M., Denis, M., and Mazoyer, B. (1996). Functional anatomy of spatial mental imagery generated from verbal instructions. *J. Neurosci.* 16, 6504–6512.

Mellet, E., Tzourio, N., Denis, M., and Mazoyer, B. (1995). A positron emission tomography study of visual, and mental, spatial exploration. *J. Cogn. Neurosci.* 7, 433–445.

Mishkin, M., and Ungerleider, L. G. (1982). Contribution of striate inputs to the visuospatial functions of parieto-preoccipital cortex in monkeys. *Behav. Brain Res.* 6, 57–77.

Mishkin, M., Ungerleider, L. G., and Macko, K. A. (1983). Object vision and spatial vision: two cortical pathways. *Trends Neurosci.* 6, 414–417.

Mitchell, K. J., Johnson, M. K., Raye, C. L., Mather, M., and D'Esposito, M. (2000). Aging and reflective processes of working memory: binding and test load deficits. *Psychol. Aging* 15, 527–541.

Moscovitch, C., Kapur, S., Kohler, S., and Houle, S. (1995). Distinct neural correlates of visual long-term memory for spatial location and object identity: a positron emission tomography study in humans. *Proc. Natl. Acad. Sci. U.S.A.* 92, 3721–3725.

Munk, M. H., Linden, D. E., Muckli, L., Lanfermann, H., Zanella, F. E., Singer, W., and Goebel, R. (2002). Distributed cortical systems in visual short-term memory revealed by event-related functional magnetic resonance imaging. *Cereb. Cortex* 12, 866–876.

O'Craven, K. M., and Kanwisher, N. (2000). Mental imagery of faces and places activates corresponding stimulus-specific brain regions. *J. Cogn. Neurosci.* 12, 1013–1023.

Paivio, A. (1971). *Imagery and Verbal Process*. New York: Holt, Rinehart, and Winston.

Palmer, S. E. (1999). *Vision Science: Photons to Phenomenology*. Cambridge: MIT Press.

Posner, M. I., and Gilbert, C. D. (1999). Attention and primary visual cortex. *Proc. Natl. Acad. Sci. U.S.A.* 96, 2585–2587.

Prabhakaran, V., Narayanan, K., Zhao, Z., and Gabrieli, J. D. (2000). Integration of diverse information in working memory within the frontal lobe. *Nat. Neurosci.* 3, 85–90.

Pylyshyn, Z. W. (1973). What the mind's eye tells the mind's brain: a critique of mental imagery. *Psychol. Bull.* 80, 1–24.

Richter, W., Somorjai, R., Summers, R., Jarmasz, M., Menon, R. S., Gati, J. S., Georgopoulos, A. P., Tegeler, C., Ugurbil, K., and Kim, S. G. (2000). Motor area activity during mental rotation studied by time-resolved single-trial fMRI. *J. Cogn. Neurosci.* 12, 310–320.

Richter, W., Ugurbil, K., Georgopoulos, A., and Kim, S. G. (1997). Time-resolved fMRI of mental rotation. *Neuroreport* 8, 3697–3702.

Riesenhuber, M., and Poggio, T. (2000). Models of object recognition. *Nat. Neurosci.* 3(Suppl.), 1199–1204.

Roland, P. E., and Gulyas, B. (1995). Visual memory, visual imagery, and visual recognition of large field patterns by the human brain: functional anatomy by positron emission tomography. *Cereb. Cortex* 5, 79–93.

Sack, A. T. (2009). Parietal cortex and spatial cognition. *Behav. Brain Res.* 202, 153–161.

Sack, A. T. (2010). Using non-invasive brain interference as a tool for mimicking spatial neglect in healthy volunteers. *Restor. Neurol. Neurosci.* 28, 485–497.

Sack, A. T., Camprodon, J. A., Pascual-Leone, A., and Goebel, R. (2005). The dynamics of interhemispheric compensatory processes in mental imagery. *Science* 308, 702–704.

Sack, A. T., Jacobs, C., De Martino, F., Staeren, N., Goebel, R., and Formisano, E. (2008). Dynamic premotor-to-parietal interactions during spatial imagery. *J. Neurosci.* 28, 8417–8429.

Sack, A. T., Lindner, M., and Linden, D. E. (2007). Object- and direction-specific interference between manual and mental rotation. *Percept. Psychophys.* 69, 1435–1449.

Sack, A. T., Sperling, J. M., Prvulovic, D., Formisano, E., Goebel, R., Di Salle, F., Dierks, T., and Linden, D. E. (2002). Tracking the mind's image in the brain II: transcranial magnetic stimulation reveals parietal asymmetry in visuospatial imagery. *Neuron* 35, 195–204.

Sengpiel, F., and Hubener, M. (1999). Visual attention: spotlight on the primary visual cortex. *Curr. Biol.* 9, R318–R321.

Seurinck, R., De Lange, F. P., Achten, E., and Vingerhoets, G. (2011). Mental rotation meets the motion aftereffect: the role of hV5/MT+ in visual mental imagery. *J. Cogn. Neurosci.* 23, 1395–1404.

Shepard, R. N., and Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science* 171, 701–703.

Slotnick, S. D., Thompson, W. L., and Kosslyn, S. M. (2005). Visual mental imagery induces retinotopically organized activation of early visual areas. *Cereb. Cortex* 15, 1570–1583.

Smith, E. E., Jonides, J., Koeppe, R. A., Awh, E., Schuhmacher, E. H., and Minoshima, S. (1995). Spatial vs. objects working memory: PET investigations. *J. Cogn. Neurosci.* 7, 337–356.

Stokes, M., Saraiva, A., Rohenkohl, G., and Nobre, A. C. (2011). Imagery for shapes activates position-invariant representations in human visual cortex. *Neuroimage* 56, 1540–1545.

Tagaris, G. A., Kim, S. G., Strupp, J. P., Andersen, P., Ugurbil, K., and Georgopoulos, A. P. (1997). Mental rotation studied by functional magnetic resonance imaging at high field (4 Tesla): performance and cortical activation. *J. Cogn. Neurosci.* 9, 419–432.

Thompson, W. L., Slotnick, S. D., Burrage, M. S., and Kosslyn, S. M. (2009). Two forms of spatial imagery: neuroimaging evidence. *Psychol. Sci.* 20, 1245–1253.

Trojano, L., and Grossi, D. (1994). A critical review of mental imagery defects. *Brain Cogn.* 24, 213–243.

Trojano, L., Grossi, D., Linden, D. E., Formisano, E., Goebel, R., Cirillo, S., Elefante, R., and Di Salle, F. (2002). Coordinate and categorical judgements in spatial imagery. An fMRI study. *Neuropsychologia* 40, 1666–1674.

Trojano, L., Grossi, D., Linden, D. E., Formisano, E., Hacker, H., Zanella, F. E., Goebel, R., and Di Salle, F. (2000). Matching two imagined clocks: the functional anatomy of spatial analysis in the absence of visual stimulation. *Cereb. Cortex* 10, 473–481.

Tulving, E., Kapur, S., Craik, F. I., Moscovitch, M., and Houle, S. (1994). Hemispheric encoding/retrieval asymmetry in episodic memory: positron emission tomography findings. *Proc. Natl. Acad. Sci. U.S.A.* 91, 2016–2020.

Yantis, S., Schwarzbach, J., Serences, J. T., Carlson, R. L., Steinmetz, M. A., Pekar, J. J., and Courtney, S. M. (2002). Transient neural activity in human parietal cortex during spatial attention shifts. *Nat. Neurosci.* 5, 995–1002.

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: 27 April 2012; accepted: 08 June 2012; published online: 28 June 2012.
*Citation: Sack AT and Schuhmann T (2012) Hemispheric differences within the fronto-parietal network dynamics underlying spatial imagery. Front. Psychology **3**:214. doi: 10.3389/fpsyg.2012.00214*
This article was submitted to Frontiers in Perception Science, a specialty of Frontiers in Psychology.

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 15 August 2012
doi: 10.3389/fpsyg.2012.00296

Unmasking the Perky effect: spatial extent of image interference on visual acuity

Adam Reeves[1]* and Catherine Craver-Lemley[2]

[1] Department of Psychology, Northeastern University, Boston, MA, USA
[2] Department of Psychology, Elizabethtown College, Elizabethtown, PA, USA

Edited by:
Joel Pearson, The University of New South Wales, Australia

Reviewed by:
John Cass, University of Western Sydney, Australia
Adriane E. Seiffert, Vanderbilt University, USA

****Correspondence:***
Adam Reeves, Department of Psychology, Northeastern University, 360 Huntingdon Avenue, Boston, MA 02115, USA.
e-mail: a.reeves@neu.edu

We have previously argued that visual mental images are not substitutable for visual percepts, because the interfering effects of visual stimuli such as line maskers on visual targets differ markedly in their properties from the interfering effects of visual images (the "Perky effect"). Imagery interference occurs over a much wider temporal and spatial extent than masking, and unlike masking, image interference is insensitive to relative orientation. The lack of substitutability is theoretically interesting because the Perky effect can be compared meaningfully to real line masking in that both types of interference are visual, not due to optical factors (accommodative blur or poor fixation) or to high-level factors (attentional distraction, demand characteristics, or effects of uncertainty). In this report, however, we question our earlier position that spatial extents of interference are markedly different: when images and real lines are matched in contrast, which was not done previously, their interference effects have very similar spatial extents. These data add weight to the view that spatial properties of images and percepts are similar in respect to extent. Along with the wider temporal extent and the insensitivity to orientation, the new results remain compatible with our older hypothesis that to create a clear mental image in a region of visual space, incoming signals from the eye must be suppressed (Craver-Lemley and Reeves, 1992). We have pursued this idea in this report using "unmasking," in which adding elements to the visual image in the region beyond the zone of suppression reduces the Perky effect.

Keywords: visual imagery, Perky effect, spatial vision, unmasking

INTRODUCTION

Whether or not visual mental images act like visual percepts is a fascinating and important question in cognitive science. Should they have the same functions and share the same anatomy, as has been argued from imaging studies of cortical area V1, one might expect them to be mutually substitutable or at least, analogous to each other. That images and percepts both have spatial layout, both contain rich visual information, and both share the functional property of reduced acuity in the periphery (Kosslyn, 1980), may be taken to argue in favor of substitutability; in Kosslyn's array theory, for example, images and percepts are thought to share the same representational array, and in Finke's (1980) theory, images and percepts are functionally equivalent for high imagers. In our earlier work, we argued in favor of analogous spatial behavior but against substitutability (Arterberry et al., 2002), based on empirical findings concerning the so-called Perky effect, in which visual images, analogous to real masking stimuli, can suppress perception of real visual targets (as discovered by Perky, 1910). The Perky effect is ubiquitous in studies of percept-image interactions, with the imagined stimulus depressing sensitivity to briefly flashed visual targets by 15% or better (0.8 d′ units or more; Segal and Fusella, 1970; Craver-Lemley and Reeves, 1987, 1992; Ishai and Sagi, 1997). In the Perky experiments, the experimenter requested that the subject project his or her visual image in the location of the visual target. Thus when the target was flashed, it was presumably presented at the same location in visual space that the image already occupied (the image-on-target, or "ON" condition). Such images always interfere with perception of the target by lowering sensitivity, not by making the response criterion sub-optimal. Interference occurs in the brain, not the eye, as controls have ruled out optical effects such as poorer accommodation of the lens during imagery or poorer fixation during imagery (Craver-Lemley and Reeves, 1987). Moreover, the Perky effect does not stem from demand characteristics; participants told that imagery aided perception produced as great a Perky effect as those told it would hinder perception (Craver-Lemley and Reeves, 1987, Experiments 1a and 1b). Other experiments have ruled out mere distraction, the notion that having any image would add a load to the cognitive system and thus lower visual performance, as auditory images have virtually no effect on visual target (Craver-Lemley and Reeves, 1992; although they do on auditory ones: Segal and Fusella, 1970), visual images located far away from the visual target have almost no effect (Craver-Lemley and Reeves, 1987; Craver-Lemley and Arterberry, 2001), and diverting attention to a distracting light does not increase the interference (Craver-Lemley and Reeves, 1992). Thus we have used the Perky effect as a test-bed for discovering whether images and real visual stimuli have the same functional properties. We adopted an acuity target as the visual stimulus, since the visual properties of acuity stimuli are well documented[1].

Experiments using real (rather than imagined) stimuli have revealed several well-known characteristics of masking of line targets, including that masking is reduced rapidly as the masking stimuli are moved away laterally from the target; that masking is reduced if the orientation of the mask differs from that of the target, disappearing if they are at right-angles to each other (although this general rule may not be true above threshold for dichoptic low-spatial frequency gratings: Meese and Hess, 2004); that masking is eliminated if the target-to-mask temporal asynchrony is more than 200 ms; and that patterned masking is cortical – unlike noise masks, patterned masks are as effective when the target is presented to the opposite as to the same eye (see, e.g., Westheimer, 1965; Westheimer and Hauske, 1975). Relying on this literature, we argued that interference due to images (the Perky effect) had quite different characteristics from real masking (Craver-Lemley and Reeves, 1992), in that the Perky effect is independent of the relative orientation of the image and the real lines, and the Perky effect persists up to 6 s after the image has been removed (Craver-Lemley and Reeves, 1987). Moreover, the Perky effect apparently covers a wider region of visual space than does the interfering effects of real lines (Craver-Lemley and Reeves, 1992).

In this past research, we asked participants to create visual images but we have not matched the contrasts of the mental images to the contrast of the physical stimuli. Typically, we employed high contrast physical stimuli, whereas participants created mental images *ad lib*, and our participants had informally reported that their imagined lines were lighter and less distinct than those pictured (Craver-Lemley and Reeves, 1987, 1992). We therefore wondered if our complete rejection of substitutability was too hasty, the difference being, not between real and imaged stimuli, but between high and low contrasts. We did not repeat our earlier work on image orientation because it is already established that masking by low-contrast gratings, like masking by high contrast ones, disappears at the orthogonal orientation, implying that our earlier result (orthogonal images interfered as much as parallel ones) would differ from real line masking no matter what real line contrast was used. We also did not repeat the work comparing timing, as the Perky effect lasts for 4–6 s after an image has been terminated, long after any masking by real lines is complete. However, we were concerned to repeat our work on the lateral extent of interference. We therefore compared the effects of images of lines against not only the real black lines used before, but also against real gray lines matched in contrast with the participant's own imagined lines (Experiment 1). We continued to use lines rather than other acuity targets for continuity with our previous work; however, we note here that the spatial structure of the stimulus may affect the lateral extent of interference. While Ishai and Sagi (1997), like us, found considerable interference by imagery in the "ON" condition when using Gabor (wavelet) targets, they also found a small but significant *enhancement* in sensitivity when the mental image was displaced just one wavelength away from the target center. With lines, interference is also reduced by increasing target-image distance, but we have not found enhancement.

EXPERIMENT 1

This experiment was conducted to find out whether gray lines that were matched in contrast to images of lines would interfere with acuity in the same manner as the mental images. Black real lines were also used to replicate previous work. Target-image and target-real line distances were varied to permit comparison of the lateral extents of image interference and real line masking.

MATERIALS AND METHODS

PARTICIPANTS

Eleven undergraduate volunteers enrolled in an introductory psychology course participated in this study for payment. All participants had normal or corrected-to-normal vision and had no previous experience in vision or imagery experiments.

APPARATUS AND STIMULI

Experiments were conducted with Model T-2B-1 two-field Gerbrands tachistoscopes. The fixation and test fields were 58 cm from the eyes, superimposed by a half-silvered mirror, and surrounded by complete darkness (the inner walls of the apparatus were covered in black velvet). The *fixation* field was a white 17 cd/m^2 rectangular index card that subtended 10° × 15° at the eye. Central fixation was aided by two small (1.22 mm) black dots placed 5.2° apart, symmetrically above and below the center of this field. Small guide dots (Experiments 1 and 2) or guide lines (Experiment 3) were added at top and bottom of the fixation field to aid the subjects in locating their images when images were requested.

There were two real line conditions, in which imaginary lines were not requested, namely, Black Lines and Matched Gray Lines. In these conditions, 1.0 cm wide and 7.6 cm high black or gray lines, subtending 0.1 × 7.5°, were added to the fixation field. Being part of the fixation field, these lines were presented continuously, unlike the vast majority of studies of masking in which masking lines are flashed; however, continuous presentation was necessary to mimic the imaginary lines, which were imagined throughout each trial, as closely as possible.

The *test* field was used to flash the vernier acuity targets, as in Craver-Lemley and Reeves (1992); Craver-Lemley et al. (1997). The test field was dimmed to 6 cd/m^2 so that flash duration at threshold would not be lower than 10 ms for the best participants.

---

[1] Recent research has re-emphasized Hume's view that the vividness of mental imagery is functional. David Hume had argued that images were less vivid than the reality they represented, and that such a difference was sufficient to distinguish truth from fantasy. Perky (1910), however, showed the participants do not always reliably distinguish images from percepts, and perhaps more tellingly, reported image vividness does not in general correlate with perceptual ability though it may with memory (Hilgard, 1981). Indeed, we (Craver-Lemley and Reeves, 1987) found no correlation between the self-reported vividness of visual imagery and the extent of the Perky effect across 125 participants. However, Rodway et al. (2006) recently reported that while self-reported high and low vividness participants did not differ in overall levels of change detection accuracy, high vividness participants were more accurate at detecting salient changes to pictures. Moreover, self-reports of image vividness correlate $r = 0.73$ with image strength as documented by the fMRI (Cui et al., 2007), and stronger visual images are more likely to survive size-scaling than are weaker ones (D'Angiulli and Reeves, 2002). The strength of simple line images, as manipulated by the participants themselves, therefore seemed like a possible factor in our experiments.

Each target was made of two, thin black (75% contrast) lines, each line subtending 2.2° vertically and 0.1° horizontally, with a vertical gap of 0.5° between them. As the vertical gap between the acuity targets is relatively large, hyperacuity was not achieved. The targets (black tape by Chart-Pak) were mounted on white index cards. The lower of the two lines in each target was, with equal probability, offset to the left or right of the upper line. There were 10 different offsets ranging from 4.2 to 22 min visual angle, with a mean offset of 9.7 min. The target presentation was randomized. The participant's task was to report whether the lower line was to the right or left of the upper line; they were told that each was equally likely to occur. The test field was jittered randomly left or right by up to 17 arc min from trial to trial, to encourage the participant to judge the lines relative to each other rather than relative to the fixation point.

Conditions

There were 10 conditions, run in separate blocks:

Imaged lines. During imagery trials, participants were asked to evoke a mental image of four black vertical lines. Participants were shown a picture of the lines they were to imagine and asked to locate their image in reference to small guide marks penciled in at the top and bottom of the fixation field. Each participant's acuity was measured while they were instructed to project mentally vertical line images either "ON" the acuity target (nearest line 0.1 cm, or 6 min arc, away from the target), "CLOSE" to the target (nearest line 0.8 cm, or 0.8°, away), or "FAR" from it (innermost line 2.0 cm, or 2.0°, away; see **Figure 1**). Guide marks indicated to the participants where to place their mental images. They were given ample time to comply, and none reported difficulty in locating their imaginary lines with respect to the guide marks. They were instructed to hold their images throughout the trial; if reported losing the image, the trial was re-run.

Black lines. During black line trials, the four real, thin black lines were shown continuously from top to bottom of the fixation field. These lines appeared either "ON" the acuity target, "CLOSE" to it, or "FAR" from it, with the same spacing as for the imaged lines.

Matched gray lines. During the initial practice session, we asked each participant to select lines from a prepared set (ranging in color from light gray to black) which best matched his or her imagined lines. All participants selected real gray lines that were lighter than the black lines, matching images to Munsell color chips between 10B 8/1 and N8/5. As in the Imagery and Black Lines conditions, the Matched Gray Lines (at the level selected by each participant) were presented in the same ON, CLOSE, and FAR conditions.

No imagery. In the baseline condition, accuracy was measured for reporting the acuity target, without either real or imagined lines.

Thus the 10 conditions comprised the baseline with no lines and no imagery, three conditions with imagined lines (ON, CLOSE, FAR), three with Black Lines (ON, CLOSE, FAR), and three with Matched Gray Lines (again, ON, CLOSE, FAR).

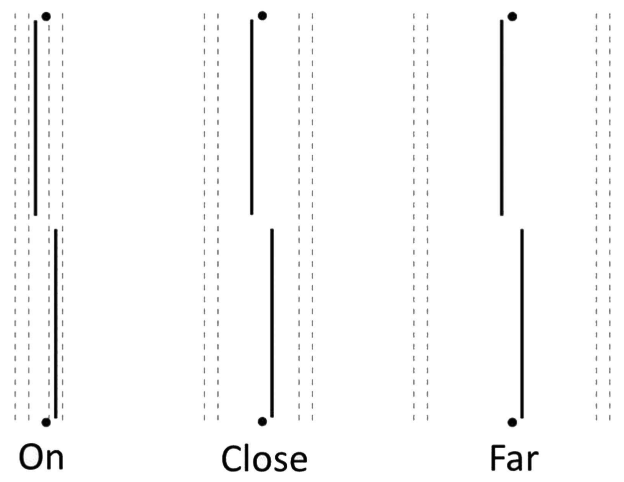


FIGURE 1 | Two-line acuity targets surrounded by imaginary lines, ON, CLOSE, or FAR away, as directed by small guide spots or lines at top and bottom of the fixation field (not shown). Real black or gray lines were positioned in the same places as the imagined ones.

PROCEDURE

Participants were individually tested during two 1 h sessions. The duration of target presentation was determined during an additional 15 min practice session. Participants were told to fixate between two fixation points where the target would be presented. The target duration was adjusted until accuracy in the baseline no imagery acuity task was close to 90%. Target duration remained fixed after the practice session; it averaged a nominal 8 ms (range: 3–25 ms; as these values are well below the critical duration, target duration exerts its effect via energy). Participants reported the direction of offset (left, right) on each trial.

Conditions were run in separate blocks of 25 trials in each of two sessions. The order of the blocks was randomized over participants. There were 25 trials per participant in the black lines ON condition, which, being near chance in the first session was not re-run, however 50 trials were included for each of the other nine conditions.

Instructions were reviewed before each block. In the No Image and the two real line conditions, the experimenter initiated each trial by saying, "go" and then the acuity target was presented. In the imagery condition, participants indicated their image was ready by saying "go," and the experimenter then presented the acuity target. Participants were requested to re-create or re-evoke their images on each of the 25 imagery trials during a block. Participants were allowed to rest between blocks. They did not receive any feedback until after the second session had been completed.

Our experimenters were undergraduate research assistants who were naïve as to the predicted outcomes of the experiments and had never before collected data in imagery experiments. They could not see the acuity target until after each trial had been completed, and so were unaware of what the correct response should be prior to stimulus presentation and the participant's response. Individual data were not tabulated or analyzed until after the experiment was completed.

RESULTS AND DISCUSSION

Accuracy in the baseline (No Image) condition averaged 91.4% correct across participants, reflecting successful choices of target durations during initial practice. Accuracy in all other conditions was lower than this, reflecting the interfering effects of real and of imagined lines. An omnibus analysis of variance (ANOVA) revealed a main effect of both distance [$F(2,9) = 191.54, p < 0.001, \eta_p^2 = 0.95$] and line type [$F(2,9) = 23.03, p < 0.001, \eta_p^2 = 0.70$], as well as an interaction [$F(4,7) = 45.66, p < 0.001, \eta_p^2 = 0.82$]. Mean percent correct in each of the conditions is plotted in **Figure 2**[2]. As expected, the results show a clear Perky effect, as imagined lines reduced acuity. The reduction is 17.6% when the imagined lines are "ON" the target, in line with Segal and Fusella (1970), Reeves (1980), and Craver-Lemley and Reeves (1987, 1992). The Perky effect was attenuated if the image was shifted off target, in the present research from 17.6 to 6.3% (i.e., baseline No Image – image OFF), also in line with Craver-Lemley and Reeves (1987).

The critical new result is that (real) Matched Gray lines have rather similar effects on acuity as imagined lines, reducing performance by 13.2% in the ON condition and 5% in the FAR condition, compared to 17.6–6.3%, respectively. Given the difficulties inherent in exactly matching the contrast and spatial position of the real lines, one is struck by the congruence between these effects and those for imagined lines. This result causes us to re-think our original claim that the spatial extent of interference by images exceeds that of real lines; contrast, not reality, appears to be the critical variable, and when contrasts are matched, the reductions in interference with distance from the target are similar.

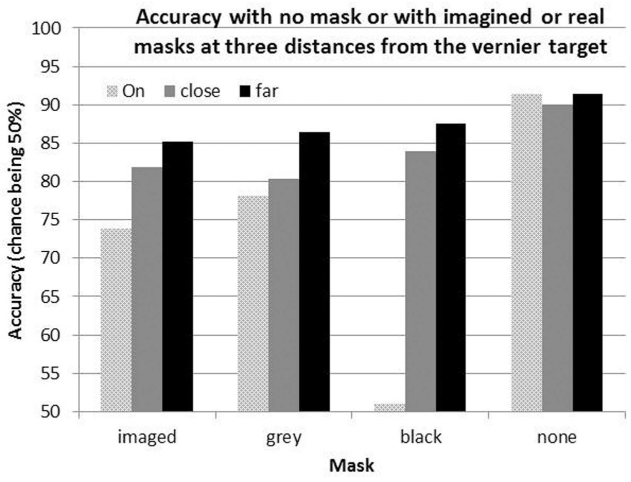


FIGURE 2 | Mean correct performance for the acuity task with imaged lines, with real gray lines matched to the participant's images, with real black lines, and in the baseline no image, no mask ("none") condition. The interfering lines are ON the acuity target, CLOSE to it, or FAR from it.

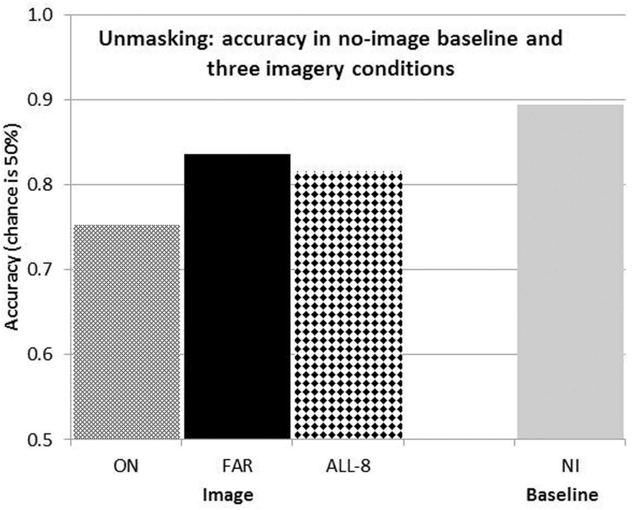


FIGURE 3 | Mean correct performance for the acuity task with ON and FAR images, each of four lines, with ALL eight imagined lines, and in the baseline no image condition.

EXPERIMENT 2

Experiment 1 suggests that contrast plays a major role in the Perky effect, but does contrast polarity also matter? We had not compared images with positive and negative contrasts in previous studies of the Perky effect, nor have we located a study that has. We therefore asked participants to create images of white vertical lines for comparison with the images of black vertical lines used before, projecting both onto a gray field. With real stimuli, acuity depends strongly on the absolute value of the contrast, but when targets are presented on a constant background, acuity is only weakly dependent on contrast polarity (Pointer, 2001), as is metacontrast (Breitmeyer et al., 2008: note the same polarity masks and targets in his **Figure 3**) and object substitution masking (Backmann and Luiga, 2008). (We note that earlier studies of polarity compared acuity for white targets on a black background with acuity for black targets on a white background, but these backgrounds alter adaptation level, light scatter, and pupil size; a constant background controls these factors.) If the analogy between imagery and real lines holds true of contrast, we therefore expect that contrast

---

[2] A detailed analysis of the data of Experiment 1 reveals the effects expected from eyeballing **Figure 2**. In the ON conditions, black lines showed significantly more interference than both matched lines [$t(10) = 7.41, p < 0.001$], and imagined lines [$t(10) = 5.71, p < 0.001$], and imagined lines showed significantly more interference than matched lines [$t(10) = -6.71, p < 0.001$]. In the CLOSE conditions, matched lines showed significantly more interference than black lines [$t(10) = -2.86, p = 0.017$], imagined lines showed significantly more interference than black lines [$t(10) = -3.32, p = 0.008$], and matched lines and imagined lines were not significantly different, $t(10) = 1.44, p = 0.181$. In the FAR conditions, black lines and matched lines were not significantly different [$t(10) = -1.35, p = 0.208$], but imagined lines showed significantly more interference than black lines [$t(10) = -2.71, p = 0.022$] and matched lines [$t(10) = -2.28, p = 0.046$]. Images ON the target showed more interference than both images CLOSE to it [$t(10) = 4.58, p = 0.001$] and FAR from it [$t(10) = 5.14, p < 0.001$]; also, CLOSE images showed more interference than FAR images [$t(10) = 4.50, p = 0.001$]. Matched Gray Lines ON and CLOSE to the target did not differ [$t(10) = 1.36, p = 0.204$], but they interfered more than those FAR from it [$t(10) = 3.9, p = 0.003$, and $t(10) = 4.54, p = 0.001$]. Real Black Lines ON the target interfered more than both those CLOSE to it [$t(10) = 11.73, p < 0.001$] and FAR from it [$t(10) = 14.66, p < 0.001$], and those CLOSE interfered more than those FAR [$t(10) = 7.61, p < 0.001$]

polarity of the image would leave the magnitude of the Perky effect relatively unchanged, given that we use a constant lit background.

MATERIALS AND METHODS

PARTICIPANTS

Twenty-two undergraduates with normal or corrected-to-normal vision volunteered to participate in exchange for course credit in an introductory psychology course. None of the students had previously participated in an imagery experiment before.

MATERIALS AND PROCEDURE

The stimuli and apparatus were as in Experiment 1. Participants were again tested individually in single 45 min. sessions. The acuity targets were presented for a duration that resulted in close to 90% correct performance without imagery, chosen for each participant during an initial 15 min practice session. The mean nominal target duration was 9.3 ms (range = 3–30 ms). The task was, as before, to report whether the lower line was offset to the left or right of the upper line.

There were three conditions: No Image (the baseline), Imagine Black Lines, and Imagine White Lines, ON the target. The three conditions were not blocked but were randomly intermixed. There were 50 trials per condition for a total of 150 trials for each participant. Participants were cued prior to the beginning of each trial as to whether they were to imagine four vertical white lines, four vertical black lines, or to have no image at all. During imagery trials, participants indicated when they were ready with the image, and then the experimenter presented the target. Participants were instructed to maintain the image for the duration of the trial. On No Image trials, the experimenter indicated it was a no image trial and then said, “go” to alert the participant that the trial was beginning. Participants received a 5 min break after 75 trials.

RESULTS AND DISCUSSION

Correct performance, averaged over targets and participants was 89.1% with no image, 80.8% for Imagine Black Lines, and 79.0% for Imagine White Lines. A significant condition effect, $F(2,42) = 23.71$, $p < 0.001$, was found in a one-way repeated measures ANOVA, with both Perky effects significant by Tukey’s honestly significant difference test ($p < 0.01$). However, accuracy with Black line images (80.8%) did not differ from that with White line images (79.0%).

The Perky effect for Imagine Black Lines, being 8.3%, was smaller than the 17.6% effect found for ON images in Experiment 1. Since otherwise these conditions were the same, this difference in magnitude may be due to randomizing rather than blocking the conditions. Randomizing may reduce magnitude because the Perky effect lasts several seconds after the participant reports turning off the image (Craver-Lemley and Reeves, 1987), and therefore some image-generated interference may leak into those baseline trials which directly follow imagery trials. However, randomizing has the advantage that the lack of a difference between black images and white images is unlikely to be due to a change in strategy between conditions. We conclude that the lack of a difference is genuine, and therefore that contrast, but not contrast polarity, influences the Perky effect. This finding for imagery appears to agree with the literature concerning the effects of contrast polarity on acuity for real stimuli, but this agreement is not definitive as we have not studied the effects of contrast polarity with real lines in our equipment and our target-line distances.

EXPERIMENT 3

If the analogy between real and imagined lines does hold up, at least for extent if not for orientation, then it is possible that imagined lines, like real ones, might show “unmasking” (also termed “release from masking.”) In this phenomenon, a real mask superimposed on the target has its masking effect reduced by adding a distant mask, even when the latter had little direct effect on target visibility (Haber, 1970). It is as if the distant mask releases the target from masking, either by inhibiting the superimposed mask (Dember and Purcell, 1967), or by strengthening the target (Briscoe et al., 1983). Herzog (2007) reviews unmasking in a more complex spatial arrangement when the vernier is masked by a spatial grating. We wondered whether the same phenomenon of unmasking could be found with visual images.

MATERIALS AND METHODS

PARTICIPANTS

Sixteen undergraduates with normal or corrected-to-normal vision volunteered to participate in exchange for course credit in an introductory psychology course. None of the students had previously participated in an imagery experiment before.

MATERIALS

A Gebrands two-field tachistoscope was used, of the same type as the one used in Experiments 1 and 2. Stimuli were the same as before. On imagery trials, participants were asked to evoke four or eight black vertical line images. Small guide lines in the fixation field helped the participants locate their images. Participants imagined four vertical lines in the ON and FAR conditions illustrated in **Figure 1**, but in the *new* condition, ALL-8, they were requested to generate an image of eight black lines, as indicated by the guide lines. Ideally this new image is the sum of ON and FAR images.

PROCEDURE

As before, participants were tested individually in a single session lasting approximately 50 min. The acuity targets were presented for a duration that resulted in close to 90% correct performance in the baseline (no imagery) condition, which was chosen for each participant during an initial 15 min practice session and thereafter fixed. The mean duration of targets was 13 ms (range = 4–33 ms). All three imagery conditions were practiced until the participant was familiar with the procedure. Participants were asked to maintain central fixation during the experiment. The task was, as in Experiments 1 and 2, to report whether the lower line was offset to the left or right of the upper line. The three imagery and one no imagery conditions were randomly intermixed with 50 trials per condition for each participant. Participants were cued prior to the beginning of each trial as to what to image, or to have no image at all. During imagery trials, participants indicated when they were ready with the image and then the experimenter presented the target. Participants were instructed to maintain the image for the duration of the trial. On baseline trials, the experimenter indicated it was a no image trial and then said, “go” to alert the participant that the trial was beginning. Participants received 5 min breaks after each block of 50 trials.

RESULTS

The differences between the four conditions were highly significant [$F(3,45) = 9.48$, $p < 0.01$] by a repeated measures, one-way ANOVA applied to the scores of the 16 participants. Mean accuracy was 89.4% in the baseline "No imagery" condition, reflecting successful manipulation of the individual target durations. Accuracy with ON images was 75.3%, giving a reduction in accuracy, or Perky effect, of 14.2% (**Figure 3**). When the lines were imaged far from the target, accuracy was 83.6%, for a much smaller Perky effect of 5.8%. These two Perky effect magnitudes are similar to those reported before by Craver-Lemley and Reeves (1987) for these two conditions. They suggested to us that the Perky effect is spatially localized. The new condition is ALL-8, in which both ON and FAR lines were imaged; in this case, mean accuracy was 81.6%, for a Perky effect of 7.8%.

The Perky effect in ALL-8 would be greater than in ON if the effect of the inner lines summed with that of the outer lines, because the inner lines were the same in ON as in ALL-8, and the outer lines had a weak (7.6%) effect of their own. However, the results showed a clear diminution of the Perky effect in ALL-8, to a mere 7.8%, which is much less than additive.

DISCUSSION

There are three possible explanations for our new result, that the Perky effect with eight imaged lines is less than additive: suppression, distraction, and weakening. We will ultimately argue for an explanation in terms of suppression, but the other two possibilities merit consideration, especially as the result from Experiment 3 is singular and we therefore do not have an adequate sampling of experimental conditions to permit a definitive conclusion.

By "weakening" we mean that relative to the strength of the ON image, that of the ALL-8 image is reduced, either by graying, blurring, making less vivid, decreasing image duration, or otherwise lowering image quality. The ALL-8 image could in principle be weaker because maintaining eight vertical line images is simply more taxing than maintaining four. However, both in initial practice and in debriefing, we found no evidence for this: all of the participants reported being able to imagine eight lines as well as four. In debriefing, we again asked whether the three images (ON, FAR, and ALL-8) had equal vividness, and had kept their vividness throughout the experiment; participants reported that they had. Such retrospective reports do not disprove the possibility of some weakening, but as explicitly weakening the ON image by asking participants to gray-out their imagery reduced the Perky effect by only 4%, as we found in a pilot experiment, we think the consequences of any slight weakening would be marginal.

By "distraction" we mean that having a vertical line image might distract attention from the primary task of reporting the acuity target. We originally argued that images required the same amount of attention, and were equally distracting, whether they were near the acuity target or far from it, so differences between them in the magnitude of the Perky effect could not be explained by attention (Craver-Lemley and Reeves, 1987, 1992). One might dispute this logic, but note that ON images demand that attention be paid to the area in which the target will be presented, as opposed to FAR images which draw attention away, so distraction would predict more, not less, of a Perky effect in FAR than in ON. Moreover, Craver-Lemley and Reeves (1992) explicitly manipulated attention by having participants attend to a blinking light in the periphery; this lowered accuracy overall but left the Perky effect unaltered. Again, Craver-Lemley and Reeves (1987) found that a vertical line image produced a Perky effect for 4–6 s after the subject indicated that he or she had removed the image, and in this situation there was no image to distract attention. Finally, Craver-Lemley et al. (1997) found that having vertical line images in front of the visual target produced a Perky effect, but having them behind the target had no interfering effect at all. Since to control location in depth, the same image was requested on one or other face of an outlined cube, it is unlikely that differential attention to the image could explain this result. We regard all this as good evidence that the Perky effect is not entirely due to distraction of attention.

By "suppression," we invoke a hypothetical mechanism in which the inner four lines in ALL-8 have their interfering effects on the target reduced by the outer four lines. Although it is somewhat counterintuitive that adding masking elements could reduce masking, this effect is known in the masking of visual targets by real maskers (Dember and Purcell, 1967; Herzog, 2007), and the same idea has been applied before to the unmasking of real stimuli by visual images (Reeves, 1980). Either the outer lines "unmask" the target by suppressing the interfering effect of the inner lines, or the outer lines act to facilitate the target. Facilitation by distant imagery is theoretically possible and has been observed using Gabor-like stimuli and images by Ishai and Sagi (1997); however, their effect, thought theoretically important, was relatively small. We are inclined to conclude that unmasking occurs when the active part of the mask is itself masked due to some form of inhibition by additional mask elements, both for real lines and for imaged ones. However, the idea that the far away image elements inhibit the ones close to or on the acuity target is based on this single result, and this may not survive further testing; and even if it does, the properties of image unmasking may turn out quite different from those of real line unmasking.

GENERAL DISCUSSION

Craver-Lemley and Reeves (1992) suggested that the Perky effect is a consequence, not of the image directly, but of a mechanism designed to *protect the image* by inhibiting visual input from the region of the visual field in which the subject was requested to have the image. Since we found the Perky effect was insensitive to the relative orientation of imaged lines and real line acuity targets, we suggested that the protection was provided by inhibitory cortical feedback to the LGN, the last relay in the visual pathway which is not sensitive to orientation. Our measurements suggested that the feedback effect was weak, amounting to a reduction in stimulus contrast of 0.24 log units, but this was enough to account for the reduction in acuity that we had reported. As already remarked, we have rejected other explanations of interference, such as poorer optical accommodation, shifts of gaze away from the target, confusion of the target with the image, distraction of attention, response bias, and experimenter effects, all based on empirical data.

However, the notion of a protective mechanism operating at the level of the LGN did not predict the finding of Craver-Lemley et al. (1997) that having vertical line images behind the visual target eliminates the Perky effect, because LGN neurons are insensitive to depth as well as to orientation. The model could be modified such that the feedback signal itself is dependent on depth, such that only frontal images need protection from visual input, although this is quite *ad hoc*. However, the new result, of unmasking, cannot be explained in such a manner. An alternative hypothesis suggested by Perky (1910) and by Segal (1971), and resurrected by Craver-Lemley et al. (1997) in order to explain the effect of depth, is that the Perky effect results from a combination of real and imagined features that makes the real features more difficult to extract. This notion is attractive in many ways, but it presupposes that imagined features have much more featural spread than real ones, since, unlike real lines, horizontal line images interfere as much with acuity for vertical lines as do vertical line images. It also supposes that some unknown after-effects of imaged lines, not just concurrent images, combine with stimulus features, as images continue to interfere for 4–6 s after being removed (Craver-Lemley and Reeves, 1987). It may be that some combination of the two hypotheses, or perhaps an altogether different one, will eventually explain the intricacies of the Perky effect.

REFERENCES

Arterberry, M. E., Craver-Lemley, C., and Reeves, A. (2002). Visual imagery is not always like visual perception. *Behav. Brain Sci.* 25, 183–184.

Backmann, T., and Luiga, I. (2008). Luminance processing in object substitution masking. *Vision Res.* 48, 937–945.

Breitmeyer, B. G., Tapia, E., Kafaligönül, H., and Ögmen, H. (2008). Metacontrast masking and stimulus contrast polarity. *Vision Res.* 48, 2433–2438.

Briscoe, G., Dember, W., and Warm, J. S. (1983). Target recovery in visual backward masking: no clear explanation in sight. *J. Exp. Psychol. Hum. Percept. Perform.* 9, 898–911.

Craver-Lemley, C., and Arterberry, M. E. (2001). Imagery-induced interference on a visual detection task. *Spat. Vis.* 14, 101–119.

Craver-Lemley, C., Arterberry, M. E., and Reeves, A. (1997). Effects of imagery on vernier acuity under conditions of induced depth. *J. Exp. Psychol. Hum. Percept. Perform.* 23, 3–13.

Craver-Lemley, C., and Reeves, A. (1987). Visual imagery selectively reduces vernier acuity. *Perception* 16, 599–614.

Craver-Lemley, C., and Reeves, A. (1992). How visual imagery interferes with vision. *Psychol. Rev.* 99, 633–649.

Cui, X., Jeter, C. B., Yang, D., Montague, P. R., and Eagleman, D. M. V. (2007). Vividness of mental imagery: individual variability can be measured objectively. *Vision Res.* 47, 474–478.

D'Angiulli, A., and Reeves, A. (2002). Generating visual mental images: latency and vividness inversely related. *Mem. Cognit.* 30, 1179–1188.

Dember, W. N., and Purcell, D. G. (1967). Recovery of masked visual targets by inhibition of the masking stimulus. *Science* 757, 1335–1336.

Finke, R. A. (1980). Levels of equivalence in imagery and perception. *Psychol. Rev.* 87, 113–132.

Haber, R. N. (1970). Note on how to choose a visual noise mask. *Psychol. Bull.* 74, 373–376.

Herzog, M. H. (2007). Spatial processing and visual backward masking. *Adv. Cogn. Psychol.* 3, 85–92.

Hilgard, E. R. (1981). Imagery and imagination in American psychology. *J. Ment. Imagery* 5, 5–66.

Ishai, A., and Sagi, D. (1997). Visual imagery facilitates visual perception: psychophysical evidence. *J. Cogn. Neurosci.* 9, 476–489.

Kosslyn, S. M. (1980). *Image and Mind.* Cambridge, MA: Harvard University Press.

Meese, T. S., and Hess, R. F. (2004). Low spatial frequencies are suppressively masked across spatial scale orientation, field position, and eye of origin. *J. Vis.* 4, 843–859.

Perky, C. (1910). An experimental study of imagination. *Am. J. Psychol.* 21, 422–452.

Pointer, J. S. (2001). The influence of level and polarity of figure-ground contrast on vision. *Acta Ophthalmol. Scand.* 79, 422–425.

Reeves, A. (1980). Visual imagery in backward masking. *Percept. Psychophys.* 28, 118–124.

Rodway, P., Gillies, K., and Schepman, A. (2006). Vivid imagers are better at detecting salient changes. *J. Individ. Differ.* 27, 218–228.

Segal, S. J. (1971). *Imagery: Current Cognitive Approaches.* New York, NY: Academic Press.

Segal, S. J., and Fusella, V. (1970). Effects of imaging on signal-to-noise ratio, with varying signal conditions. *Br. J. Psychol.* 60, 459–464.

Westheimer, G. (1965). Visual acuity. *Annu. Rev. Psychol.* 16, 359.

Westheimer, G., and Hauske, G. (1975). Temporal and spatial interference with vernier acuity. *Vision Res.* 15, 1137–1141.

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 May 2012; accepted: 30 July 2012; published online: 15 August 2012.
Citation: Reeves A and Craver-Lemley C (2012) Unmasking the Perky effect: spatial extent of image interference on visual acuity. Front. Psychology ***3****:296. doi: 10.3389/fpsyg.2012.00296*
This article was submitted to Frontiers in Perception Science, a specialty of Frontiers in Psychology.

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 10 July 2012
doi: 10.3389/fpsyg.2012.00224

Training visual imagery: improvements of metacognition, but not imagery strength

***Rosanne L. Rademaker* [1,2,3] ** and Joel Pearson* [2,3,4]**

[1] *Cognitive Neuroscience Department, Maastricht University, Maastricht, Netherlands*
[2] *Psychology Department, Vanderbilt University, Nashville, TN, USA*
[3] *Vanderbilt Vision Research Center, Vanderbilt University, Nashville, TN, USA*
[4] *School of Psychology, The University of New South Wales, Sydney, NSW, Australia*

Edited by:
Stephen Michael Kosslyn, Stanford University, USA

Reviewed by:
Stephen Michael Kosslyn, Stanford University, USA
Gregoire Borst, Université Paris Descartes, France

****Correspondence:***
Rosanne L. Rademaker, Cognitive Neuroscience Department, Maastricht University, Universiteitssingel 40, 6229 ER Maastricht, Netherlands.
e-mail: rosanne.rademaker@maastrichtuniversity.nl

Visual imagery has been closely linked to brain mechanisms involved in perception. Can visual imagery, like visual perception, improve by means of training? Previous research has demonstrated that people can reliably evaluate the vividness of single episodes of imagination – might the metacognition of imagery also improve over the course of training? We had participants imagine colored Gabor patterns for an hour a day, over the course of five consecutive days, and again 2 weeks after training. Participants rated the subjective vividness and effort of their mental imagery on each trial. The influence of imagery on subsequent binocular rivalry dominance was taken as our measure of imagery strength. We found no overall effect of training on imagery strength. Training did, however, improve participant's metacognition of imagery. Trial-by-trial ratings of vividness gained predictive power on subsequent rivalry dominance as a function of training. These data suggest that, while imagery strength might be immune to training in the current context, people's metacognitive understanding of mental imagery can improve with practice.

Keywords: visual imagery, training, learning, metacognition, introspection, binocular rivalry, consciousness

INTRODUCTION

Mental imagery can be described as the retrieval of perceptual information from memory, and the subsequent examination of this information in the "minds eye." Research has provided a growing body of behavioral and neuroimaging evidence that there is considerable overlap between the "minds eye" and actual perception (Chen et al., 1998; Kreiman et al., 2000; O'Craven and Kanwisher, 2000; Zatorre and Halpern, 2005). For example, behavioral studies have demonstrated that imagery content can selectively influence perception (Perky, 1910; McDermott and Roediger, 1994; Pearson et al., 2008). Imagery has been shown to affect visual detection thresholds (Ishai and Sagi, 1997), performance on a visual acuity task (Craver-Lemley and Reeves, 1992), and to induce negative aftereffects (Gilden et al., 1995) in much the same way as a sensory stimulus. Recent neuroimaging studies show that there is considerable spatial overlap between activated areas of the brain during both visual perception and visual imagery, for example information about a pattern held in mind during working memory or imagery can be present in visual sensory cortex (Kosslyn et al., 1995; Slotnick et al., 2005; Harrison and Tong, 2009; Serences et al., 2009; Stokes et al., 2009). Like perception, visual imagery is impaired when visual cortical activity is disturbed by means of transcranial magnetic stimulation (Kosslyn et al., 1999).

If visual imagery can indeed be defined as the recreation of a perceptual representation in the absence of retinal input (Ishai and Sagi, 1995), one may wonder exactly how similar imagery is to perception. Specifically, prolonged visual practice can improve perceptual skill (Fahle and Poggio, 2002; Fine and Jacobs, 2002; Sasaki et al., 2010); can imagery also improve with daily practice? There is some evidence to suggest that perceptual learning can occur from training without physical stimulation. Repetitively imagining the crucial part of a visual bisection stimulus (visual spatial judgment) or imagining a low-contrast Gabor pattern (contrast judgment) can improve performance on subsequent perceptual tasks (Tartaglia et al., 2009). Similarly, imagining motor-acts facilitates performance on corresponding tasks by training relevant parts of motor cortex, and by strengthening associations between processes and actions (Driskell et al., 1994; Weiss et al., 1994; Feltz and Landers, 2007). To date, research has mainly focused on the effects imagery training has on subsequent perceptual tasks. Here, we look directly at the influence of imagery training on the strength of imagery itself.

One of the hallmarks of mental imagery is the considerable difference in reported imagery strength and vividness observed across individuals (Galton, 1883; McKellar, 1965; Marks, 1973; Amedi et al., 2005; Cui et al., 2007). Some individuals claim veridical, vivid imagery, while others doubt its entire existence (McKellar, 1965). The factors causing such differences in imagery strength remain largely unknown. One hypothesis is that individuals who actively practice, or whose everyday activities involve strong use of imagery, might have strengthened their imagery through training and practice (Sacks, 2010). We sought to examine such a proposal in the lab by engaging individuals in an imagery task daily, over a period of 5 days. Can repeated instances of forming visual imagery lead to improved imagery strength?

To address this question researchers must be able to reliably measure imagery strength from 1 day to the next. Previous work demonstrated that sustained imagery has a pronounced and visually specific impact on subsequent perception (Pearson et al., 2008, 2011). These studies utilized a visual phenomenon called binocular rivalry; when two different patterns are presented one to each eye, only one of the patterns is consciously perceived. Subtle experimental manipulations, such as attention (Meng and Tong, 2004; Mitchell et al., 2004; Chong and Blake, 2006; Kamphuisen et al., 2007), sensory memory (Pearson and Brascamp, 2008), or imagery (Pearson et al., 2008), can bring about a slight imbalance in the neural states, creating a bias that helps one pattern win the race for dominance at the expense of the other.

We have previously demonstrated that imagery can alter future competitive visual interactions in favor of the imagined stimulus on a large percentage of trials (Pearson et al., 2008, 2011), while catch-trial presentations of mock rivalry stimuli do not reveal such bias, ruling out the possibility of demand characteristics (Pylyshyn, 2003). Indeed, scores on offline imagery questionnaires predict imagery strength measured using rivalry (Pearson et al., 2011) and rivalry has been utilized to examine the role of imagery during visual working memory (Keogh and Pearson, 2011). Thus, there is compelling evidence that rivalry bias (or "perceptual bias") is a useful way to measure imagery strength in general (e.g., encompassing perceptual elements and sensations of vividness). In the current study, imagery strength is the underlying construct of interest, and the extent to which imagery biases perception is taken as a reliable measure of imagery strength. The subjective experiences associated with imagery strength are probed by having participants report the "vividness" of their mental images.

Can people evaluate the phenomenal qualities of internally generated experiences, such as whether a mental image is vivid or detailed? Recently, an attempt was made to answer the question of knowing ones own thoughts (exemplifying the problem of "metacognition;" Flavell, 1979) in relation to mental imagery (Pearson et al., 2011). This study provided compelling new evidence that people have accurate metacognitive knowledge at fine-grained scale, regarding specific instances of imagery: On individual trials, higher ratings of imagery vividness predicted a greater likelihood that the imagined pattern would appear dominant during subsequent rivalry (Pearson et al., 2011). Interestingly, repeated attempts to form a particular visual image can lead to different degrees of success with each try, causing imagery strength to fluctuate from one moment to the next. Despite this variance in imagery strength, people demonstrate good metacognitive understanding of their imagery, and can readily evaluate how vivid their mental images are on a particular occasion.

At a general level, there has been a growing interest in metacognitive judgments of memory and sensory decision-making (Kiani and Shadlen, 2009; Fleming et al., 2010; Rounis et al., 2010; Song et al., 2011). Frontal brain regions are important for introspective or metacognitive ability (Kepecs et al., 2008; Fleming et al., 2010), which suggests that the neural substrates of metacognitive ability are distinct from those supporting primary perception. Although the ability to introspect varies substantially across individuals, within a single individual metacognitive ability seems to be stable and task independent, suggesting a common cognitive process (Song et al., 2011).

Little is known regarding the stability and independence of metacognition of mental imagery. If metacognition for perceptual tasks originates from a common cognitive process, might a similar process allow people to have metacognition of mental imagery? Despite the highly subjective and volitional nature of imagery, people are reasonably good at imagery metacognition (Pearson et al., 2011). Is this ability stable, or might metacognition of imagery improve with repeated practice? Here, we also investigated the degree of imagery metacognition as a function of daily training.

To assess metacognition we use a method derived from signal detection theory (Swets, 1986; Macmillan and Creelman, 1991; Galvin et al., 2003; Kornbrot, 2006) that has been successfully employed in a variety of recent metacognition studies (Fleming et al., 2010; Song et al., 2011). Using this method, we looked at the likelihood that imagery biased subsequent rivalry, given a certain level of imagery vividness. Signal detection allows us to estimate a single quantitative "sensitivity" measure of metacognitive ability, derived from these objective (amount of perceptual bias) and subjective (ratings of vividness) variables. This measure of sensitivity is criterion free, which means that it is not prone to changes in criterion (rating-magnitude), and it is not affected by irregular use of the rating scale (which generally results in unequal numbers of observations across the various conditions).

By way of preview, here we report that imagery strength – measured as the extent to which imagery biases perception during binocular rivalry – did not increase over the 5-day training period. Interestingly, participant's metacognition of imagery did significantly improve over the training period. This dissociation between imagery strength and metacognitive ability suggests a degree of independence between the two processes.

MATERIALS AND METHODS

PARTICIPANTS

Nine observers (six female) participated in the experiment. All had normal or corrected-to-normal visual acuity and normal stereovision, and all provided written informed consent. Observers received payment for their participation ($10 per hour, plus a $5 per hour bonus upon completion) with the exception of a participating author (RR) and participant BW. The study was carried out with the approval of the Institutional Review Board at Vanderbilt University.

MATERIALS

Observers viewed the stimuli on a luminance-calibrated CRT monitor with 1152 × 870 resolution and a 75-Hz refresh rate in an otherwise darkened room. Visual stimuli were generated with Matlab 7.5.0 (R2007b) and the Psychophysics toolbox (Brainard, 1997; Pelli, 1997) under Mac OSX. Observers sat at a viewing distance of 56 cm, and used a chinrest to maintain a stable head position. A mirror stereoscope was used to present a different pattern to each eye, and binocular convergence of the two images was aided by a white bull's eye fixation dot (0.95°) at the center of each monocular half-image. Participants were instructed to maintain steady fixation throughout all experimental trials.

Rivalry stimuli consisted of a green and a red grating (spatial frequency = 1.23 c/°) surrounding a central fixation point, presented against a black background with a mean luminance of 0.09 cd/m^2. CIE color values of the stimuli were as follows – green: $x = 0.293$, $y = 0.572$; red: $x = 0.602$, $y = 0.353$. Gratings were presented at 75% contrast and had a Gaussian-shaped luminance profile (mean luminance = 6.95 cd/m^2) that faded to black at the stimulus edge (Gaussian $\sigma = 4.29°$). Five observers were trained with a green grating of orientation 112.5° and a red grating of orientation 22.5°, while on generalization blocks they were presented with 67.5° green, and 157.5° red gratings. The opposite was true for the remaining four observers, meaning that we counterbalanced which grating-pairs were used for training and generalization between participants. On catch trials, a mock rivalry stimulus was presented consisting of a physical blend of the green and red rivalry patterns. This stimulus was presented to both eyes simultaneously in order to avoid interocular competition. Presentation of the mock-stimulus allowed us to test for decisional bias and demand characteristics (Landsberger, 1958).

The dominant eye plays a key role in determining which of two monocular images is likely to be perceived at the onset of binocular rivalry. Therefore, individual fine-tuning of stimulus contrast was done before the start of the experiment, and before each daily session, to control for differences in ocular dominance between observers. We used the same procedure as in previous research (Pearson et al., 2008, 2011; Keogh and Pearson, 2011), matching the relative strength of the rivalry gratings to the point at which perceptual competition is most balanced, and thus most susceptible to disruption.

PROCEDURE

To investigate whether visual imagery can be improved by means of training, and to see how this relates to metacognition of imagery over time, we had observers perform a visual imagery task on five consecutive days, for about an hour a day. A sixth follow-up session was conducted 2–3 weeks after training. Participants came into the lab at or around the same time on each day of training, and were dark-adapted for a couple of minutes before the start of each experimental session.

During the experiment, participants were briefly presented with a randomly chosen (equal number of both) central cue ("G" for green, or "R" for red) at the beginning of each imagery-trial (**Figure 1A**). Subsequently, participants would engage in visual imagery of the cued pattern for an 8-s period. After completing this imagery period, the word "vividness?" cued participants to first report the quality of their imagery by means of left-handed button presses (1 = *almost no imagery*, 2 = *some weak imagery*, 3 = *moderate imagery*, 4 = *strong imagery almost like perception*), after which they were cued by the word "effort?" to report the amount of vigor with which they had tried to imagine the pattern (1 = *almost no effort*, 2 = *some effort*, 3 = *moderate effort*, 4 = *tried very hard to form a mental image*). Observers were instructed to use the full range of the rating scale to the best of their abilities.

As soon as a participant had responded to both questions, a rivalry display (90% of trials) or a mock display (10% of trials) was presented for 750 ms. On rivalry trials, the green grating was presented to the left eye, and the red grating to the right eye. On mock trials, the plaid-stimulus was presented to both eyes simultaneously. Participants reported which image had appeared most dominant, by pressing one of three buttons (1 = *green*, 2 = *mixed*, 3 = *red*). For this response, the right hand was used in order to minimize potential response conflict between the two hands. A "mixed" response could be made on all trials (rivalry and mock trials). On rivalry trials, the observer could give a mixed response in case he or she was unable to distinguish which grating had appeared more dominant due to binocular combination or piecemeal rivalry. This type of mixed percept was reported on 6.49% of rivalry trials (SEM = 2.49%).

A single training session consisted of two blocks of 70 trials each. Within each block, seven catch trials were randomly interleaved between the rivalry trials. We tested potential generalization of learning to non-trained orientations on day 1 and 5 of training, and during follow-up. On these days, observers performed

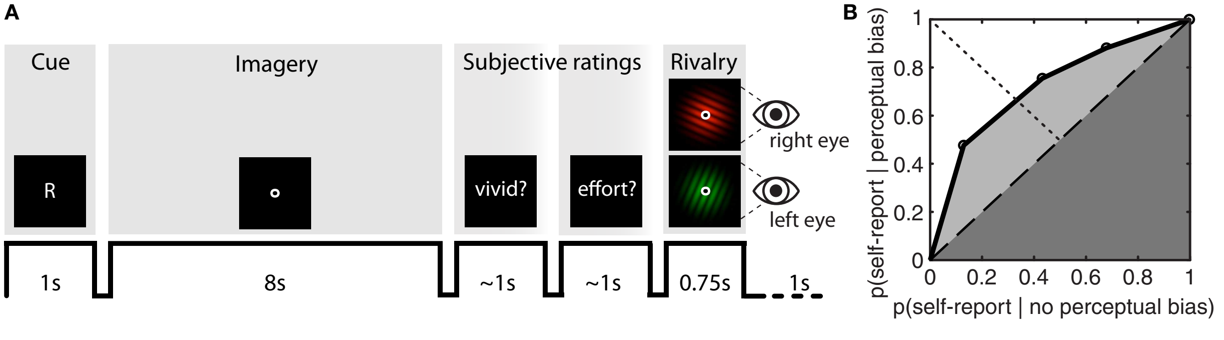


FIGURE 1 | Sample trial sequence and example ROC II curve. (A) Participants were presented with a random cue ("G" or "R") after which they formed a mental image of the green or red grating over an 8-s period. Participants were then cued to report the vividness of the imagined item, and the effort exerted while imagining the item, on an absolute scale from 1 to 4. After a brief flash of the rivalry display, participants reported which grating had appeared perceptually dominant, or whether their percept was an equal mix of the two. On 10% of the trials – instead of the rivalry display – a mock-stimulus was presented to both eyes simultaneously, consisting of a physical combination of both the green and red grating. **(B)** To determine how well subjective ratings predict perceptual bias, type II ROC sensitivity was calculated by taking the area under the ROC curve (A_{roc}). This area is the sum of the area of the half-square triangle (dark-gray shaded region) and the area between the diagonal and the ROC function (light-gray shaded region).

twice the amount of trials, with training and generalization blocks presented separately and in a randomized order.

ANALYSES

To assess the strength of visual imagery, we looked at the perceptual facilitation (or bias) of imagery on rivalry. This was calculated as the percentage of trials in which the imagined grating matched subsequent perception during rivalry (Pearson et al., 2008), excluding trials on which a mixed percept was reported. A perceptual bias greater than 50% (chance) on the rivalry trials but not on the catch trials suggests facilitation due to imagery content. Due to experimenter error, a small number of runs (7 out of 108) were missing from the data. Where necessary, we used trilinear interpolation to infer the mean percentage of bias. For the day-by-day analysis (**Figure 2**) only one data point was interpolated (percentage perceptual bias for participant CB on day 4); the session-by-session analysis of the same data required interpolation of all seven missing runs.

Data obtained from subjective ratings of vividness (and effort) were analyzed by constructing type II receiver operating characteristic (ROC) curves for each participant on each day of training. This method of assessing metacognitive ability is derived from signal detection methods (Swets, 1986; Macmillan and Creelman, 1991; Galvin et al., 2003; Kornbrot, 2006) and has been successfully employed in a variety of recent inquiries about metacognition (Fleming et al., 2010; Song et al., 2011). Essentially, the only difference between type I and type II tasks lies in the event that is being discriminated: Type I decisions are about the occurrence of events independent of the observer (so, distinguishing between signal and noise), whereas type II decisions are about whether a decision was correct or incorrect (so, making a metacognitive judgment).

Since it is not possible to be "correct" or "incorrect" about an internally generated image, we adapted the definition of the type II decision to include judgments about the vividness of single episodes of mental imagery and its effect on the perceptual outcome during brief instances of subsequent binocular rivalry. Applying the signal detection logic, we categorized trials where participants reported high vividness and where imagery subsequently biased perception as "hits." Trials where participants reported high vividness but perception was not subsequently biased were categorized as "false alarms." Here, the ROC II characterizes the probability of a participant being perceptually biased during rivalry, given a certain level of self-reported vividness.

To construct the ROC II curves, we calculated $p(\text{rating} = i \mid \text{perceptual bias})$ and $p(\text{rating} = i \mid \text{no perceptual bias})$ for all i, and transformed these into cumulative probabilities before plotting them against each other (anchored at [0,0] and [1,1]). Distribution-free methods were employed to characterize type II ROC sensitivity by calculating the area under the ROC curve (A_{roc}), and type II ROC bias (B_{roc}). These parameters are derived from simple geometry and do not make assumptions about the shape of the distribution (Kornbrot, 2006). The area under the ROC curve (A_{roc}) quantifies the extent to which metacognitive judgments are predictive of perceptual bias during rivalry (**Figure 1B**); a diagonally flat ROC function indicates little predictive value of the metacognitive judgment on the subsequent perceptual outcome during rivalry. The area under the ROC curve is the sum of the area of the half-square triangle (dark-gray shaded region in **Figure 1B**) and the area between the diagonal and the ROC function (light-gray shaded region in **Figure 1B**):

$$A_{\text{roc}} = 0.25 \sum_{i=1}^{4} \left[\left(h_{i+1} - f_i\right)^2 - \left(h_i - f_{i+1}\right)^2 \right] + 0.5 \tag{1}$$

The bias of the ROC II curve (B_{roc}) was defined as the ratio $K_{\text{B}}/K_{\text{A}}$, where K_{B} is the area between the ROC curve and the major diagonal (dashed line in **Figure 1B**) to the right of the minor diagonal (dotted line in **Figure 1B**), and K_{A} is the area between the ROC curve and the major diagonal to the left of the minor diagonal. A neutral bias would give B_{roc} equal to zero, while a negative or positive B_{roc} indicates a bias toward lower or higher ratings respectively.

$$B_{\text{roc}} = \ln \left(\frac{0.25 \sum_{i=1}^{2} \left[\left(h_{i+1} - f_i\right)^2 - \left(h_i - f_{i+1}\right)^2 \right]}{0.25 \sum_{i=3}^{4} \left[\left(h_{i+1} - f_i\right)^2 - \left(h_i - f_{i+1}\right)^2 \right]} \right) \tag{2}$$

Previous research has shown that subjective ratings of vividness – but not effort – are predictive of how much perceptual bias someone experiences (Pearson et al., 2011). To confirm this, we also applied the ROC II methods described above to participant's ratings of exerted effort. This effort-based ROC thus characterizes the probability of a participant being perceptually biased given a certain level of self-reported effort. Finally, to determine whether the ROC II model did a good job accounting for our metacognitive data, we fit a linear regression model:

$$z(h) = \beta_0 + \beta_1 z\left(f\right) + \varepsilon \tag{3}$$

Where z is the inverse of the cumulative normal distribution function. The ROC II model provided a good fit to the self-reported vividness (mean $R^2 = 0.976 \pm 0.004$) and effort data (mean $R^2 = 0.981 \pm 0.007$).

RESULTS

IMAGERY TRAINING

Sustained mental imagery can bias the perception of an ambiguous display, resulting in a reliable measure of imagery strength on a trial-to-trial basis (Pearson et al., 2008, 2011). When people rate their imagery as more vivid, the likelihood that imagery influences perception is larger (Pearson et al., 2011). Thus, if training mental imagery would result in more vivid images, one would expect to see an increase of perceptual bias over time. **Figure 2** shows the mean imagery strength (or "perceptual bias") as a function of days of training and again 2 weeks later. A within-subjects ANOVA revealed that training did not increase the amount of perceptual bias over time [$F_{(5,40)} < 1$].

Mental imagery did bias perception in favor of the imagined grating [$F_{(5,40)} = 8.861$; $p = 0.018$] which is consistent with previous work demonstrating the effect of mental imagery on rivalry (Pearson et al., 2008, 2011). Unsurprisingly – considering the lack of a training effect – gratings of both trained and untrained

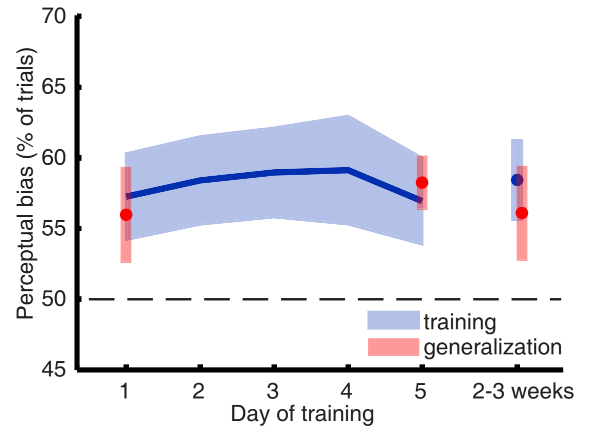


FIGURE 2 | Imagery biased perception during rivalry in favor of the previously imagined grating, but this perceptual bias did not significantly change over the course of a 5-day training. The blue line represents the mean training data, and the blue shaded regions represent ±1 SEM. The parts of the plot depicted in red represent data from blocks where we tested training generalization to other orientations.

(generalization) orientations yielded similar perceptual biases: a within-subjects ANOVA for training days 1, 5, and follow-up revealed no main effect of orientation [$F_{(1,8)} < 1$]. Analyzing the data by session did not unveil any hidden differences in perceptual bias over time [$F_{(11,88)} = 1.106$; $p = 0.366$], which excludes the possibility that most learning took place between the first couple of sessions.

Additional evidence that mental imagery was not improved by training comes from participant's introspective judgments of imagery vividness. Mean self-reported vividness of mental imagery was statistically the same on all days of training [$F_{(5,40)} = 1.224$; $p = 0.316$]. Self-reports of exerted effort did not change over the course of training either [$F_{(5,40)} < 1$]. In summary, neither the perceptual measure of imagery strength ("perceptual bias") nor ratings of vividness showed any significant change over the 5-days of training. Thus, it appears that training in this study was unable to increase imagery strength over time.

CATCH TRIALS

Catch trials were presented in a randomly interleaved fashion on 10% of all experimental trials, to determine whether observers showed response bias in favor of the imagined grating. On these trials, a mock rivalry display was presented consisting of a balanced physical combination of the green and red gratings shown to both eyes simultaneously. If the effects observed during rivalry were due to decisional bias or demand characteristics, we expect to find the same degree of response bias on catch trials. We analyzed bias by coding veridical "mixed" responses to the catch trials as 50%, while responses that matched the cued pattern were coded as 100%, and responses opposite to the cued grating were coded as 0%. The percentage of catch trials during which participant's responses were biased in favor of the cued grating are shown in **Figure 3** (for all days of training). On average, this bias was 50.79%. This indicates that demand characteristics and decisional bias have a negligible influence on participant's reports of rivalry dominance, as previously documented (Pearson et al., 2011).

METACOGNITIVE JUDGMENTS

To assess whether people's metacognitive insights about imagery strength improve over the course of training, we constructed ROC II curves for each individual observer, on each day of training (**Figure 4**; Materials and Methods). The extent to which metacognitive judgments of vividness predict perceptual bias was quantified as the area under each ROC II curve. Data presented in **Figure 4** demonstrate that on earlier days of training (darker green lines) the area under the curve is smaller than on later days of training (lighter green lines). The upward bowing profile of the curves observable in over half of our participants demonstrates that vividness judgments indeed predict perceptual bias.

There is a clear trend toward more metacognitive ability over time for both the trained [$F_{(5,40)} = 1.742$; $p = 0.147$] and untrained (generalization) orientation [$F_{(2,16)} = 7.416$; $p = 0.005$]. Trained and untrained orientations do not statistically differ [$F_{(1,8)} < 1$]. This lack of orientation specificity may not be surprising considering that metacognition for perception is something presumably supported by higher-level frontal areas of the brain (Fleming et al., 2010), where responses are invariant to fine-grained orientation information. Hence, we simplified our analysis by collapsing the data from all orientations before constructing the ROC II curves as displayed in **Figure 4**. Estimates of the type II ROC sensitivity A_{roc} are therefore slightly more reliable on day 1, day 5, and during follow-up, since they are constructed based on more data.

The information from **Figure 4** is summarized in **Figure 5**, showing the main effect of training: vividness judgments predict perceptual bias increasingly better over time [$F_{(5,40)} = 3.075$; $p = 0.019$]. This trend is linear when only looking at training days 1–5 [$F_{(1,8)} = 5.846$; $p = 0.042$] but becomes quadratic when follow-up is included [$F_{(1,8)} = 8.778$; $p = 0.018$], indicating a drop of the proportion A_{roc} at follow-up. Nevertheless, planned comparisons (uncorrected t-tests) show that – with the exception of day 1 – the predictive value of self-reported vividness on the perceptual outcome is larger than would be expected by chance (one-tailed one-sample t-test day 1: $p = 0.243$; all others: $p < 0.029$), and this ability is still present 2–3 weeks after training ($p = 0.021$).

In previous work we demonstrated that people can reliably evaluate the vividness of their mental imagery from one trial to the next (Pearson et al., 2011). Participants in that previous study were tested only once. Thus, one might expect to find evidence for metacognition of imagery strength on day 1 of training in the current experiment. However, the difference in analyses used to determine metacognition in our previous and current work (within-subjects analysis of variance, and area under ROC II curve respectively), make it hard to directly compare the findings. A within-subjects analysis of variance performed on the current data shows that on day 1 of training, participants marginally (but not significantly) showed a main effect of vividness on perceptual bias [$F_{(3,15)} = 2.83$; $p = 0.074$]. However, a lack of observed power (0.558) indicates that at this sample size there is only a small (44%) chance of finding a significant effect (at $\alpha = 0.05$) when assuming that people have metacognitive insights into their own imagery

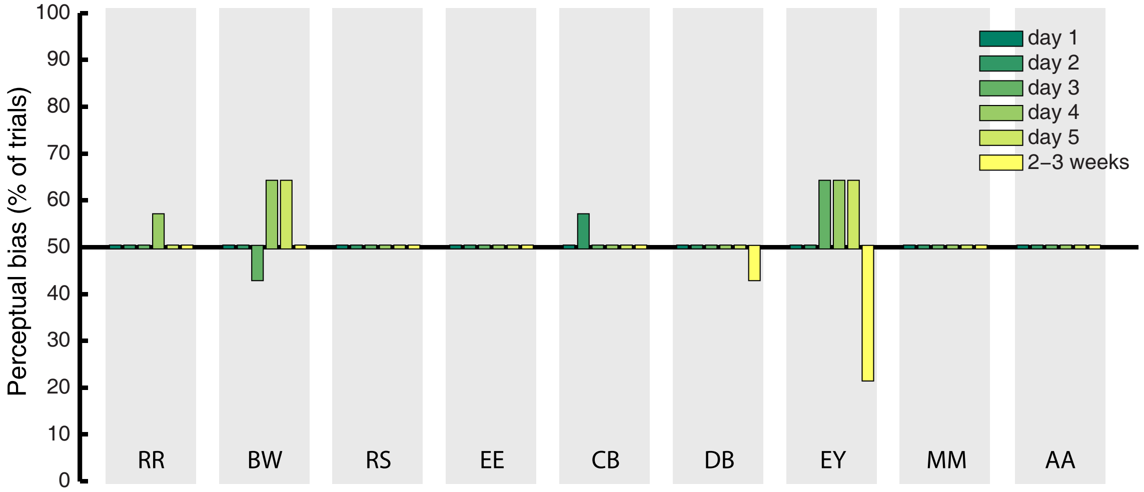


FIGURE 3 | Mean percentage of catch trials on which participant's response to the fake-rivalry display was biased in favor of the imagined grating. A score of 50% indicates a lack of bias. Gray shaded areas show each individual participant, the colored bars represent the 5 days of training plus follow-up 2–3 weeks later.

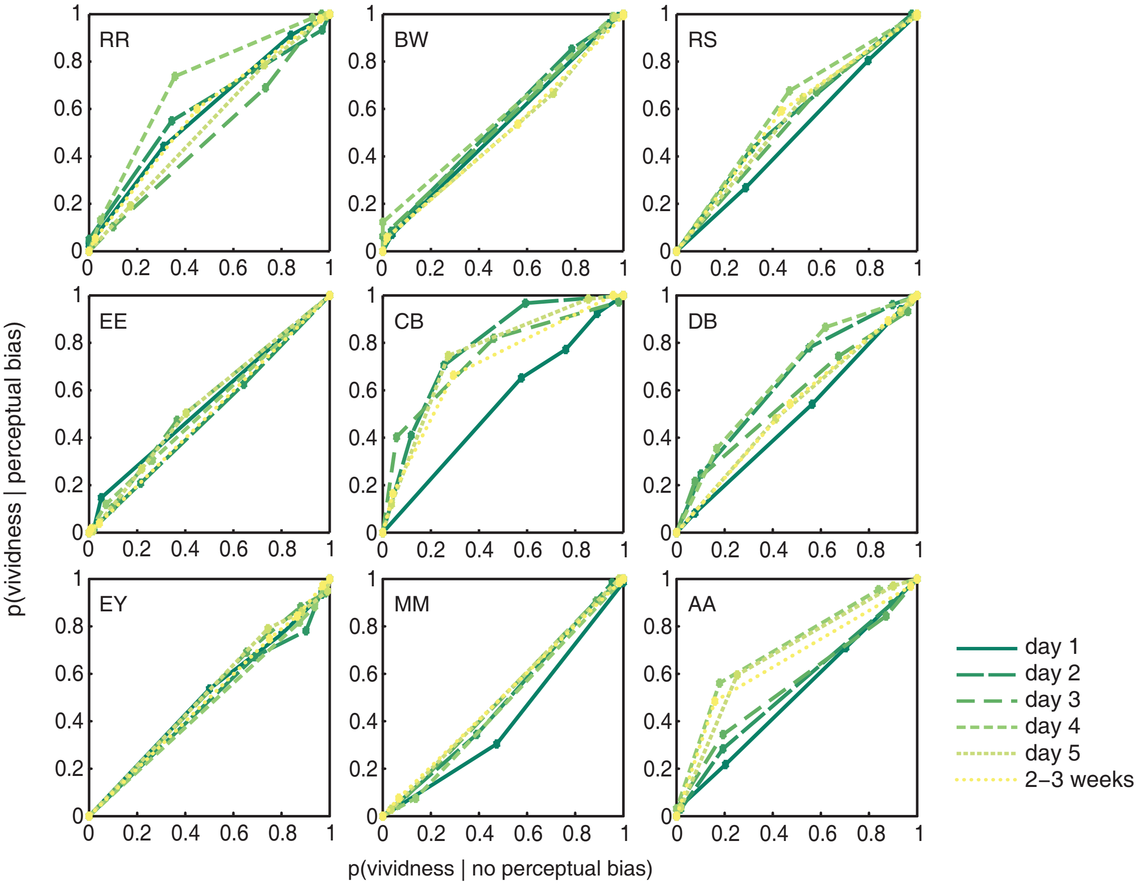


FIGURE 4 | Participant's metacognition of imagery over the course of training. ROC II curves based on the vividness ratings reported by our participants (collapsed over Gabor orientations). Each plot represents an individual participant; the colored lines represent the ROC curves for the different days of training.

strength at the population level. An *a priori* power analysis indicates that, assuming a medium effect size, 21 subjects would be required to obtain a power of 0.95.

Vividness ratings are predictive of the efficacy that mental imagery has at biasing the perception of rivaling stimuli. By contrast, self-reported effort for imagery was not hypothesized to predict perceptual bias. Attempts to exert greater effort do not necessarily result in highly effective imagery, as demonstrated by previous work (Pearson et al., 2011). To ensure that our findings were specific to introspective vividness – and not effort – we constructed ROC II curves (as in **Figure 4**; Materials and Methods) based on the effort ratings reported by our participants. The pooled (across participants) curves per day are shown in **Figure 6A**; the diagonally flat function indicates a weak link between self-reported effort ratings and perceptual bias during rivalry. **Figure 6B** demonstrates that, as expected, effort did not predict perceptual bias [$F_{(1,8)} < 1$]. The area under the ROC II curve (A_{roc}), which quantifies the degree to which self-reported effort predicts perceptual bias during rivalry, did not differ from chance on any of the training days (two-tailed one-sample t-tests all $p > 0.081$). Neither did we observe a change over time for the trained [$F_{(5,40)} < 1$], untrained (generalized) [$F_{(2,16)} = 2.711$; $p = 0.097$], or collapsed [$F_{(5,40)} < 1$] grating orientations.

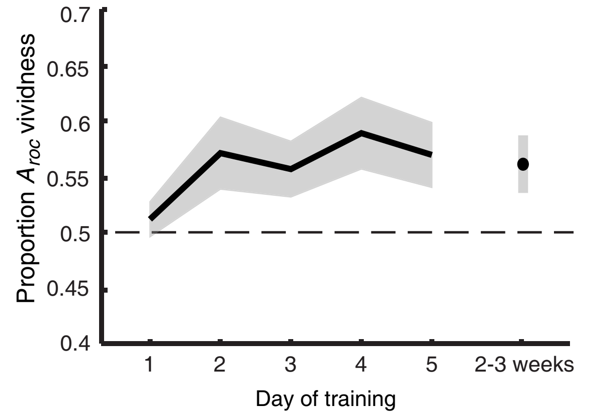


FIGURE 5 | Imagery metacognition improves over the course of training. The extent to which self-reported vividness predicts perceptual bias (quantified as the area under the ROC II curve after collapsing all grating orientations, as also shown in **Figure 4**) plotted against day of training. Shaded areas represent ±1 SEM.

Vividness appears to predict perceptual bias more strongly over the course of training, whereas effort does not predict perceptual bias at all. Can this finding be explained by the way participants used the rating scales? Participant's average reported vividness (2.57 ± 0.21) and effort (2.86 ± 0.14) did not significantly differ ($p = 0.261$). In other words, subjective vividness – but not effort – is predictive of how well something was imagined independent of rating-magnitude. Signal detection theory considers metacognitive ability (sensitivity) and rating-magnitude (bias) as two independent properties (Galvin et al., 2003). In accordance with this notion, we find that individuals with higher self-reported vividness were not better at evaluating their imagery strength and vice versa. Specifically, participant's ability to make accurate metacognitive judgments of their mental imagery (A_{roc} vividness) and participant's mean vividness ratings were not correlated ($r = 0.063$; $p = 0.873$).

The type II bias of the ROC curve (B_{roc}) provides us with a distribution-free estimate of the criterion used by participants to provide their subjective ratings. A neutral bias would give B_{roc} equal to zero, while a negative or positive B_{roc} indicates a bias toward lower or higher ratings respectively. This estimate corresponded very well with the actual rating-magnitude collected during the experiment: self-reported vividness ratings and estimated vividness bias (B_{roc} vividness) were highly correlated ($r = 0.913$; $p < 0.001$), as were self-reported effort and estimated effort bias ($r = 0.889$; $p = 0.001$). The close resemblance between these two variables – both measures of participant's criterion – helps validate the distribution-free approach used to determine ROC estimates in the current paradigm.

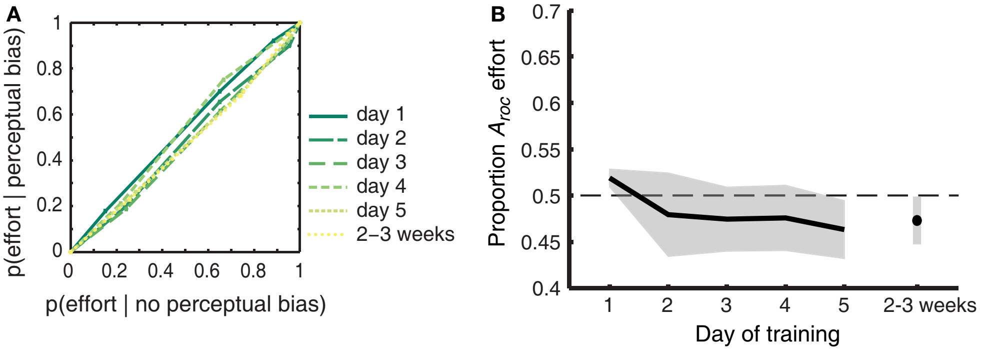


FIGURE 6 | Participant's self-reported effort with which the gratings were imagined does not predict perceptual bias. (A) ROC II curves based on the effort ratings pooled across participants. Green colored lines represent the different days of training. The diagonally flat functions indicate a weak link between self-reported effort ratings and perceptual bias during rivalry. **(B)** The extent to which self-reported effort predicts perceptual bias (quantified as the area under the ROC II curve) plotted against each day of training. Self-reported effort does not predict perceptual bias, and this does not change over time. Data was collapsed over Gabor orientations, since outcomes did not differ between the two sets of orientations [$F_{(1,8)} < 1$]. Shaded areas represent ±1 SEM.

Strikingly, **Figure 4** shows large differences between individuals: the degree to which metacognitive vividness judgments predict perceptual bias varies quite a bit from one person to the next ($A_{\mathrm{roc}} = 0.48–0.71$). This type of variability is not uncommon, and previous studies have reported similarly large individual differences in metacognitive ability for perceptual tasks (Song et al., 2011). Besides large differences related to metacognition, also the overall amount of perceptual bias experienced by our participants varied widely (46–76%). Nonetheless, participant's overall metacognitive ability and the percent perceptual bias they experienced throughout the experiment, were uncorrelated ($r = -0.027$; $p = 0.945$). This suggests that participant's metacognitive ability in this task is independent of imagery strength, as measured by binocular rivalry.

DISCUSSION

The research presented here suggests that mental imagery strength does not improve over the course of our 5-day training regime. Over the 5-days, no changes were observed relating to imagery strength as measured by rivalry, nor were there any changes in the average introspective judgments of imagery vividness. We further demonstrated that self-reported vividness of mental imagery predicts the perceptual consequences of single epochs of imagery. More importantly, this prediction becomes stronger with practice, implying increased metacognition of imagery over the course of training. Self-reported effort of mental imagery on the other hand, did not predict perceptual outcomes.

There have been reports of visual imagery increasing performance on subsequent perceptual tasks (Tartaglia et al., 2009). Yet we were unable to find an increase in facilitation of rivalry dominance after 5 days of training. The question is of course, why? The emphasis of the research presented here was on improving imagery strength over time. This is a notably different emphasis from studies that have investigated how imagery training changes perceptual skills (Tartaglia et al., 2009). One obvious explanation for the lack of an imagery training effect in this study is that imagery strength simply cannot improve with practice. This idea is corroborated by the fact that neither imagery bias, nor subjective ratings of imagery strength showed a significant increase as a function of training. Introspective ratings of imagery strength are reflected in the perceptual outcomes during rivalry, and the close relationship between the two implies they measure the same underlying construct (Pearson et al., 2011). Thus, several aspects of the data support the hypothesis that it is not possible – or very difficult – to improve imagery strength by means of training.

The idea that training cannot easily change imagery strength might be explained by the manner in which imagery strength is linked to brain anatomy. The Tartaglia study (Tartaglia et al., 2009) had participants repetitively imagine the crucial part of a bisection stimulus (spatial judgment) or a low-contrast Gabor pattern (contrast judgment). They found improved perceptual performance on a subsequent perceptual bisection task and a Gabor detection task after imagery training, and this improvement generalized to untrained orientations. This lack of orientation specificity implies that learning through imagery did not involve plastic changes in early visual cortex, but probably involved higher-level extra-striate areas. Higher-level changes may boost perceptual performance through imagery training, yet, changes at this cortical level may not be sufficient to improve imagery strength itself.

Historically, mental imagery has been considered a fainter form of perception (Hume, 1739). Evidence to support this notion comes from functional magnetic resonance imaging (fMRI) studies demonstrating that the magnitude of brain activity is lower during imagery than during bottom-up perception (Goebel et al., 1998; O'Craven and Kanwisher, 2000). Likewise, single neuron recordings in the medial temporal lobe of humans found fewer neurons that were recruited during imagery than during perception, and that the firing rate of these cells was lower during imagery compared to perception (Kreiman et al., 2000). In the case of perceptual bias during rivalry, imagery is presumed to influence or boost the memory trace that exists between one rivalry presentation and the next, and the location and orientation specificity of this memory trace implies that it is composed of primarily low-level characteristics (Ishai and Sagi, 1995; Pearson et al., 2008; Slotnick, 2008). Mechanisms such as a gain in sensitivity for the imagined pattern, or the strengthening of sensory traces, would be needed to modify population activity in lower visual areas mediating alternations of conscious perception during rivalry. Imagery may simply lack sufficient impact to induce permanent plastic changes at these lowest sensory levels. Future research directions aiming to improve imagery strength could investigate the necessity of bottom-up information for learning. Specifically, it would be interesting to see if there is a transfer from improving visual perception by means of prolonged training with actual sensory stimuli, to improvements of imagery strength.

One could hypothesize that imagery strength is liable to improvement, but we simply failed to find any in this study due to the configuration of our task. Research into the process of improving perceptual skill – or perceptual learning – provides useful context in support of this hypothesis. One influential view known as the reverse hierarchy theory (Ahissar and Hochstein, 2004), states that learning is gated by top-down, task-related factors: Learning begins at high-level areas of the brain, after which it trickles down the hierarchy, fine-tuning the read out from lower level areas. This theory invokes a number of detailed predictions, namely, early (fast) learning should be related to high-level changes, whereas asymptotic (slow) learning should involve plasticity in low-level sensory areas – if required by the task. There is considerable evidence supporting this view (Ahissar and Hochstein, 1993, 1997; Dosher and Lu, 1998; Dupuis-Roy and Gosselin, 2007).

In light of the reverse hierarchy framework, our training regime is suspect to a critical vulnerability. Namely: training duration. Five days may have been insufficient time to reach the asymptotic learning phase. The Tartaglia study previously mentioned (Tartaglia et al., 2009) trained participant's imagery for 10 days, twice as long as in our study, and found an improvement on perceptual tasks. Assuming that specific cellular plastic processes at the hierarchical level of ocular dominance columns can only occur during asymptotic learning, longer training might be necessary when aiming to influence rivalry perception.

Recent research has demonstrated that perceptual learning can also occur without a specific task and outside of awareness, as long as the information of interest is paired with feedback or a reward signal (Seitz and Watanabe, 2003, 2005; Seitz et al., 2009) or with online-feedback via decoded fMRI signals (Shibata et al., 2011). Our experimental design lacked a direct reward signal. Perhaps if successful epochs of imagery were paired with a reward signal, this could facilitate learning. In practice the implementation of a reward may prove difficult to realize. Often, measures of imagery strength are dependent on subjective reports, and offering rewards based on only self-reports could induce strong response and observer biases. Nevertheless, it is possible that our training was insufficient to obtain an effect, and providing feedback, rewards, or some manner of getting participants to intentionally try and increase their imagery strength, could have been a more effective way to train mental imagery.

During memory consolidation, initially fragile memory traces become stabilized due to practice-induced plasticity in task relevant brain areas (Karni, 1996; Dudai, 2004). Can the ineffectiveness of imagery training be due to somehow disrupted memory consolidation? Classically, consolidation has been defined as a time limited process directly following learning (Dudai, 2004). However, recent studies indicate that interference is rather time independent, and can occur at long intervals after training (Goedert and Willingham, 2002; Caithness et al., 2004; Zhang et al., 2008). Interference can be considered strongly stimulus dependent, resulting from similarity between the learned and interfering stimulus, and the corresponding neuronal populations recruited by these stimuli (Seitz et al., 2005; Been et al., 2011). Specifically, for Gabor patterns most interference occurs when interfering stimuli differ from the learned orientation by 30°, while no interference is observed from orthogonal orientations (Been et al., 2011). Considering the orthogonal training orientations of our experiment, disruption of consolidation seems an unlikely explanation for the ineffectiveness of imagery training.

Can people become better at knowing their own thoughts? We were able to improve subject's ability to judge the vividness of their imagery. This improvement was still present during a follow-up test, implying a long lasting effect of training on metacognitive evaluation of mental imagery. Furthermore, training of metacognition was not orientation specific: metacognition was improved for both trained and untrained sets of orientations. It is likely that the improvement of metacognition reported here originates from higher-level brain areas. This is in concordance with the suspected high-level neural locus of metacognitive ability for perception (Kepecs et al., 2008; Kiani and Shadlen, 2009; Fleming et al., 2010) as well as the idea that networks in high-level cortical regions orchestrate strategic choices during early learning, allocating attention and motivation in response to specific task demands (Willingham, 1999; Hochstein and Ahissar, 2002; Doyon et al., 2003).

Our study demonstrated improvements of the metacognition of imagery, whereas no changes in imagery strength itself were observed. This dissociation suggests distinct brain mechanisms underlying metacognition and visual imagery respectively. Similar distinctions have been made regarding metacognition of perception: Neuroanatomical substrates of introspective ability are distinct from those supporting primary perception (Fleming et al., 2010), and there is a marked dissociation between metacognitive ability and performance on visual perceptual tasks (Lau and Passingham, 2006; Lau, 2008; Rounis et al., 2010; Song et al., 2011). Thus, metacognitive ability can be viewed as a stable and task independent cognitive process that can be improved with practice, independent of performance on other tasks. Changes in high-level neuronal populations are likely candidates for this learning.

The ability to introspect on private thoughts is key to human subjective experience. Yet, people's ability to evaluate internally generated experiences – such as imagery – is not as self-evident as it may appear. Although a large number of studies now demonstrate that something as private as a mental image can be successfully studied from a third-person perspective (Ishai and Sagi, 1995, 1997; Kosslyn et al., 2001; Pearson et al., 2008; Tartaglia et al., 2009), research has only recently begun to tackle issues related to the first-person perspective (Pearson et al., 2011). The core problem from the first-person perspective of the imaginer is that self-generated instances of imagery, unlike perception, cannot be directly compared with a perceptual template. Nevertheless, people seem quite capable of knowing if a mental image is accurate, vivid, or detailed. And practice further improves this first-person introspective ability. Why might such metacognitive knowledge be important?

Introspective or "metacognitive" sensitivity is important to guide actions and to make decisions (Vickers, 1979; Daw et al., 2005; Dayan and Daw, 2008) and being able to adequately estimate ones confidence can help drive adaptive behavior (Kepecs et al., 2008). In its simplest form, low confidence that a recent decision was correct may prompt reexamination of the evidence, or seeking a second-opinion. In the event of internally generated experiences such as mental imagery, low confidence that an image was veridical and life-like may lead someone to reconsider such an experience. A better metacognitive understanding may help the imaginer bridge the gap between first and third-person perspective. For example, people can resolve potential ambiguities about perception by comparing their own perceptual experience with the subjective experience of another person (Bahrami et al., 2010). Similarly, when the imaginer has a better understanding of the authenticity of his or her mental image, it will be easier to communicate its content to another person. In sum, increasing the efficiency with which people introspect the quality of their mental images can prove a novel and important finding.

In conclusion, we discussed a variety of reasons why training did not lead to an improvement of imagery strength in the current study. Such an improvement may simply be very difficult to document, or our task may not have been optimally suited to detect improvements of imagery strength. Nevertheless, we demonstrated that people's ability to introspect their own imagery strength does improve with training, which suggests distinct mechanisms underlying imagery and metacognition. Being able to improve metacognition by means of practice can have important implications for real-life situations. It would be interesting to know if training metacognition could help people improve certain cognitive functions, such as decision-making or planning actions. If so, this may prove especially helpful for specific patient populations. Finally, future investigations of prolonged training of imagery can prove advantageous in outlining the overlap between mechanisms of perception and imagery. Imagery as defined here is a highly voluntary process that allows introspection in the absence of direct perceptual input. As such, imagery can provide a unique gateway to understanding how perceptual and introspective processes are represented in the brain.

ACKNOWLEDGMENTS

We thank Frank Tong for discussions and helpful input with design and execution of this work. We also thank Jan Brascamp, Elias Cohen, Janneke Jehee, Jan Schepers, and Jascha Swisher for helpful comments and discussions regarding this research. This work was funded by grants from the National Health and Medical Research Council of Australia CJ Martin Fellowship 457146 and project grant APP1024800 to Joel Pearson.

REFERENCES

Ahissar, M., and Hochstein, S. (1993). Attentional control of early perceptual learning. *Proc. Natl. Acad. Sci. U.S.A.* 90, 5718–5722.

Ahissar, M., and Hochstein, S. (1997). Task difficulty and the specificity of perceptual learning. *Nature* 387, 401–406.

Ahissar, M., and Hochstein, S. (2004). The reverse hierarchy theory of visual perceptual learning. *Trends Cogn. Sci. (Regul. Ed.)* 8, 457–464.

Amedi, A., Malach, R., and Pascual-Leone, A. (2005). Negative BOLD differentiates visual imagery and perception. *Neuron* 48, 859–872.

Bahrami, B., Olsen, K., Latham, P. E., Roepstorff, A., Rees, G., and Frith, C. D. (2010). Optimally interacting minds. *Science* 329, 1081–1085.

Been, M., Jans, B., and De Weerd, P. (2011). Time-limited consolidation and task interference: no direct link. *J. Neurosci.* 31, 14944–14951.

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

Caithness, G., Osu, R., Bays, P., Chase, H., Klassen, J., Kawato, M., Wolpert, D. M., and Flanagan, J. R. (2004). Failure to consolidate the consolidation theory of learning for sensorimotor adaptation tasks. *J. Neurosci.* 24, 8662–8671.

Chen, W., Kato, T., Zhu, X. H., Ogawa, S., Tank, D. W., and Ugurbil, K. (1998). Human primary visual cortex and lateral geniculate nucleus activation during visual imagery. *Neuroreport* 9, 3669–3674.

Chong, S., and Blake, R. (2006). Exogenous attention and endogenous attention influence initial dominance in binocular rivalry. *Vision Res.* 46, 1794–1803.

Craver-Lemley, C., and Reeves, A. (1992). How visual imagery interferes with vision. *Psychol. Rev.* 99, 633–649.

Cui, X., Jeter, C. B., Yang, D., Montague, P. R., and Eagleman, D. M. (2007). Vividness of mental imagery: individual variability can be measured objectively. *Vision Res.* 47, 474–478.

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.

Dayan, P., and Daw, N. D. (2008). Decision theory, reinforcement learning, and the brain. *Cogn. Affect. Behav. Neurosci.* 8, 429–453.

Dosher, B. A., and Lu, Z. L. (1998). Perceptual learning reflects external noise filtering and internal noise reduction through channel reweighting. *Proc. Natl. Acad. Sci. U.S.A.* 95, 13988–13993.

Doyon, J., Penhune, V., and Ungerleider, L. G. (2003). Distinct contribution of the cortico-striatal and cortico-cerebellar systems to motor skill learning. *Neuropsychologia* 41, 252–262.

Driskell, J., Copper, C., and Moran, A. (1994). Does mental practice enhance performance? *J. of Appl. Psychol.* 79, 481–491.

Dudai, Y. (2004). The neurobiology of consolidations, or, how stable is the engram? *Annu. Rev. Psychol.* 55, 51–86.

Dupuis-Roy, N., and Gosselin, F. (2007). Perceptual learning without signal. *Vision Res.* 47, 349–356.

Fahle, M., and Poggio, T. (2002). *Perceptual Learning.* Cambridge, MA: MIT Press.

Feltz, D. L., and Landers, D. M. (2007). *The Effects of Mental Practice on Motor Skill Learning and Performance: A Meta-analysis.* Champaign, IL: Human Kinetics.

Fine, I., and Jacobs, R. A. (2002). Comparing perceptual learning tasks: a review. *J. Vis.* 2, 190–203.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: a new area of cognitive-developmental inquiry. *Am. Psychol.* 34, 906–911.

Fleming, S. M., Weil, R. S., Nagy, Z., Dolan, R. J., and Rees, G. (2010). Relating introspective accuracy to individual differences in brain structure. *Science* 329, 1541–1543.

Galton, F. (1883). *Inquiries into Human Faculty and its Development.* London: Macmillan.

Galvin, S. J., Podd, J. V., Drga, V., and Whitmore, J. (2003). Type 2 tasks in the theory of signal detectability: discrimination between correct and incorrect decisions. *Psychon. Bull. Rev.* 10, 843–876.

Gilden, D., Blake, R., and Hurst, G. (1995). Neural adaptation of imaginary visual motion. *Cogn. Psychol.* 28, 1–16.

Goebel, R., Khorram-Sefat, D., Muckli, L., Hacker, H., and Singer, W. (1998). The constructive nature of vision: direct evidence from functional magnetic resonance imaging studies of apparent motion and motion imagery. *Eur. J. Neurosci.* 10, 1563–1573.

Goedert, K. M., and Willingham, D. B. (2002). Patterns of interference in sequence learning and prism adaptation inconsistent with the consolidation hypothesis. *Learn. Mem.* 9, 279–292.

Harrison, S., and Tong, F. (2009). Decoding reveals the contents of visual working memory in early visual areas. *Nature* 458, 632–635.

Hochstein, S., and Ahissar, M. (2002). View from the top: hierarchies and reverse hierarchies in the visual system. *Neuron* 36, 791–804.

Hume, D. (1739). *A Treatise on Human Nature*, Vol. 1. London: John Noon.

Ishai, A., and Sagi, D. (1995). Common mechanisms of visual imagery and perception. *Science* 268, 1772–1774.

Ishai, A., and Sagi, D. (1997). Visual imagery: effects of short-and long-term memory. *J. Cogn. Neurosci.* 9, 734–742.

Kamphuisen, A., Van Wezel, R., and Van Ee, R. (2007). Inter-ocular transfer of stimulus cueing in dominance selection at the onset of binocular rivalry. *Vision Res.* 47, 1142–1144.

Karni, A. (1996). The acquisition of perceptual and motor skills: a memory system in the adult human cortex. *Brain Res. Cogn. Brain Res.* 5, 39–48.

Keogh, R., and Pearson, J. (2011). Mental imagery and visual working memory. *PLoS ONE* 6, e29221. doi:10.1371/journal.pone.0029221

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

Kiani, R., and Shadlen, M. N. (2009). Representation of confidence associated with a decision by neurons in the parietal cortex. *Science* 324, 759–764.

Kornbrot, D. E. (2006). Signal detection theory, the approach of choice: model-based and distribution-free measures and evaluation. *Percept. Psychophys.* 68, 393–414.

Kosslyn, S. M., Ganis, G., and Thompson, W. L. (2001). Neural foundations of imagery. *Nat. Rev. Neurosci.* 2, 635–642.

Kosslyn, S. M., Pascual-Leone, A., Felician, O., Camposano, S., Keenan, J. P., Thompson, W. L., Ganis, G., Sukel, K. E., and Alpert, N. M. (1999). The role of area 17 in visual imagery: convergent evidence from PET and rTMS. *Science* 284, 167–170.

Kosslyn, S. M., Thompson, W. L., Kim, I. J., and Alpert, N. M. (1995). Topographical representations of mental images in primary visual cortex. *Nature* 378, 496–498.

Kreiman, G., Koch, C., and Fried, I. (2000). Imagery neurons in the human brain. *Nature* 408, 357–361.

Landsberger, H. A. (1958). *Hawthorne Revisited.* Ithaca, NY: Cornell University.

Lau, H. C. (2008). A higher order Bayesian decision theory of consciousness. *Prog. Brain Res.* 168, 35–48.

Lau, H. C., and Passingham, R. E. (2006). Relative blindsight in normal observers and the neural correlate of visual consciousness. *Proc. Natl. Acad. Sci. U.S.A.* 103, 18763–18768.

Macmillan, N. A., and Creelman, C. D. (1991). *Detection Theory: A User's Guide.* Cambridge: Cambridge University Press.

Marks, D. F. (1973). Visual imagery differences in the recall of pictures. *Br. J. Psychol.* 64, 17–24.

McDermott, K. B., and Roediger, H. L. (1994). Effects of imagery on perceptual implicit memory tests. *J. Exp. Psychol. Learn Mem. Cogn.* 20, 1379–1390.

McKellar, P. (1965). *The Investigation of Mental Images*. Harmondsworth: Penguin Books.

Meng, M., and Tong, F. (2004). Can attention selectively bias bistable perception? Differences between binocular rivalry and ambiguous figures. *J. Vis.* 4, 539–551.

Mitchell, J., Stoner, G., and Reynolds, J. (2004). Object-based attention determines dominance in binocular rivalry. *Nature* 429, 410–413.

O'Craven, K., and Kanwisher, N. (2000). Mental imagery of faces and places activates corresponding stimulus-specific brain regions. *J. Cogn. Neurosci.* 12, 1013–1023.

Pearson, J., and Brascamp, J. (2008). Sensory memory for ambiguous vision. *Trends Cogn. Sci. (Regul. Ed.)* 12, 334–341.

Pearson, J., Clifford, C., and Tong, F. (2008). The functional impact of mental imagery on conscious perception. *Curr. Biol.* 18, 982–986.

Pearson, J., Rademaker, R. L., and Tong, F. (2011). Evaluating the mind's eye: the metacognition of visual imagery. *Psychol. Sci.* 22, 1535–1542.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat. Vis.* 10, 437–442.

Perky, C. (1910). An experimental study of imagination?. *Am. J. Psychol.* 21, 422–452.

Pylyshyn, Z. (2003). Return of the mental image: are there really pictures in the brain? *Trends Cogn. Sci. (Regul. Ed.)* 7, 113–118.

Rounis, E., Maniscalco, B., Rothwell, J. C., Passingham, R. E., and Lau, H. (2010). Theta-burst transcranial magnetic stimulation to the prefrontal cortex impairs metacognitive visual awareness. *Cogn. Neurosci.* 1, 165–175.

Sacks, O. (2010). *The Mind's Eye*. New York: Alfred A. Knopf.

Sasaki, Y., Nanez, J. E., and Watanabe, T. (2010). Advances in visual perceptual learning and plasticity. *Nat. Rev. Neurosci.* 11, 53–60.

Seitz, A., and Watanabe, T. (2005). A unified model for perceptual learning. *Trends Cogn. Sci. (Regul. Ed.)* 9, 329–334.

Seitz, A. R., Kim, D., and Watanabe, T. (2009). Rewards evoke learning of unconsciously processed visual stimuli in adult humans. *Neuron* 61, 700–707.

Seitz, A. R., and Watanabe, T. (2003). Psychophysics: is subliminal learning really passive? *Nature* 422, 36.

Seitz, A. R., Yamagishi, N., Werner, B., Goda, N., Kawato, M., and Watanabe, T. (2005). Task-specific disruption of perceptual learning. *Proc. Natl. Acad. Sci. U.S.A.* 102, 14895–14900.

Serences, J. T., Ester, E. F., Vogel, E. K., and Awh, E. (2009). Stimulus-specific delay activity in human primary visual cortex. *Psychol. Sci.* 20, 207–214.

Shibata, K., Watanabe, T., Sasaki, Y., and Kawato, M. (2011). Perceptual learning incepted by decoded fMRI neurofeedback without stimulus presentation. *Science* 334, 1413–1415.

Slotnick, S. D. (2008). Imagery: mental pictures disrupt perceptual rivalry. *Curr. Biol.* 18, R603–R605.

Slotnick, S. D., Thompson, W. L., and Kosslyn, S. M. (2005). Visual mental imagery induces retinotopically organized activation of early visual areas. *Cereb. Cortex* 15, 1570–1583.

Song, C., Kanai, R., Fleming, S. M., Weil, R. S., Schwarzkopf, D. S., and Rees, G. (2011). Relating inter-individual differences in metacognitive performance on different perceptual tasks. *Conscious. Cogn.* 20, 1787–1792.

Stokes, M., Thompson, R., Cusack, R., and Duncan, J. (2009). Top-down activation of shape-specific population codes in visual cortex during mental imagery. *J. Neurosci.* 29, 1565–1572.

Swets, J. A. (1986). Form of empirical ROCs in discrimination and diagnostic tasks: implications for theory and measurement of performance. *Psychol. Bull.* 99, 181–198.

Tartaglia, E. M., Bamert, L., Mast, F. W., and Herzog, M. H. (2009). Human perceptual learning by mental imagery. *Curr. Biol.* 19, 2081–2085.

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

Weiss, T., Hansen, E., Rost, R., Beyer, L., Merten, F., Nichelmann, C., and Zippel, C. (1994). Mental practice of motor skills used in poststroke rehabilitation has own effects on central nervous activation. *Int. J. Neurosci.* 78, 157–166.

Willingham, D. B. (1999). Implicit motor sequence learning is not purely perceptual. *Mem. Cognit.* 27, 561–572.

Zatorre, R. J., and Halpern, A. R. (2005). Mental concerts: musical imagery and auditory cortex. *Neuron* 47, 9–12.

Zhang, J.-Y., Kuai, S.-G., Xiao, L.-Q., Klein, S. A., Levi, D. M., and Yu, C. (2008). Stimulus coding rules for perceptual learning. *PLoS Biol.* 6, e197. doi:10.1371/journal.pbio.0060197

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: 27 April 2012; accepted: 16 June 2012; published online: 10 July 2012.

*Citation: Rademaker RL and Pearson J (2012) Training visual imagery: improvements of metacognition, but not imagery strength. Front. Psychology **3**:224. doi: 10.3389/fpsyg.2012.00224*

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

frontiers in
PSYCHOLOGY

REVIEW ARTICLE
published: 02 October 2012
doi: 10.3389/fpsyg.2012.00360

New percepts via mental imagery?

Fred W. Mast[1,2]*, **Elisa M. Tartaglia**[3,4] **and Michael H. Herzog**[3]

[1] Department of Psychology, University of Bern, Bern, Switzerland
[2] Center for Cognition, Learning and Memory, University of Bern, Bern, Switzerland
[3] Laboratory of Psychophysics, Brain and Mind Institute, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland
[4] Laboratory of Neurophysics and Physiology, UMR 8119 CNRS, Université René Descartes, Paris, France

Edited by:
Joel Pearson, The University of New South Wales, Australia

Reviewed by:
Duje Tadin, University of Rochester, USA
Mark M. Schira, University of New South Wales, Australia

Correspondence:
Fred W. Mast, Department of Psychology, University of Bern, Muesmattstrasse 45, Bern CH-3012, Switzerland.
e-mail: fred.mast@psy.unibe.ch

We are able to extract detailed information from mental images that we were not explicitly aware of during encoding. For example, we can discover a new figure when we rotate a previously seen image in our mind. However, such discoveries are not "really" new but just new "interpretations." In two recent publications, we have shown that mental imagery can lead to perceptual learning (Tartaglia et al., 2009, 2012). Observers imagined the central line of a bisection stimulus for thousands of trials. This training enabled observers to perceive bisection offsets that were invisible before training. Hence, it seems that perceptual learning via mental imagery leads to new percepts. We will argue, however, that these new percepts can occur only within "known" models. In this sense, perceptual learning via mental imagery exceeds new discoveries in mental images. Still, the effects of mental imagery on perceptual learning are limited. Only perception can lead to really new perceptual experience.

Keywords: ambiguous figure, bisection, combinatorics, interpolation, mental rotation, top-down processing, perceptual learning

> *Imagination is more important than knowledge. For knowledge is limited, whereas imagination embraces the entire world, stimulating progress, giving birth to evolution.*
>
> *(Albert Einstein, 1929)*

MENTAL IMAGERY

Mental imagery is usually referred to as seeing with the mind's eye. We are able to create images in our mind despite the absence of appropriate sensory stimulation. Mental images are usually described introspectively, using a sensory terminology such as form, color, spatial extent, and so forth. Even though introspection does not necessarily allow for drawing conclusions about the underlying mechanisms recent research has shown that people have a good metacognitive understanding of the vividness of their mental images (Pearson and Tong, 2011).

The mechanisms that underlie mental imagery have become the topic of the "Imagery Debate" (see for example Kosslyn et al., 2003; Pylyshyn, 2003). Cognitive psychologists have carried out a large amount of behavioral experiments to tap into the mechanisms that underlie mental imagery, and the results provide compelling evidence that mental imagery shares common mechanisms with visual perception. For example, it has been shown that mental images preserve spatial distances. In experiments on image scanning, participants had to first memorize visual information shown in a picture (e.g., a map of an island) and later answer questions about landmarks (Kosslyn et al., 1978). Interestingly, response times were proportional to the actual distances between the landmarks. It took participants more time to scan longer distances. Hence, participants were able to extract metric distance information from memory. Yet other experiments have shown that participants are well able to judge and compare distances in mental images. For example, when participants indicate whether a previously seen object is higher than wide (e.g., a picture of a sailboat) they are able to extract this information from mental images even though they did not attend to the spatial dimensions when they encoded the visual stimulus (Kosslyn et al., 1995). It has been argued that participants learned the stimuli beforehand, and that tacit knowledge about the purpose of the experiment could have led to the findings. Finke and Pinker (1982) showed participants an array of dots. After the dots disappeared, an arrow was presented and participants decided whether or not the arrow pointed to one of the dots they just saw. Response times increased linearly with increasing distance between the arrow and the target dot. It has been proposed that attentional crowding could account for the distance effect, leading to more difficult discriminations for further distances (Pylyshyn, 2002). However, this possible explanation has been ruled out by Dror and Kosslyn (1993) who conducted a modified version of Finke and Pinker's experiment. They replaced dots by black (3) and white (17) squares which were arranged in a square-like configuration (six squares on each side). The task remained the same as in Finke and Pinker's experiment but the distance between the black squares was chosen so that it exceeded the distance known to produce attentional crowding (Intriligator and Cavanagh, 2001). As expected, Dror and Kosslyn (1993) demonstrated that response times increased with increasing distance between the arrow and the square. Hence, their results confirmed that mental images embody metric properties even when no spatial relationships are encoded explicitly during the presentation of the stimulus. Yet other evidence comes from neuroimaging studies showing a large overlap in activated areas during mental imagery and visual perception, including early visual cortex (e.g., Slotnick et al., 2005). Activation in the latter has been found when participants were engaged in mental imagery tasks that required the extraction of high-resolution visual information. In addition, TMS over the occipital cortex disrupted performance in those tasks (Kosslyn et al., 1999). Thus, there is compelling evidence for commonalities between mental imagery and perception. However, it has to be pointed out that there are also differences, and, for example, only rarely do we confuse images with percepts. The discovery of new information is yet another conceivable difference between mental imagery and perception. When we inspect with our eyes a visual scene new information is picked up continuously. We discover more and more details that are part of the visual scene. It has been argued that new discoveries cannot be made in mental imagery (Chambers and Reisberg, 1985, 1992). A mental image contains nothing new besides the information the observer is aware of when generating the image. Ambiguous figures, for example, demonstrate compellingly that the perceptual interpretation of the same visual stimulus can switch from one explanation to another. Observers usually do not notice the alternative perceptual explanation when first viewing the ambiguous figure. Obviously, new discoveries occur in perception. What about mental images? Are people able to discover new information in mental images? Mast and Kosslyn (2002) used an ambiguous figure in which the alternative perceptual explanation was upside down (young lady, old lady). It was almost impossible to discover the second explanation (e.g., young lady upside down) accidentally during encoding. However, roughly 50% of participants were able to discover in their mental image a new upright interpretation after having mentally rotated the image by 180°. None of the participants was aware about the alternative explanation when they learned the image (in one orientation only). Hence, observers discovered a new interpretation that they were not aware of when encoding the image. Thus, the ability to extract new interpretations is not bound to the process of perception as it can occur just as well during mental imagery. The ability to discover new interpretations in images is not the crucial difference between imagery and perception. Mental images are not interpreted entities as claimed by Fodor (1975) and they are not tied to their initial interpretation assigned during image generation. Indeed, participants were not aware of the second interpretation in the image even though all the information was available in the stimulus they encoded. Hence, only a new interpretation but no new information was discovered. Here is another example. Imagine a letter "D" and rotate it counterclockwise by 90°, and then take the letter "J" and attach it to the rotated "D" so that the upper end of the "J" is attached exactly at the middle of the horizontal line in the rotated "D." What does it look like? An umbrella. The umbrella does not make part of its individual constituents (D, J) but still people are able to extract the shape from the configuration created in imagery. Indeed, a major strength of mental imagery is the ability to recombine information, thus, going beyond perceptual experience. Therefore, mental imagery serves an important function in creative thinking. However, the detection of emergent new shapes in imagery still remains in the realm of potential interpretations.

This article focuses on the relation between mental imagery and perceptual learning. At first glance, these topics appear somewhat unrelated but we were able to demonstrate in a series of experiments the existence of perceptual learning via mental imagery. Perceptual learning without perceptual input opens new ways of using mental imagery in learning paradigms, thus challenging classical views on perceptual learning. The fact that mental imagery and visual perception share – at least in part – the same mechanisms extends to learning and memory. We gathered evidence that learning in imagery transfers to perceptual performance (Tartaglia et al., 2009, 2012). The next paragraph will provide an overview of perceptual learning, and focus on the implications that learning via mental imagery has on the current understanding of mental imagery and perception alike.

PERCEPTUAL LEARNING

Perceptual learning is learning to perceive. Sommeliers can taste not only the grape of a wine but often the year of its production and vineyard. Years of training make the master. In the laboratory, training improves basic visual skills including the discrimination of motion direction (Ball and Sekuler, 1987), the detection of Gabor gratings (Sowden et al., 2002), stereoacuity (Fendick and Westheimer, 1983; Kumar and Glaser, 1993), line orientation judgments (Vogels and Orban, 1985), texture discrimination (Karni and Sagi, 1991; Pourtois et al., 2008) and the discrimination of fine spatial features, such as in vernier acuity (Fahle and Edelman, 1993). Another example are bisection stimuli where a central line is closer to one of two outer lines (**Figure 1**). Observers indicate the offset direction (to the left or to the right). Training can strongly improve offset discrimination (**Figure 1**).

Without phylogenetic or ontogenetic perceptual learning, there is no perception (Herzog and Esfeld, 2009). Kittens that were raised in an environment missing horizontal lines during a critical period were blind for the horizontal orientation for the rest of their entire life (Blakemore and Cooper, 1970; Hirsh and Spinelli, 1970). In this respect, perceptual learning is often thought to be the most basic type of learning.

The mechanisms that underlie perceptual learning are controversially debated. However, there is clear consensus that perceptual learning is driven by the repeated exposure to the stimuli. Only repeated wine tastings make a good sommelier. On the neural level, the presentation of a stimulus leads to changes of synaptic weights which may, for example, fine tune receptive fields (e.g., Fahle, 2005) or enhance gating to decision stages (e.g., Herzog and Fahle, 1998; Lu and Dosher, 2004). Top-down aspects, such as attention, may be crucial in perceptual learning, but they are only modifying, but not driving the synaptic changes. For example, attention selects aspects of the stimuli to be learned, but the learning itself is driven by the stimuli.

In all mathematical models of perceptual learning, there is a stimulus related term essential for learning. In unsupervised learning algorithms, as for example in Hebbian learning, the synaptic changes Δw_j depend only on the concurrent activation of neurons:

$$\Delta w_j = \eta \cdot y^{\mu} x_j^{\mu}$$

where x_j^{μ} is the activity of the j-th input neuron in response to stimulus μ and y_j^{μ} is the activity of the output neuron in response to the same stimulus (η is the learning rate); y_j^{μ} is the weighted

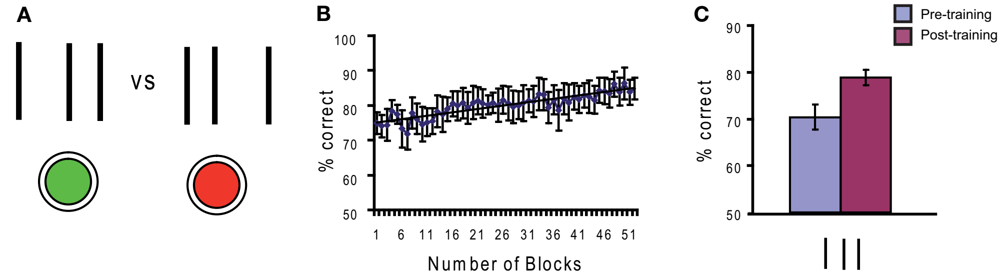


FIGURE 1 | (A) In line bisection, a spatial interval, delineated by two outer lines, is bisected in two unequal components by a central line. The task of the observer is to judge whether the central line is closer to the left or right outer line. Observers are asked to give their response by using one of two buttons. **(B)** Performance improves with training, i.e., observers become able to discriminate offset sizes indiscriminable before training. **(C)** Another way of measuring improvements of performance is to compare observers' performance before ("pre-training") and after ("post-training") the training phase. Post-training baselines are significantly higher than pre-training baselines.

sum of the input $y_j^\mu = \Sigma w_j * x_j^\mu$ and hence fully determined by the actual input and the synaptic weights w_j. If no stimuli are presented $x_j = 0$, for all j, i.e., there is no activation of the input layer and hence no learning since $\Delta w_j = 0$, for all j. Hence, no perceptual learning is expected in the absence of stimulus presentation. Similar considerations hold for other types of neural networks of perceptual learning.

However, contrary to all previous thinking in perceptual learning research, our recent experiments showed that perceptual learning can occur in the absence of (proper) stimulus presentation when participants imagine the missing perceptual information (Tartaglia et al., 2009, 2012). For example, we used a modified bisection stimulus in which only the outer lines were presented during training (**Figure 2**). Observers were asked to imagine the center line to be offset either to the left or right depending on a tone provided in addition (a high tone was associated to the right offset, a low tone to the left one). Hence, the physical stimulus was always the same in all trials. There was "nothing" to learn from the stimulus itself. Still, performance improved during mental imagery, as determined by pre- and post-training measurements in which the proper bisection stimulus was presented (**Figure 2**). Hence, mental imagery is sufficient to enable perceptual learning.

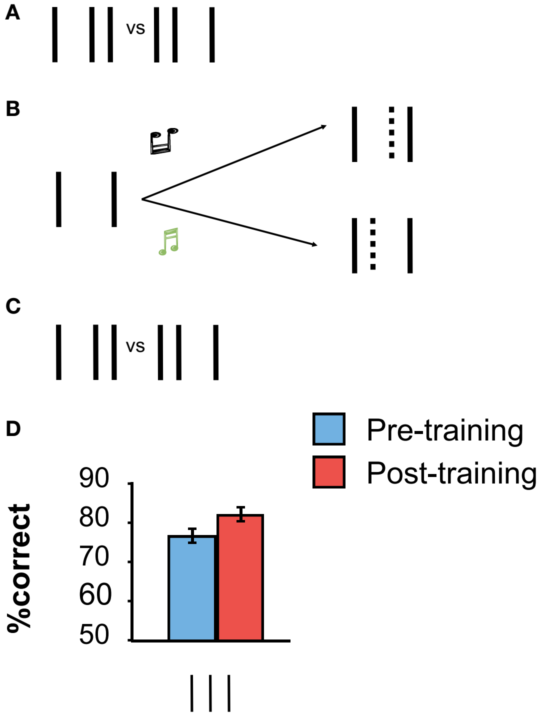


FIGURE 2 | (A,C) Pre- and Post-training baseline measurements with the "proper" vertical bisection stimulus. **(B)** Imagery training. Only the two outer lines of the bisection stimulus are presented. Observers are asked to imagine the absent central line to be offset either to the left or to the right, depending on a tone, and to push buttons accordingly. We asked observers to imagine the smallest possible offset; the imagined line is indicated by the dashed line (which was not shown in the actual display). **(D)** Performance improved significantly through training. The error bars indicate the standard error of the mean (SEM) for nine observers. Adapted from Tartaglia et al. (2009).

The improvements during training are not caused by the pre- and post-measurements with the full bisection stimulus because measurements without training did not yield any performance gains (see Tartaglia et al., 2009). But is it really mental imagery that improves performance in perceptual learning? Unspecific effects such as better coping with attention and improved decision making need to be controlled for. We performed a control experiment with (almost) the same stimuli but without imagery during training. During the training phase, again only the two vertical outer lines of the bisection were presented, together with the two different tones. Observers had to press the right button when a high-frequency tone was presented and the left button when a low-frequency tone was presented. We did not ask observers to imagine the central line of the bisection stimulus. To ensure observers' attention to the stimuli, the outer lines were only half as long as normal in a few trials (0, 1, or 2 within a block of 80 trials). At the end of each block, observers were required to report the number of trials with shorter lines. The physical stimulation was identical to the imagery experiment (except for the 0–2 lines deviating per

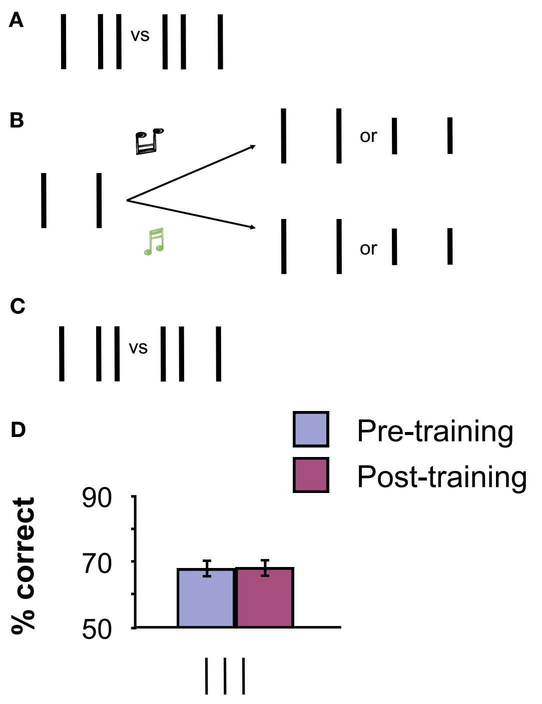


FIGURE 3 | (A,C) Pre and Post-training baseline measurements with the proper vertical bisection stimulus. **(B)** Counting lines. Only the two outer lines were presented. Observers were asked to press the right button when a high-frequency tone was presented and the left button for a low-frequency tone. In addition, observers were asked to count the number of trials with shorter lines (0, 1, or 2 out of 80 trials per block). No imagery was involved in this experiment. **(D)** No significant improvement of performance occurred through training for the vertical bisection. The error bars indicate the standard error of the mean (SEM) for six observers. Adapted from Tartaglia et al. (2009).

block; **Figure 3**). Performance did not improve in this condition (**Figure 3**). Hence, it is the mental imagery training that leads to perceptual learning. Further experiments demonstrated perceptual learning via mental imagery for Gabor and motion stimuli (Tartaglia et al., 2009, 2012).

Interestingly, perceptual learning via motion imagery improves coherent motion discrimination of moving dots when a blank screen is presented. However, no improvement occurs when, instead of a blank screen, randomly moving dots are presented during training (Tartaglia et al., 2012). Contrary to motion stimuli perceptual learning in the bisection task requires the presence of the two outer lines during training. Without the lines, performance does not improve. In mental imagery studies with static stimuli, it is often the case that a perceptual reference is needed (see the role of perceptual assistance shown in other studies, e.g., Intons-Peterson, 1981; Mast et al., 1999). Future research will better define the conditions under which mental imagery training will unfold its impact on perceptual learning. Taken together, it is possible to learn to see things that were not visible before imagery training. Thus, is perceptual learning via mental imagery truly creating new percepts? We will argue: yes and no.

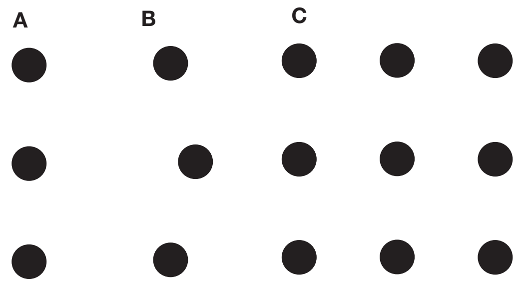


FIGURE 4 | *The dot world.* (A) Three dots are presented that appear to be aligned. **(B)** After extended training, observers can easily discriminate the offset direction of the center dot. Perceptual learning has made spatial differences visible that were not visible before. **(C)** Nine dots are presented. In this display, $\binom{9}{3} = 84$ tasks with positional changes of three dots can be defined (the other six dots constitute the background). In general, 2^9 tasks can be defined for 9 dots. In displays with n dots, hence, 2^n tasks can be defined leading to 2^n objects that one can potentially learn to perceive. $\binom{n}{k} = \frac{n!}{k(n-k)!}$ with $n! = 1*2*3*\ldots*n$. From Herzog and Esfeld (2009).

COMBINATORICS

In **Figures 4A,B** a 3-dot vernier is shown for which training improves performance (Fahle and Morgan, 1996). In **Figure 4C**, a set of nine dots is shown for which 84 3-dot vernier tasks can be defined and trained. Because of the specificity of perceptual learning, we expect no or very little transfer between tasks. Also 2 or 8-dot vernier tasks can be defined. In general, 256 vernier tasks can be defined by the 9-dot display. For 52 dots, there are more tasks than all the milliseconds in the universe that passed by since big bang. Hence, very small sets of elements can create very large combinatorial spaces.

We propose that mental imagery operates within such combinatorial spaces. The landmarks of the island that were memorized in the Kosslyn et al. (1978) experiment span a combinatorial space. Imagery does not add any new basic elements (e.g., a light tower) but computes, for example, second order aspects such as path lengths. In this sense, imagery does not create anything new. It just makes implicit aspects explicit. However, combinatorial spaces are large, often virtually infinite. Hence, the majority of combinatorial facts will remain unknown because only a small margin can be computed. In this sense, imagery produces new knowledge that may be unique to a certain person at a given moment in time.

On the first glance, it may appear as if perceptual learning via mental imagery can lead to really new percepts, as in the bisection stimulus where participants learn to perceive offsets that were not visible before. However, as mentioned above, without perceptual assistance provided by the outer lines, mental imagery training did not lead to perceptual learning. It seems that perceptual learning via mental imagery occurs within a model consisting of the two outer lines and the metric space in between. It is possible that the brain learns to better read out information from coding the center line. Hence, perceptual learning via mental imagery creates new percepts. However, these basic percepts are not so much different from detecting a new face in a mental image that was not visible during encoding. Still, all these percepts are new in the virtually infinite combinatorial space of things that can be learned and perceived. Once we have made the perceptual experience new knowledge can emerge from the combinatorial space. Really new percepts, however, are based on experience and they cannot be traced back to previous visual percepts. At this point, however, we enter the realm of philosophy and, admittedly, the definition of what a new percept is remains a debatable issue.

CONCLUSION

Mental imagery can create new percepts via perceptual learning. The examples we summarized here can be explained by mental imagery relying on a generating model combining and changing past percepts. This would allow for creating infinite numbers of new percepts within vast combinatorial spaces. We suggest that only perception can generate really new percepts in the sense of model free extrapolation.

ACKNOWLEDGMENTS

This research is supported by SNF grants awarded to Fred W. Mast (135303), Michael H. Herzog (114404), and Elisa M. Tartaglia (PBELP3-135838). We thank the reviewers for helpful suggestions.

REFERENCES

Ball, K., and Sekuler, R. (1987). Direction-specific improvement in motion discrimination. *Vision Res.* 27, 953–965.

Blakemore, C., and Cooper, G. F. (1970). Visual experience modifies distribution of horizontally and vertically oriented receptive fields in cats. *Science* 168, 869–871.

Chambers, D., and Reisberg, D. (1985). Can mental images be ambiguous? *J. Exp. Psychol. Hum. Percept. Perform.* 11, 317–328.

Chambers, D., and Reisberg, D. (1992). What an image depicts depends on what an image means. *Cogn. Psychol.* 24, 145–174.

Dror, I., and Kosslyn, S. M. (1993). Mental imagery and aging. *Psychol. Aging* 9, 90–102.

Einstein, A. (1929). *What Life Means to Einstein.* Saturday Evening Post, 26, October.

Fahle, M. (2005). Perceptual learning: specificity versus generalization. *Curr. Opin. Neurobiol.* 15, 154–160.

Fahle, M., and Edelman, S. (1993). Long-term learning in vernier acuity: effects of stimulus orientation, range and of feedback. *Vision Res.* 33, 397–412.

Fahle, M., and Morgan, M. (1996). No transfer of perceptual learning between similar stimuli in the same retinal position. *Curr. Biol.* 6, 292–297.

Fendick, A., and Westheimer, G. (1983). Effect of practice and the separation of test targets on foveal and peripheral stereoacuity. *Vision Res.* 23, 145–150.

Finke, R. A., and Pinker, S. (1982). Spontaneous imagery scanning in mental extrapolation. *J. Exp. Psychol. Learn Mem. Cogn.* 8, 142–147.

Fodor, J. A. (1975). *The Language of Thought.* New York: Crowell.

Herzog, M. H., and Esfeld, M. (2009). How the mind constitutes itself through perceptual learning. *Learn. Percept.* 1, 147–154.

Herzog, M. H., and Fahle, M. (1998). Modeling perceptual learning: difficulties and how they can be overcome. *Biol. Cybern.* 78, 107–117.

Hirsh, H. V. B., and Spinelli, D. N. (1970). Development of the brain depends on the visual environment. *Nature* 228, 477–478.

Intons-Peterson, M. J. (1981). Constructing and using unusual and common images. *J. Exp. Psychol. Hum. Learn.* 7, 133–144.

Intriligator, J., and Cavanagh, P. (2001). The spatial resolution of visual attention. *Cogn. Psychol.* 43, 171–216.

Karni, A., and Sagi, D. (1991). Where practice makes perfect in texture discrimination: evidence for primary visual cortex plasticity. *Proc. Natl. Acad. Sci. U.S.A.* 88, 4966–4970.

Kosslyn, S. M., Ball, T. M., and Reiser, B. J. (1978). Visual images preserve metric spatial information: evidence from studies of image scanning. *J. Exp. Psychol. Hum. Percept. Perform.* 4, 47–60.

Kosslyn, S. M., Ganis, G., and Thompson, W. L. (2003). Mental imagery: against the nihilistic hypothesis. *Trends Cogn. Sci. (Regul. Ed.)* 7, 109–111.

Kosslyn, S. M., Pascual-Leone, A., Felician, O., Camposano, S., Keenan, J. P., and Thompson, W. L., Ganis, G., Sukel, K. E., and Alpert, N. M. (1999). The role of area 17 in visual imagery: convergent evidence from PET and rTMS. *Science* 284, 167–170.

Kosslyn, S. M., Thompson, W. L., Kim, I. J., and Alpert, N. M. (1995). Topographical representations of mental images in primary visual cortex. *Nature* 378, 496–498.

Kumar, T., and Glaser, D. (1993). Initial performance, learning, and observer variability for hyperacuity tasks. *Vision Res.* 33, 2287–2300.

Lu, Z. L., and Dosher, B. A. (2004). Perceptual learning retunes the perceptual template in foveal orientation identification. *J. Vis.* 4, 44–56.

Mast, F., Kosslyn, S. M., and Berthoz, A. (1999). Visual mental imagery interferes with allocentric orientation judgements. *Neuroreport* 10, 3549–3553.

Mast, F. W., and Kosslyn, S. M. (2002). Visual mental images can be ambiguous: insights from individual differences in spatial transformation abilities. *Cognition* 86, 57–70.

Pearson, J., and Tong, F. (2011). Evaluating the mind's eye: the metacognition of visual imagery. *Psychol. Sci.* 22, 1535–1542.

Pourtois, G., Rauss, K. S., Vuilleumier, P., and Schwartz, S. (2008). Effects of perceptual learning on primary visual cortex activity in humans. *Vision Res.* 48, 55–62.

Pylyshyn, Z. W. (2002). Mental imagery: in search of a theory. *Behav. Brain Sci.* 25, 157–238.

Pylyshyn, Z. W. (2003). Return of the mental image: are there pictures in the brain? *Trends Cogn. Sci.* 7, 113–118.

Slotnick, S. D., Thompson, W. L., and Kosslyn, S. M. (2005). Visual mental imagery induces retinotopically organized activation of early visual areas. *Cereb. Cortex* 1570–1583.

Sowden, P. T., Rose, D., and Davies, I. R. (2002). Perceptual learning of luminance contrast detection: specific for spatial frequency and retinal location but not orientation. *Vision Res.* 42, 1249–1258.

Tartaglia, E. M., Bamert, L., Herzog, M. H., and Mast, F. W. (2012). Perceptual learning of motion discrimination by mental imagery. *J. Vis.* 12, 1–10.

Tartaglia, E. M., Bamert, L., Mast, F. W., and Herzog, M. H. (2009). Human perceptual learning by mental imagery. *Curr. Biol.* 19, 2081–2085.

Vogels, R., and Orban, G. A. (1985). The effect of practice on the oblique effect in line orientation judgements. *Vision Res.* 25, 1679–1687.

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: 01 May 2012; accepted: 04 September 2012; published online: 02 October 2012.

*Citation: Mast FW, Tartaglia EM and Herzog MH (2012) New percepts via mental imagery? Front. Psychology **3**:360. doi: 10.3389/fpsyg.2012.00360*

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

frontiers in
HUMAN NEUROSCIENCE

REVIEW ARTICLE
published: 01 April 2013
doi: 10.3389/fnhum.2013.00104

An emerging paradigm: a strength-based approach to exploring mental imagery

Tadhg E. MacIntyre[1]*, Aidan P. Moran[2], Christian Collet[3] and Aymeric Guillot[3]

[1] *Department of Physical Education and Sport Sciences, University of Limerick, Limerick, Ireland*
[2] *School of Psychology, University College Dublin, Ireland*
[3] *Centre de Recherche et d'Innovation sur le Sport, Université Claude Bernard Lyon 1, France*

Edited by:
Joel Pearson, The University of New South Wales, Australia

Reviewed by:
Joel Pearson, The University of New South Wales, Australia
Julie Krans, The University of New South Wales, Australia

***Correspondence:**
Tadhg E. MacIntyre, Department of Physical Education and Sport Sciences, University of Limerick, Limerick, Ireland.
e-mail: tadhg.macintyre@ul.ie

Mental imagery, or the ability to simulate in the mind information that is not currently perceived by the senses, has attracted considerable research interest in psychology since the early 1970's. Within the past two decades, research in this field—as in cognitive psychology more generally—has been dominated by neuroscientific methods that typically involve comparisons between imagery performance of participants from clinical populations with those who exhibit apparently normal cognitive functioning. Although this approach has been valuable in identifying key neural substrates of visual imagery, it has been less successful in understanding the possible mechanisms underlying another simulation process, namely, motor imagery or the mental rehearsal of actions without engaging in the actual movements involved. In order to address this oversight, a "strength-based" approach has been postulated which is concerned with understanding those on the high ability end of the imagery performance spectrum. Guided by the expert performance approach and principles of ecological validity, converging methods have the potential to enable imagery researchers to investigate the neural "signature" of elite performers, for example. Therefore, the purpose of this paper is to explain the origin, nature, and implications of the strength-based approach to mental imagery. Following a brief explanation of the background to this latter approach, we highlight some important theoretical advances yielded by recent research on mental practice, mental travel, and meta-imagery processes in expert athletes and dancers. Next, we consider the methodological implications of using a strength-based approach to investigate imagery processes. The implications for the field of motor cognition are outlined and specific research questions, in dynamic imagery, imagery perspective, measurement, multi-sensory imagery, and metacognition that may benefit from this approach in the future are sketched briefly.

Keywords: expertise, mental imagery, metacognition, motor cogniton, converging methods, mental practice, mental travel, mental rotation

INTRODUCTION

Since the classic study on mental rotations by Shepard and Metzler (1971) four decades ago, research on mental imagery, or "an internal representation that gives rise to the experience of perception in the absence of the appropriate sensory input" (Wraga and Kosslyn, 2002, p. 466), has flourished in cognitive psychology (e.g., Paivio et al., 1968), cognitive neuropsychology (e.g., Farah, 2000a), cognitive neuroscience (e.g., Kosslyn et al., 2006), and more recently, in the study of motor cognition (e.g., Jeannerod, 2001, 2006a). The common concern of researchers in these fields conveys the extent to which exploring mental imagery has become part of the cognitive science zeitgeist (Cornoldi and De Beni, 2012). Furthermore, *mental imagery* has also been of interest to those concerned with its application in the domains of skill acquisition (e.g., Sevdalis et al., 2013), rehabilitation sciences (e.g., Malouin and Richards, 2013) and professional expertise (e.g., Kozhevnikov and Blajenkova, 2013). What is it about mental imagery that has captured the interest of scientists across an array of disciplines for four decades? The answer to that question, as we shall see, is perhaps as complex as the construct of mental imagery itself.

Imagery, due to its ephemeral nature, has provided significant challenges for psychology since it first received formal research attention over a century ago (e.g., Galton, 1883). Chief among these challenges were the empirical question of how best to measure imagery given the limitations of the introspective methods (Kosslyn, 1980) and the theoretical question of whether the mental representation underlying imagery is propositional or analogical in nature (Block, 1983). Kosslyn et al. (2006) have provided a comprehensive proposal to resolve this latter issue. Despite such challenges, scientists have made significant progress in capturing the intricacies of mental imagery processes and applications by using new approaches (e.g., neuroscientific methods such as fMRI; Slotnick et al., 2005). Building on this need for new paradigms, the purpose of the present paper is to explore the possibility that a *strength-based* approach to imagery—one that focuses explicitly on recruiting participants at the high-ability end of the continuum of expertise—can augment conventional approaches to this construct. We argue that a strength-based approach may be valuable in illuminating both fundamental and applied questions (e.g., do expert athletes and/or dancers possess distinctive insights into meta-imagery processes in their domain?) which have so far evaded imagery researchers. In pursuit of this objective, the paper is organized as follows. In the first section, we shall review the strengths and weaknesses of conventional approaches to sampling that have dominated research on mental imagery for over a century. Next, we shall present an alternative approach to mental imagery research that we call the "strength-based" paradigm. The significant factors that led to the emergence of this paradigm will also be summarized. In the final section, we shall outline some potentially fruitful new directions in imagery research that can be addressed by supplementing traditional approaches with this strength-based paradigm.

KEY SAMPLES IN MENTAL IMAGERY RESEARCH

The study of imagery has profited greatly from neuroscientific methods (Behrmann, 2000). Given that it was a neglected topic for almost 50 years, its return from obscurity is remarkable (Kosslyn et al., 1995). Integral to the popularity of imagery as a legitimate scientific topic has been the increased strength in the inferences drawn from the research findings. These inferences are based on two aspects of the research process, the methodological tools and the approach to sampling (**Table 1**). The latter of these will be discussed in light of the converging methods approach of neuroscience (Kosslyn and Koenig, 1992).

Smith and Kosslyn (2007) describe the development of three main approaches in studying cognition: behavioral measures (e.g., latencies on computer-based tasks), correlational neural methods (e.g., fMRI), and causal neural methods (e.g., neuropsychological studies). Typical of this latter approach are comparisons between the performances of two groups on a task that requires imagery (i.e., behavioral data). The groups may comprise those who exhibit normal cognitive functioning or participants from clinical populations (Senior et al., 2006). Furthermore, the methods are often combined (i.e., behavioral data from normal and patient samples during brain imaging). Studying the deficits of patients with brain damage and impairments can establish that certain brain areas (e.g., by double dissociation) are at least partially responsible for a particular cognitive function. Neuropsychological evidence of both deficits following brain damage and brain activation patterns have been fundamental to the accumulating knowledge base of mental imagery (Kosslyn et al., 2006). Robust inferences can obviously be determined from monitoring changes in intact brain regions (Sarter et al., 1996), particularly when combined with theory-driven experimental paradigms (Kosslyn et al., 2006). Consequently, the neural-based studies with patients suffer from issues of generalizability but relying on converging evidence these limitations have largely been overcome.

From the above one could conclude that the cognitive neuroscience approach has encompassed participants across the performance spectrum, from those with deficits to those with high abilities. The nomothetic method has ultimately emphasized normative scores and even sophisticated statistical models are used to average date from neural imaging studies (Senior et al., 2006). Individual differences have not been central to the mental imagery research program in recent decades although as we shall see later in this paper, they provided an impetus for the strength-based approach. What is apparent from the aforementioned investigations is the paucity of research with participants at the high end of the continuum. Ability measures have been used widely to assess handedness (e.g., lateralization inventories) and they have been employed to discriminate good from poor imagers (e.g., Isaac and Marks, 1994). Indeed, highly skilled imagers may have been among the student samples in the classic imagery experiments of Kosslyn et al. (1978) and Shepard and Metzler (1971). The *deficit-based approach* which included both patients and healthy subjects may have overlooked potential insights, as we will propose later, from those who are highly skilled on imagery ability measures.

Nevertheless, the success of the prevailing paradigm has been the accumulation over four decades of substantive empirical evidence from case studies, behavioral data, and meta-analytic reviews (Kosslyn et al., 2006). Furthermore, evidence has emerged over the past two decades substantiating the neural basis of motor imagery (Munzert and Zentgraf, 2009; Munzert et al., 2009), olfactory imagery (Bensafi et al., 2003; Djordjevic et al., 2005), auditory imagery (Hubbard, 2010), gustatory imagery (Nitschke et al., 2006), spatial imagery (Kosslyn et al., 2007; Kozhevnikov and Blajenkova, 2013), and the role of emotions in imagery (Kosslyn, 2010).

However, some limitations in the key methods of cognitive neuroscience are also evident, and they may explain why our understanding of visual imagery is more comprehensive than that of other modalities (Moran et al., 2012). The scanning technologies to date (e.g., fMRI) restrict movement and thus limit maximal contractions during measurement. As a result, typically only discrete movements can be performed (e.g., small finger or foot movements; Pascual-Leone et al., 1995; Brown et al., 2005). Interestingly, transcranial magnetic stimulation (TMS) offers a different approach to studying action-one that is far more flexible as cortical functions can be disrupted even among normal controls (Walsh and Cowey, 2000; Stewart and Walsh, 2006). These methods may have obscured some aspects of imagery but we posit that in combination with the strength-based approach, they will ultimately illuminate the construct further.

ANTECEDENTS OF A "STRENGTH-BASED" APPROACH

In many respects, the "strength-based" approach to understanding mental imagery is not a new phenomenon. In fact, we believe that it has three key antecedents—the individual differences approach to imagery (e.g., Galton, 1883), the idea of imagery as a cognitive simulation process (e.g., Jeannerod, 1994), and the emergence of a movement (known as "motor cognition"; see Jeannerod, 2006a) which is committed to the investigation of imagery-action relationships. As we shall explain below, we believe that these three antecedents have contributed certain key ideas which come together in the strength-based approach (**Table 2**). For example, the individual differences tradition led to the discovery that certain mental imagery processes are trainable.

Table 1 | Summary of traditional approaches to sampling in mental imagery research with human participants.

Samples	Paradigm	Method	Goal	Strengths	Limitations	Landmark study
Experimenters were included in the sample of healthy subjects	Individual differences approach	Introspection	Understand phenomenology	Information appears to be easy to test	Highly subjective	Perky (1910) study of visual imagery and perception
Healthy subjects Large no. of subjects	Psychometric approach	Questionnaires	Focus on individual differences and sub-scale performance		Subjective Potential response bias (e.g., central tendency)	Marks (1973) Vividness of Visual Imagery Questionnaire (VVIQ)
Healthy subjects and typically right-handed (based on laterality tests) Large no. of subjects	Experimental approach	Response times (e.g., Mental scanning task)	Normalized data	Objective and easy to test large *N*	Trade-off between structure and process Some criticism of potential demand characteristics (Pylyshyn, 2003) and risk in use of sophisticated subjects	it took longer for subjects to consciously scan between image features that were relatively further apart than between those that appeared close together (Kosslyn et al., 1978)
	Cognitive neuropsychological approach	Correlational neural methods	Typically subtractive methods with normalized data combined with behavioral data			
Healthy subjects and clinical patients Systematic exclusion based on certain characteristics (e.g., imagery test performance)	nomothetic	EEG/ERP	Records changes in electrical potentials via electrodes	Relatively non-invasive measure of brain activity	Imprecise in resolution Only taps cortical changes Relatively expensive	EEG activity was measured in four tasks after sample was divided on the basis of VVIQ scores (Isaac and Marks, 1994)
Healthy subjects and clinical patients Small no. of subjects (<20)	nomothetic	PET/fMRI	Differences in the amounts of oxgyen consumed form basis for technique.	High spatial extent and resolution	Restricted movement for subject Highly expensive	Area 17 used in visual imagery and perception (Kosslyn et al., 1999)
Healthy subjects and clinical patients Small no. of subjects (<10)	nomothetic	TMS	Magnetic pulse can inhibit or facilitate function	Possible to target precise area	Low spatial extent Highly expensive	Used to inhibit imagery during rest condition prior to fMRI condition (Kosslyn et al., 1999)
Clinical patients 1-2 subjects with rare disorder	Neuropsychological case studies idiographic	Causal neural methods	Observation of patterns of dissociation and association	Ability to relate brain loci to functions	Uniqueness limits ability to generalize	Dissociation reported between mental imagery and object recognition in patient with brain damage (Behrmann et al., 1992)

Table 2 | Antecedents to the strength-based approach in mental imagery research.

Issue	Source	Contribution
Mental Practice (MP) effect	Vandell et al., 1943	Richardson (1967a,b) first narrative review of MP
Trainability of imagery	Shepard and Metzler, 1971	Subjects performed several thousand trials over 8–10 weeks
Mental travel effect	Decety et al., 1989	Congruence found between the duration of locomotion task and imagery
Converging methods approach	Kosslyn and Koenig, 1992	Kosslyn (1994) Neurally based computational model had superseded computational account (Kosslyn, 1980)
Role of deliberate practice in expertise	Ericsson et al., 1993	Expertise is domain-specific and is acquired through practice
Motor cognition approach	Jeannerod, 1994	Imagery is integral to motor preparation and action (Jeannerod, 2001, 2006a,b)
Individual differences in imagery ability	Kosslyn et al., 1998	MRT and rCBF study demonstrating two different ways to perform mental rotation, one that involves processes that execute movements and one that may not

Similarly, the theory of "imagery as simulation" has helped researchers to postulate theoretical explanations for certain robust effects in imagery research (e.g., expertise effects in research on the phenomena of "mental practice" and "mental travel"; see explanation of terms below). Finally, the motor cognition movement has enabled theoretical bridges to be built between imagery research in cognitive sport psychology and that in neuroscience (see also Moran et al., 2012).

INDIVIDUAL DIFFERENCES

It has long-been known (e.g., since Galton, 1883) that people differ reliably from each other in their imagery experiences and skills. For instance, whereas some people can report intense visual and motor images, akin to actually "seeing" the experience and "feeling" the effort (e.g., running up a stairs), others report little detail in their recollection of imagery. This discovery of individual differences in imagery has led to several important if somewhat under-appreciated breakthroughs in our understanding of this construct. For example, consider Shepard and Metzler's (1971) study of mental rotations or people's ability to turn things over in their mind in order to answer questions about the spatial orientation of certain pairs of shapes. It is worth noting that the participants in this study were not naïve subjects but highly practiced individual who had performed several thousand trials over 8–10 weeks. Crucially, this study highlighted the fact that mental rotation processes were not only measurable but *trainable* too. Interestingly, the extensive practice trials sparked follow-up research on the role of *practice* in mental rotation effects (Steiger and Yuille, 1983). Results showed that mental rotation effects were robust—regardless of the amount of training received by participants (Leone et al., 1993). Another notable finding from imagery studies in this era concerns Kosslyn's (1994) report that Jacky Metzler (co-author with Roger Shepard) had commented that some subjects had experienced what appeared to be "kinesthetic" sensations during imagery. It now seems that the mental rotations task, which was originally assumed to be visuo-spatial in nature, may actually involve the *motor* system. This discovery of motor involvement in visuospatial imagery could account for the early findings in neuroimaging studies of mental rotation that reported multiple brain areas activated during the task (Kosslyn et al., 2001a). Indeed, Kosslyn had argued that "visual mental images are transformed in part via motor processes" (1994, p. 345). A key question arose for researchers: do different strategies influence performance on mental rotations ability? In 1998, an rCBF technique was employed while subjects mentally rotated either hands or the original 3-D block objects (Kosslyn et al., 1998). The results highlighted that two mechanisms could be applied—"one mechanism that recruits processes that prepare motor movements and another that does not" (Kosslyn et al., 1998, p. 151). Findings on the neural basis of mental rotation, while dependent on the type of stimuli (animate vs. inanimate), may be *contaminated* by some subjects using a motor-based strategy. In a follow-up study, subjects either imagined rotating 3-D block isomers by hand or by a motor and the neural evidence demonstrated differential activation of the motor and visual cortex, respectively (Kosslyn et al., 2001b).

Another consequence of the individual differences movement was the range of expert samples used (e.g., pilots; Dror et al., 1993) and the range of methods expanded to explore in an ecological way the application of mental imagery in everyday life. For example, a weeklong diary study found that students reported using imagery for a range of purposes including problem solving (e.g., navigation) and mental practice (Kosslyn et al., 1990).

IMAGERY AND SIMULATION

The idea of imagery as a cognitive simulation technique may be traced back to William James' (1890) ideomotor principle (Jeannerod, 2006b). Since then, two different imagery effects have come to the attention of researchers—firstly, "mental practice," and secondly, "mental travel."

Firstly, mental practice, or "the systematic use of mental imagery to rehearse an action in the minds eye without engaging the actual physical movements involved" (Moran, 2012, p. 349), has been subjected to formal study since the 1940's (Vandell et al., 1943). In the intervening decades several hundred studies have investigated the mental practice effect, which based on meta-analytic research has a significant positive effect on the performance of motor skills (Driskell et al., 1994). A number of moderating variables were reported including the type of task and experience level. To explain, experienced athletes appear to benefit more from mental practice than do novices. This finding suggests a possible constraint on the efficacy of mental practice for novice learners. Specifically, as Driskell et al. (1994) proposed, that for "novices, who have not formed an approximation of the skill, the symbolic rehearsal provided by mental practice may not be sufficient to guide performance" (p. 489). One explanation for this expertise effect within mental practice is that the enhanced imagery abilities of experts may enable them to develop greater implicit knowledge of the spatial and kinesthetic requirements of the tasks than is possible for relative novices. Alternatively, experts may simply possess greater meta-cognitive knowledge of how to *employ* imagery effectively for skill improvement as compared with novices.

Since the early 1990's researchers have continued to show how mental practice can improve sport skills (e.g., flip-turn in swimming; Casby and Moran, 1998), surgical skills (Arora et al., 2011), and can accelerate the process of neurorehabilitation (McEwan et al., 2009). However, many questions remain unanswered. For example, what are the exact psychological mechanisms underlying mental practice effects (Kosslyn and Moulton, 2009)? How best does one apply imagery optimally in motor skill rehearsal (Weinberg, 2008)? And how do elite athletes employ mental practice in field settings (Moran, 2012)? Nevertheless, the mental practice literature highlighted the clever application of imagery among elite sport participants (Morris et al., 2005). And furthermore, the awareness of the mental practice effect created a common interest for sport psychologists and cognitive neuroscientists alike. While the former were primarily concerned with performance as the dependent variable, the latter were interested in the neural basis. In fact, researchers within the field of sport psychology have called for models of mental imagery in sport to be grounded in the neuroscientific literature and attempts at such theorizing have been tried (Holmes and Collins, 2001; Guillot and Collet, 2008; Wakefield et al., 2013).

The second imagery effect that has been reported in the literature is the mental travel effect, which is concerned with the relationship between the duration of a simulated movement (i.e., through imagery) and its executed counterpart (Decety et al., 1989). For example, in the early study by Decety et al. (1989), participants had to perform a blindfold walk and then mentally simulate the walk. Congruence between the duration for the tasks in each condition was reported in this study. Since then a robust mental travel effect has consistently been found while both underestimation and overestimation may occur in certain conditions, the temporal coupling of simulated and executed action is strongly influenced by expertise. Reviews have indicated that elite performers in sport are highly proficient at this skill (Guillot and Collet, 2005; Guillot et al., 2012a,b). For example, strong correlations have been reported between the time taken to rehearse completing a canoe-slalom course and the motor execution time (Moran and MacIntyre, 1998). Conversely, patients with developmental movement disorder (e.g., Gabbard et al., 2012), and a variety of neurological disorders of the motor system have been demonstrated to perform poorly on tests comparing the duration of simulated and executed movements (Guillot and Collet, 2005).

MOTOR COGNITION

As a consequence of the robust findings on imagery effects, motor imagery rose to prominence at the interface between cognitive neuroscience and sport psychology (Moran et al., 2012). Since Jeannerod's landmark paper in 1994, the implications of motor imagery as a window into the representation of action have been acknowledged by numerous researchers (e.g., Guillot and Collet, 2005; Smith and Kosslyn, 2007; Bläsing et al., 2012). The adoption of this new umbrella term, *motor cognition,* may now ensure that the study of action is within the realm of cognitive neuroscience after a degree of neglect by researchers within psychology (Rosenbaum, 2005). By implication the domains of dance (Bläsing et al., 2010), sport (Kosslyn and Moulton, 2009) have all, as a consequence, evolved as natural laboratories for the study of motor cognition. Why the neglect of motor control by psychology? Rosenbaum (2005) argues that "motor control has had the status of Cinderella in psychological research" (p. 308). Among the reasons for the failure of psychology to engage in attempts to understand aspects of motor control were the *"too hard to study hypothesis."* Undoubtedly the complexity of psychological aspects of motor control was a barrier to research, but the *raison d'être* may be that the methodological barriers hindered its exploration. It seems plausible to propose that motor imagery has only become a topic of study since researchers have had the methodological tools to explore it (i.e., through mental kinematics). The aforementioned factors, allied with the expertise paradigm, based on the theory of deliberate practice (Ericsson et al., 1993; Ericsson, 2009) provides the impetus for the strength-based approach.

THE STRENGTH-BASED APPROACH TO STUDYING MENTAL IMAGERY

The "strength-based" approach provides an extension of the prevailing paradigm of cognitive neuroscience, which has focused primarily upon investigations with both healthy subjects and patient samples. Instead, the emphasis is on those at the high ability end of the continuum. Furthermore, it posits that an expertise paradigm should be employed, to ensure that a comprehensive rationale is provided for the selection of participants (i.e., they meet multiple criteria to establish their level of performance). To tap this expertise efficiently, it is proposed that the converging methods approach is applied and moreover, that ecological validity is considered in the experimental design (**Table 3**).

Two key issues integral to the application of the strength-based approach are now highlighted. These are quantifying expertise and applying the principle of ecological validity. Subsequently, recent examples of this approach in music, sport, and dance samples are discussed. And finally, we highlight where we foresee the "strength-based" approach can illuminate the interaction between cognition and action (motor cogntion), the integration of multi-sensory information in perception and simulation, and the role of conscious thought and knowledge in imagery (metacognition).

EXPERTISE PARADIGM

Expert performance is defined as consistently superior performance on a specified set of representative tasks for the domain that can be administered to any subject (Ericsson, 2009). Ericsson and Smith (1991) proposed the expert performance approach to understand the critical mechanisms underlying expertise. Thus, the "strength-based" approach advances the prevailing research paradigms by explicitly exploring experts to enhance our understanding of mental imagery. As we have noted, this had been done

Table 3 | Proposed assumptions of the strength-based approach.

Assumption	Implications
Expertise approach (Ericsson et al., 1993)	• Use the dimension of expertise to choose samples • Employ multiple criteria to establish level of expertise • Explore expertise across domains relevant to mental imagery processes
Ecological validity (Neisser, 1976, 1978)	• To preserve the domain-specific expertise elements of ecological validity should be included in the study design • This should occur across the different dimensions-nature of the setting, stimuli, and response set
Converging methods (Kosslyn and Koenig, 1992)	• This approach should be employed to explore the interaction between abilities, the brain and computation
Theory-based approach	• Research questions should be guided by theory • Functional-equivalence and structural equivalent accounts of mental imagery provide a road map for research questions

previously but it occurred in an implicit fashion. For instance, the training in the original mental rotation study hinted at the development of expertise (Shepard and Metzler, 1971). Only recently have experts been targeted consistently as samples by cognitive neuroscientists (**Table 4**). Notable exceptions include the use of expert samples (i.e., artists) by Blajenkova et al. (2006) in their validation of the *Object-Spatial Imagery Questionnaire*. Another example of the implicit expertise approach was the inclusion of US air force pilots in a study by Dror et al. (1993) on their visual-spatial ability. However, in that case, we would assert that their expertise should be quantifiable across multiple criteria including their memory, metacognitive skills, and not simply the number of flying hours accrued.

The challenge for researchers, therefore, is to quantify precisely what constitutes an expert. Standardized ranking systems within domains like chess make this possible (Saarliluoma et al., 2004) but in sport, dance, and music, it is more difficult to have a standard metric. Typically researchers in sport have defined an expertise as a function of competitive level (e.g., novice, collegiate, elite, professional). This simple rubric, with elite being denoted as those competing at the highest possible level (Van den Auweele et al., 1993) is vague and may not adequately reflect the nature of the expertise. For example, is an elite athlete equivalent to a chess Grandmaster? Consequently, there is a need to fractionate experts from one another in terms of expertise. In domains such as music, dance, and sport "performance can be publically observed and even objectively measured in open competition and public performances" (Ericsson, 2009, p. 18). And moreover, given that expertise is distinguishable according to criteria, including metacognition (Ericsson, 2009), the precise performance across a matrix of measures of expertise should be explored (e.g., declarative knowledge, predictive ability).

Furthermore, comparisons across expert groups should be applied rather than simply exploring expert-novice differences. Experts and novices can easily be discriminated from one another in many domains on such an array of variables that the comparison can be meaningless. Instead the emphasis should be on inclusion criteria that target participants on the expertise spectrum based upon the research question. To explain, while the studies listed in **Table 4** have samples that reflect a range of abilities, it is evident that purposive sampling was applied. For example, to understand learning a novice sample were used (Bezzola et al., 2012) and on the other hand, a spectrum of abilities were represented in Ross et al.'s (2003) investigation of links between activated brain areas and golf skill.

The primacy of the expertise approach within the "strength-based" approach supplements the prevailing approach with the intention of exploring the role of expertise in mental imagery processes. Experts can be recruited, for instance, on the basis of their special imagery abilities or their sport, or professional activity expertise. Moreover, choosing appropriate activities from which to recruit samples should be based upon the cognitive task-demands of the actions, rather than an *ad-hoc* decision. One paradox is that clinical patients may develop specialized abilities in order to cope with the demands of their deficits. One such example is the case of IW, a patient with peripheral neuropathy, who has shown diminished motor imagery ability but enhanced visual imagery ability relative to controls (ter Horst et al., 2012). The inclusion of expertise as a variable in mental imagery research can shed light on both the processes underlying imagery and its potential application.

ECOLOGICAL VALIDITY

The expert performance approach of Ericsson and Smith (1991) proposes that field or laboratory tasks are designed in order to retain a high level of ecological validity. As Saarliluoma et al. (2004) states "it is important to vary the way basic concepts such as mental imagery are operationalized to avoid the metascientific *Ebbinghaus effect*" (p. 753–754). In other words, complex and dynamic tasks in which imagery processes are important should be subject to scientific scrutiny or aspects of the underlying processes may be overlooked. One persistent criticism of imagery research within the sport context is the use of tasks lacking ecological validity (Morris et al., 2005; Moran, 2012). This challenge applies to perhaps a greater degree within neuroscience where simple laboratory tasks (usually involving constrained movements of the fingers) that are chosen not only for their amenability to computational modeling but also for the ease with which they can be mastered after a relatively small amount of practice (Nielsen and Cohen, 2008; Yarrow et al., 2009). Ecological validity is a necessary component of studies that are targeting expertise as if we are to determine their abilities within a converging methods approach, some transfer of skill and process should occur (Moran, 2012). The dimensions of ecological validity include the nature of the setting, the stimuli, and the response (Schmuckler, 2001). These interlinked dimensions can be considered within the converging methods approach that is integral to the "strength-based" approach. For example, the diary study methodology conducted by Kosslyn et al. (1990) could be

Table 4 | Recent studies using neuroimaging methods in mental imagery with music, sport, and dance samples.

Question	Method	Authors	Sample	Findings
An attempt to define any association between activated brain areas and golf skill	fMRI of imagery of golf swing	Ross et al., 2003	6 golfers from novice to elite level	Decreased brain activation occurred with increased golf skill level for the SMA and cerebellum with little activation of basal ganglia
Investigation of the cortical network which mediates music performance compared to music imagery	fMRI	Meister et al., 2004	12 music academy students	Activations of premotor areas and precuneus were found in both conditions, contralateral M1 and posterior parietal cortex were active during performance only
Comparison of neural networks of expert and novice golfers during simulation of pre-shot routine	fMRI of imagined pre-shot routine	Milton et al., 2007	6 expert and 7 novice golfers	Extensive practice leads experts to develop a focused and efficient organization of task-related neural networks, whereas novices have difficulty filtering out irrelevant information
To investigate differences in brain activity between groups and to effect of the use of internal vs. external perspective	fMRI of motor imagery of a high jump	Olsson et al., 2008	High-jumpers (12 elite and 12 novices)	Imagery training reduces the activity in parietal cortex suggesting that imagery is performed more automatic and results in a more efficient motor representation more easily accessed during motor performance
Role of experience in facilitating corticospinal representations of actions	TMS of familiar and unfamiliar skills	Fourkas et al., 2008	3 expert tennis players	Subjective reports indicated that only in the tennis imagery condition did experts differ from novices in the ability to form proprioceptive images
To investigate multi-modal musical imagery performed by expert pianist	fMRI during imagery and simulated motor performance of a memorized extract	Davidson-Kelly et al., 2011	42-year-old expert pianist	Pattern of activation for performed and imagined piano music was similar, with the motor system of the brain showing similar activation during both conditions (except for M1)
Study of dynamic neurofunctional changes induced by a physical training	fMRI of golf putts in longitudinal study over 40 h training	Bezzola et al., 2012	11 novice golfers and age-matched controls	Training induces functional neuroplasticity and skill improvement is associated with a modified activation pattern

utilized to evaluate experts' imagery use over time (e.g., professional dancers). This approach offers researchers a naturalistic laboratory for investigating imagery and action (Moran, 2009).

NEW APPROACHES TO OLD QUESTIONS

Borrowing from Boring (1957), it may be argued that imagery research has a long past but only a short history and that some enduring questions in this field remain unanswered. As we have noted, perennial issues around imagery measurement, theory, and function have been central to enquiry for over a century (Roekelein, 2004). For instance, Jeannerod (2006b) states that the "mental conception of action" or the "motor idea," "to account for the role of memory images or remote impressions in shaping an action" has been with us since the time of William James (p. 360). The key issue is how the prevailing paradigm of cognitive neuroscience has shifted in recent decades. We have seen how the field of cognitive psychology has been reconstituted since the 1990s with an increasing emphasis on neural implementation (Smith and Kosslyn, 2007; Anderson, 2010). Moreover, the key topics that delineated cognitive psychology in the early textbooks have been extended to include neuroscience methods, emotion, and action (Smith and Kosslyn, 2007). Furthermore, Martha Farah has stated that the present paradigm of cognitive neuroscience "is far from having outlived its usefulness…and I'd like to see it continue to move towards the edge of our understanding" (2000b, p. 362).

The evidence presented heretofore, on the growth of the "strength-based" approach, conveys how it has augmented the study of abilities within the cognitive neuroscience triangle (Kosslyn, 1994). The expansion of the "strength-based" approach within mental imagery research opens up new modes of enquiry for mental imagery and perception (e.g., Tartaglia et al., 2009) and specifically in our understanding of action processes. Motor cognition research has the potential to shed light on imagery processes, the representation of action, and the role of imagery processes in experts.

MOTOR COGNITION

Three key questions within the field of motor cognition can be understood by applying the "strength-based" approach. One issue surrounds the role of action coupled with motor imagery, what has been termed dynamic imagery (MacIntyre and Moran, 2010). Motor imagery, by definition, occurs in the absence of any overt movement or motor output (Guillot and Collet, 2010). However, on the basis that athletes often engage in movements while engaging in imagery, sport psychologists have recommended that performers apply this in their imagery practice (Holmes and Collins, 2001) and moreover, they have amended their definition of mental imagery to include possible motor output (Morris et al., 2005). Researchers had noted that athletes engaged in either synchronous movements (e.g., simulating the task) or asynchronous movements (e.g., other movements) during imagery (MacIntyre and Moran, 2010). The role of these quasi-movements (Nikulin et al., 2008) has yet to be rigorously evaluated. Preliminary evidence suggests that athletes find this beneficial (MacIntyre and Moran, 2010) but to date research has not investigated the complex issue of coupling action and motor imagery by athletes (for an exception see Guillot et al., 2013). Evidence from other samples suggests that this is a topic worthy of further research. For example, Ionta et al. (2012) reported that variations in the hand position of participants' during mental rotations tasks influenced the latencies for congruent stimuli. They concluded that sensory-motor and postural information coming from the body might influence mental rotation of body parts according to specific, somatotopic rules. These preliminary findings were congruent with the *body-specificity hypothesis* that claims that body-specific patterns of motor experience shape the way we think (Casasanto, 2011). Furthermore, future findings from this line of enquiry may have ramifications for the recent accounts of embodied cognition (Borghi and Cimatti, 2010; Gallese and Sinigaglia, 2012). One confound that has been noted in research on the action-motor imagery coupling is the imagery perspective (the viewpoint adopted during visual imagery). This has resulted in debate in both sport psychology and cognitive neuroscience. Moran (2012) noted that the complexity of agency, visual perspective, and confounds with kinesthetic or motor imagery were reflected in the findings emanating from sport psychology. Researchers had developed sophisticated methodologies to attempt to control and measure the visual perspective adopted during testing. The "strength-based" approach is one possible route to understanding this topic further. A recent special issue of the *Journal of Mental Imagery* on whether the internal viewpoint is a default hypothesis is testament to the continuing interest in this topic (Morris and Spittle, 2012). The issue of visual perspective in mental imagery, because of the conflation with both visual imagery and motor imagery in the past, has been noted as a topic that necessitates further research (Madan and Singhal, 2012).

A second question that relates to motor cognition is the process of multi-sensory integration, which is an area of current debate (Foxe and Molholm, 2009). More specifically, the relative contribution of different senses in the simulation of action is of direct concern to neuroscientists (Lacey and Lawson, 2013). Recently, this topic has received increased attention because of concerns with the traditional approach in the understanding and application of multi-sensory imagery. For example, within sport psychology, it was traditionally assumed that a multi-modal approach was optimal (Morris et al., 2005). However, this has recently been questioned on the basis that it may be more important to only imagine the pertinent senses (Holmes and Collins, 2001; Weinberg, 2008; MacIntyre and Moran, 2010). Initial evidence, from fMRI studies (Ross et al., 2003), dual-task studies (Smyth and Waller, 1998), and qualitative accounts of how elite performers employ imagery (Munroe et al., 2000; MacIntyre and Moran, 2007a,b) suggests that the simplistic multi-sensory application of imagery merits further investigation. Interestingly, this topic would be of interest across the domains of music, sport, and dance that have been referred to in this paper. From another perspective, the importance of different senses underlying the reported imagery effects has yet to be fully ascertained and both inhibitory processes during imagery and consolidation effects are only beginning to be explored (e.g., Guillot et al., 2012a,b).

Fourthly, as discussed in the introduction, measuring imagery ability has been an issue of controversy for the field since Galton's early attempts at quantification (Galton, 1883). The question "why do people differ so much in their imagery abilities" (Kosslyn et al., 2001a, p. 641) is still pertinent today. The influence of implicit or explicit expertise on mental rotations findings, in mental travel research, and in mental practice studies has been established (Guillot and Collet, 2005). And moreover, a trend in imagery research has been the expansion of the imagery ability from one unitary construct to a number of distinct abilities which reflect different neural processes (e.g., dorsal vs. ventral stream; Blajenkova et al., 2006). What is less clear is how imagery abilities are developed and what is the precise role of these imagery abilities in moderating imagery effects (Madan and Singhal, 2012). One alternative to the plethora of pencil and paper imagery tests employed in sport (e.g., Williams and Cumming, 2011) or neuroscience settings (e.g., McAvinue and Robertson, 2007) and is to employ a compound measure, the motor imagery index, which combines psychometric, behavioral (e.g., mental travel), and psychophysiological measures (Collet et al., 2011). It is vital that we are able to quantify imagery abilities if we are to match participants for competence or if we wish to evaluate the trainability of imagery abilities. Recent research has explored the trainability of imagery vividness using robust measures and interestingly, the only reported changes were in the metacognitive understanding of their imagery (Rademaker and Pearson, 2012).

META-IMAGERY

Theoretically, a potentially valuable new route for imagery researchers in cognitive neuroscience concerns the investigation of the neglected topic of "meta-imagery processes"—"their beliefs about the nature and regulation of their own imagery skills" (Moran, 2002, p. 415). Research in the expertise literature suggests that meta-cognition, people's insight into, and control over their own mental processes, is a factor that differentiates novices from experts (Moran, 2012). Interest in this topic surprisingly emerged from a survey by Denis and Carfantan (1985) who surveyed undergraduate students on their knowledge about imagery research findings in psychology. The motive for their study was to quantify the participants' *tacit knowledge* of imagery effects (Denis, 2012). In order to assess the level of *tacit knowledge* among experimental subjects they were asked to predict the outcomes of various imagery experiments that were described but not formally named (e.g., is memory for pictures better than memory for words?). The findings indicated that although the majority of participants predicted correctly that imagery would have beneficial effects on learning and reasoning, few subjects were able to predict accurately the results of mental rotation experiments (whereby more time is required to accomplish greater amounts of rotation of images) or mental scanning studies (whereby longer distances between points in an image take longer to scan than shorter distances). Furthermore, a majority of participants rejected the idea that mental imagery could enhance the performance of motor skills (the mental practice effect). This latter finding led Denis and Carfantan (1985) to conclude "how counterintuitive the idea is that motor skills may be affected by purely mental practice" (p. 56). The naïve responses of the participants in this study are in stark contrast to the evidence that has recently emerged from sporting samples. Researchers have asked athletes and dancers to indicate why, where, how, what, and when they use mental imagery processes (e.g., Nordin and Cumming, 2005; MacIntyre and Moran, 2007a,b). Athletes' responses from both interviews and surveys demonstrated a comprehensive knowledge of the multi-model potential of imagery, their awareness of both mental practice and mental travel effect, and highlighted the sophisticated nature of their understanding of imagery (MacIntyre and Moran, 2010).

In 2002, Moran suggested that it would be interesting to discover if top athletes have greater insight into and control over their use of imagery compared with their less successful counterparts. To date the preponderance of the evidence favors an expertise effect for meta-imagery. In fact, a model of meta-imagery in action suggests that there are three components-knowledge, monitoring and control, which opens up possibilities of developing a test of meta-imagery (MacIntyre and Moran, 2010). Furthermore, contemporary evidence from cognitive psychology supports the role of meta-cognitive knowledge of imagery ability and relates it to our ability to judge individual episodes of imagery (Pearson et al., 2011). The voluntary nature of imagery and the role of conscious awareness during imagery tasks make it amenable to introspection (Pearson et al., 2008), ironically the method that was central to the demise of the scientific study of imagery, a century ago (see Block, 1983). While in the past the study of metacognition has targeted intellectual skills "if intellectual and perceptual-motor skills rely on similar mechanisms, one would expect metacognition to apply to the guidance of perceptual-motor skills, just as it does to the guidance of intellectual skills" (Augustyn and Rosenbaum, 2005, p. 911). Consequently, armed with a comprehensive roadmap of imagery processes and an increased understanding of action, the study of meta-imagery could provide a back door into the typically impenetrable realm of sensory perception (Pearson et al., 2011).

SOME OBSERVATIONS ON THE STRENGTH-BASED APPROACH WITHIN COGNITIVE NEUROSCIENCE

In this paper, we have argued that a strength-based approach to mental imagery can augment rather than replace the traditional approach to this construct. After all, this latter approach has been highly successful in, for example, illuminating both visual imagery and visual cognition (Kosslyn et al., 2006). More recently, this latter approach has led to advances in our understanding of the overlap between visual perception and imagery in scanning tasks (Borst and Kosslyn, 2008). Based on such progress, there is an imperative for researchers to maintain the traditional approach in order to answer key questions. For example, the use of randomized controlled trials to explore the role of motor imagery in patients with sub-acute neglect (Welfringer et al., 2011) and with stroke victims (Ietswaart et al., 2011) are essential for the validation of imagery interventions. Similarly, the role of mental imagery among clinical patient samples (Pearson et al., 2012) and stroke victims (Confalonieri et al., 2012) requires continued investigation. Unsurprisingly, studies of patients will, in all likelihood, continue to inform our appreciation of the deficits, challenges, and recovery strategies of those with specific with rare disorders (ter Horst et al., 2012). This again raises aforementioned paradox of expertise among patient samples. A recent study with a Paralympic athlete led to the conclusion that only tasks that we have physical experience of recruit the motor system (Olson and Nyberg, 2011).

Despite the success of traditional approaches to imagery, however, a strength-based approach may contribute in the development and refinement of imagery inventories, as implicitly employed in the case of the *Object-Spatial Imagery Questionnaire* (Blajenkova et al., 2006). Similarly, investigations of the neural processes underlying imagery effects in expert samples will help to elucidate both applied and theoretical aspects of mental imagery. However, the new approach advocated in this paper has its limits too. For example, analyzing the imagery skills of motor experts in an atheoretical manner will not advance conceptual understanding and may provide spurious findings. Instead, a rigorous, theory-driven approach with converging methods is required for strength-based approaches to yield benefits to imagery researchers. As anticipated by Kosslyn et al. (2002) over a decade ago, "individual differences can actually help to reveal the nature of underlying mechanisms (p. 342)."

CONCLUSIONS

Cognitive neuroscience has made impressive progress in the illumination of the nature, function, and neural basis of mental imagery (Kosslyn, 2010). Nevertheless, certain aspects of this construct (e.g., its relationship to skilled performance) have been relatively neglected by mainstream imagery researchers. In this article, certain significant trends—antecedents of a strength-based approach to imagery—that may be detected over more than a century of imagery research have been highlighted. These trends provide signposts for at least four potentially fruitful new avenues of inquiry. Firstly, imagery research continues to evolve from an excessively narrow focus on the visual sense (e.g., Galton, 1883; Perky, 1910) to a modern concern with investigation of the true multi-sensory character of this construct. Next, theoretical understanding of mental imagery has deepened with increasing awareness of the multi-dimensional nature of this construct. Thirdly, major methodological advances, both in technology and sampling, have illuminated both fundamental and applied questions. Finally, the preliminary application of the "strength-based" approach has been shown to be useful in enriching our understanding of the neural basis of expert performance (Ross et al., 2003; Milton et al., 2007; Olsson et al., 2008) and the "clever application" of imagery in sport (Moran et al., 2012), dance (Bläsing et al., 2012), and music contexts (Meister et al., 2004). It is noteworthy that the interest in exploring the naturalistic expertise in imagery (e.g., dance) with neuroscientific methods stems is bi-directional. For example, "neuroscientists have turned to dancers as a valuable human resource in possession of a rich skill set…to address issues of how the brain coordinates perception with action" (Cross, 2010, p. 197). The implications of augmenting existing paradigms with the "strength-based" approach will be most obvious in the domain of motor cognition where issues of dynamic imagery, visual imagery perspective, and multi-modal integration can be explored. Also, the area of metacognition research is a rapidly growing field and meta-imagery as a topic is evolving. Meta-imagery research is a topic than can benefit from the "strength-based" approach. In summary, through the lens of motor cognition, further exploration of the construct of mental imagery will ensure that the topic of will remain part of the cognitive neuroscience lexicon for many decades to come.

REFERENCES

Anderson, J. R. (2010). *Cognitive Psychology and its Implications.* New York, NY: Worth Publishers.

Arora, S., Aggarwal, R., Sirimanna, P., Moran, A., Grantcharov, T., Kneebone, R., et al. (2011). Mental practice enhances surgical technical skills: a randomized controlled study. *Ann. Surg.* 253, 265–270.

Augustyn, J. S., and Rosenbaum, D. A. (2005). Metacognitive control of action: preparation for aiming reflects knowledge of Fitts' Law. *Psychon. Bull. Rev.* 12, 911–916.

Behrmann, M. (2000). The mind's eye mapped onto the brain's matter. *Curr. Dir. Psychol. Sci.* 9, 50–54.

Behrmann, M., Winocur, G., and Moscovitch, M. (1992). Dissociation between mental imagery and object recognition in a brain-damaged patient. *Nature* 359, 636–637.

Bensafi, M., Porter, J., Pouliot, S., Mainland, J., Johnson, B., Zelano, C., et al. (2003). Olfactory activity during imagery mimics that during perception. *Nat. Neurosci.* 6, 1142–1144.

Bezzola, L., Merillat, S., and Jancke, L. (2012). The effect of leisure activity golf practice on motor imagery: an fMRI study in middle adulthood. *Front. Hum. Neurosci.* 6:67. doi: 10.3389/fnhum.2012.00067

Blajenkova, O., Kozhevnikov, M., and Motes, M. A. (2006). Object-spatial imagery: a new self-report imagery questionnaire. *Appl. Cogn. Psychol.* 20, 239–263.

Bläsing, B. E., Calvo-Merino, B., Cross, E. S., Jola, C., Honisch, J., and Stevens, C. J. (2012). Neurocognitive control in dance perception and performance. *Acta Psychol.* 139, 300–308.

Bläsing, B., Puttke, M., and Schack, T. (2010). "The neurocognition of dance. Introduction," in *The Neurocognition of Dance: Mind, Movement and Motor Skills,* eds B. Bläsing, M. Puttke, and T. Schack (Hove, East Sussex: Psychology Press), 1–8.

Block, N. (1983). Mental pictures and cognitive science. *Philos. Rev.* 92, 499–541.

Borghi, A. M., and Cimatti, F. (2010). Embodied cognition and beyond: acting and sensing the body. *Neuropsychologia* 48, 763–773.

Boring, E. G. (1957). *A History of Experimental Psychology.* New York, NY: Prenctice Hall.

Borst, G., and Kosslyn, S. M. (2008). Visual mental imagery and visual perception: structural equivalence revealed by scanning processes. *Mem. Cognit.* 36, 849–862.

Brown, S., Martinez, M. J., and Parsons, L. M. (2005). The neural basis of human dance. *Cereb. Cortex* 16, 1157–1167.

Casasanto, D. (2011). Different bodies, different minds: the body specificity of language and thought. *Curr. Dir. Psychol. Sci.* 20, 378–383.

Casby, A., and Moran, A. (1998). Exploring mental imagery in swimmers: a single-case study design. *Ir. J. Psychol.* 19, 525–531.

Collet, C., Guillot, A., Lebon, F., MacIntyre, T., and Moran, A. (2011). Measuring motor imagery: combining psychometric, qualitative, chronometric, and psychophysiological techniques. *Exerc. Sport Sci. Rev.* 39, 89–92.

Confalonieri, L., Pagnoni, G., Barsalou, L. W., Rajendra, J., Eickhoff, S. B., and Butler, A. J. (2012). Brain activation in primary motor and somatosensory cortices during motor imagery correlates with motor imagery ability in stroke patients. *ISRN Neurol.* 2012, 1–17.

Cornoldi, C., and De Beni, R. (2012). "Foreword," in *Essays in Honour of Michel Denis: From Mental Imagery to Spatial Cognition and Language,* eds V. Gyselinck and F. Pazzaglia (Hove, East Sussex: Psychology Press), ix–x.

Cross, E. S. (2010). "Building a dance in the human brain: Insights from expert and novice dancers," in *The Neurocognition of Dance: Mind, Movement and Motor Skills* eds B. Bläsing, M. Puttke, and T. Schack (Hove, East Sussex: Psychology Press), 177–202.

Davidson-Kelly, K., Hong, S., Dhinakaran, J., Joseph, M., and Gray, C. (2011). "An fMRI study of expert musical imagery: to what extent do imagined and executed performance share the same neural substrate," in *Proceedings of the International Symposium on Performance Science 2011,* eds A. Williamon, D. Edwards, and L. Bartel (Utrecht: European Association of Conservatoires (AEC), Toronto), 613–617.

Decety, J., Jeannerod, M., and Prablanc, C. (1989). The timing of mentally represented actions. *Behav. Brain Res.* 34, 35–42.

Denis, M. (2012). "Decades of imagery. Reminisences of a shared journey," in *Essays in honour of Michel Denis: From Mental Imagery to Spatial Cognition and Language,* eds V. Gyselinck and F. Pazzaglia (Hove, East Sussex: Psychology Press), 203–254.

Denis, M., and Carfantan, M. (1985). People's knowledge about images. *Cognition* 20, 49–60.

Djordjevic, J., Zatorre, R. J., Petrides, M., Boyle, J. A., and Jones-Gotman, M. (2005). Functional neuroimaging of odor imagery. *Neuroimage* 24, 791–801.

Driskell, J., Copper, C., and Moran, A. (1994). Does mental practice enhance performance? A meta-analysis. *J. Appl. Psychol.* 79, 481–492.

Dror, I., Kosslyn, S. M., and Waag, W. (1993). Visual-spatial abilities of pilots. *J. Appl. Psychol.* 78, 763–773.

Ericsson, K. A. (2009). *Development of Professional Expertise: Toward Measurement of Expert Performance and Design of Optimal Learning Environments*. Cambridge, MA: Cambridge University Press.

Ericsson, K. A., Krampe, R. T., and Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychol. Rev.* 100, 363–406.

Ericsson, K. A., and Smith, J. (1991). "Prospects and limits in the empirical study of expertise: an introduction," in *Toward a General Theory of Expertise: Prospects and Limits,* eds K. A. Ericsson and J. Smith (Cambridge: Cambridge University Press), 1–38.

Farah, M. J. (2000a). "The neural basis of mental imagery," in *The Cognitive Neurosciences*, ed M. S. Gazzaniga (Cambridge, MA: MIT Press), 965–974.

Farah, M. J. (2000b). Interview with Martha Farah. *J. Cogn. Neurosci.* 12, 360–363.

Fourkas, A. D., Bonavolonta, V., Avenanta, A., and Aglioti, S. M. (2008). Kinesthetic imagery and tool-specific modulation of corticospinal representations in expert tennis players. *Cereb. Cortex* 18, 2382–2390.

Foxe, J. J., and Molholm, S. (2009). Ten years at the Multisensory Forum: musings on the evolution of a field. *Brain Topogr.* 21, 149–154.

Gabbard, P., Cacola, P., and Bobbio, T. (2012). The ability to mentally represent action is associated with low motor imagery ability in children: a preliminary investigation. *Child Care Health Dev.* 38, 390–393.

Gallese, V., and Sinigaglia, C. (2012). What is so special about embodied simulation? *Trends Cogn. Neurosci.* 15, 512–519.

Galton, F. (1883). *Enquiries into Human Faculty*. London: Dent.

Guillot, A., and Collet, C. (2005). Duration of mentally simulated movement: a review. *J. Mot. Behav.* 37, 10–20.

Guillot, A., and Collet, C. (2008). Construction of the motor imagery integrative model in sport: a review and theoretical investigation of motor imagery use. *Int. Rev. Sport Exerc. Psychol.* 1, 31–44.

Guillot, A., and Collet, C. (2010). *The Neurophysiological Foundations of Mental and Motor Imagery*. Oxford: Oxford University Press.

Guillot, A., Di Rienzo, F., MacIntyre, T., Moran, A., and Christian, C. (2012a). Imagining is not doing but involves specific motor commands: a review of experimental data related to motor inhibition. *Front. Hum. Neurosci.* 6:247. doi: 10.3389/fnhum.2012.00247

Guillot, A., Hoyek, N., Louis, M., and Collet, C. (2012b). Learning by thinking: review and future directions for the timing of imagined movements. *Int. Rev. Sport Exerc. Psychol.* 5, 3–22.

Guillot, A., Moschberger, K., and Collet, C. (2013). Coupling movement with imagery as a new perspective for motor imagery practice. *Behav. Brain Funct.* 9, 1–17.

Holmes, P. S., and Collins, D. J. (2001). The PETTLEP approach to motor imagery: a functionalequivalence model for sport psychologists. *J. Appl. Sport Psychol.* 13, 60–83.

Hubbard, T. L. (2010). Auditory imagery: empirical findings. *Psychol. Bull.* 136, 302–329.

Ietswaart, M., Johnston, M., Dijkerman, H. C., Joice, S., Scott, C. L., MacWalter, R. S. et al. (2011). Mental practice with motor imagery in stroke recovery: randomized controlled trial of efficacy. *Brain* 134, 1373–1386.

Ionta, S., Perruchoud, D., Draganski, B., and Blanke, O. (2012). Body context and posture affect mental imagery of hands. *PLoS ONE* 7:e34382. doi: 10.1371/journal.pone.0034382

Isaac, A. R., and Marks, D. F. (1994). Individual differences in mental imagery experience: developmental changes and specialisation. *Br. J. Psychol.* 85, 479–500.

James, W. (1890). *Principles of Psychology.* New York, NY: Holt, Rinehart and Winston.

Jeannerod, M. (1994). The representing brain: neural correlates of motor intention and imagery. *Behav. Brain Sci.* 17, 187–202.

Jeannerod, M. (2001). Neural simulation of action: a unifying mechanism for motor cognition. *Neuroimage* 14, S103–S109.

Jeannerod, M. (2006a). *Motor Cognition*. New York, NY: Oxford University Press.

Jeannerod, M. (2006b). The origin of voluntary action. History of a physiological concept. *C. R. Biol.* 329, 354–362.

Kosslyn, S. M. (1980). *Image and Mind*. Cambridge, MA: Harvard University Press.

Kosslyn, S. M. (1994). *Image and Brain*. Cambridge, MA: MIT Press.

Kosslyn, S. M. (2010). "Multimodal images in the brain," in *The Neurophysiological Foundations of Mental and Motor Imagery,* eds A. Guillot and C. Collet (New York, NY: Oxford University Press), 3–16.

Kosslyn, S. M., Ball, T. M., and Reiser, B. J. (1978). Visual images preserve metric spatial information: evidence from studies of image scanning. *J. Exp. Psychol. Hum.* 4, 47–60.

Kosslyn, S. M., Behrmann, M., and Jeannerod, M. (1995). The cognitive neuroscience of mental imagery. *Neuropsychologia* 33, 1335–1344.

Kosslyn, S. M., Cacioppo, J. T., Davidson, R. J., Hugdahl, K., Lovallo, W. R., Spiegel, D., et al. (2002). Bridging psychology and biology: the analysis of individuals in groups. *Am. Psychol.* 57, 341–351.

Kosslyn, S. M., Digirolamo, G. J., Thompson, W. L., and Alpert, N. M. (1998). Mental rotation of objects versus hands: neural mechanisms revealed by positron emission tomography. *Psychophysiology* 35, 151–161.

Kosslyn, S. M., Ganis, G., and Thompson, W. L. (2001a). Neural foundations of imagery. *Nat. Rev. Neurosci.* 2, 635–642.

Kosslyn, S. M., Thompson, W. L., Wraga, M., and Alpert, N. M. (2001b). Imagining rotation by endogenous and exogenous forces: Distinct neural mechanisms for different strategies. *Neuroreport* 12, 2519–2525.

Kosslyn, S. M., and Koenig, O. (1992). *Wet Mind. The New Cognitive Neuroscience.* New York/Toronto: The Free Press.

Kosslyn, S. M., and Moulton, S. T. (2009). "Mental imagery and implicit memory," in *Handbook of Imagination and Mental Simulation,* eds K. D. Markman, W. M. P. Klein, and J. A. Suhr (New York, NY: Psychology Press), 35–52.

Kosslyn, S. M., Pascual-Leone, A., Felician, O., Camposano, S., Keenan, J. P., Thompson, W. L., et al. (1999). The role of area 17 in visual imagery: convergent evidence from PET and rTMS. *Science* 284, 167–170.

Kosslyn, S. M., Seger, C., Pani, J., and Hillger, L. A. (1990). When is imagery used in everyday life? A diary study. *J. Ment. Imagery* 14, 131–152.

Kosslyn, S. M., Shephard, J. M., and Thompson, W. L. (2007). "Spatial processing during mental imagery: a neurofunctional theory," in *Spatial Processing in Navigation, Imagery, and Perception,* eds F. Mast and L. Jancke (New York, NY: Springer), 1–16.

Kosslyn, S. M., Thompson, W. L., and Ganis, G. (2006). *The Case for Mental Imagery.* Oxford: Oxford University Press.

Kozhevnikov, M., and Blajenkova, O. (2013). "Individual differences in object versus spatial imagery: from neural correlates to real-world applications," in *Theoretical and Applied Perspectives on Multisensory Imagery,* eds S. Lacey and R. Lawson (New York, NY: Springer), 299–318.

Lacey, S., and Lawson, R. (2013). *Theoretical and Applied Perspectives on Multisensory Imagery*. New York, NY: Springer.

Leone, G., Taine, M. C., and Droulez, J. (1993). The influence of long-term practice on mental rotation strategies of 3-D objects. *Brain Res. Cogn. Brain Res.* 1, 241–255.

MacIntyre, T., and Moran, A. (2007a). A qualitative investigation of imagery use and meta-imagery processes among elite canoe-slalom competitors. *J. Imagery Res. Sport Phys. Act.* 2, 1–23.

MacIntyre, T., and Moran, A. (2007b). A qualitative investigation of meta-imagery processes and imagery direction among elite athletes. *J. Imagery Res. Sport Phys. Act.* 2, 1–20.

MacIntyre, T., and Moran, A. (2010). "Meta-imagery processes among elite sports performers," in *The Neurophysiological Foundations of Mental and Motor Imagery,* eds A. Guillot and C. Collet (New York, NY: Oxford University Press), 227–244.

Madan, C. R., and Singhal, A. (2012). Motor imagery and higher level cognition: four hurdles before research can sprint forward. *Cogn. Process.* 13, 211–229.

Malouin, F., and Richards, C. L. (2013). "Clinical applications of motor imagery in rehabilitation," in *Theoretical and Applied Perspectives on Multisensory Imagery,* eds S. Lacey and R. Lawson (New York, NY: Springer), 397–420.

Marks, D. F. (1973). Visual imagery differences in the recall of pictures. *Br. J. Psychol.* 64, 17–24.

McAvinue, L. P., and Robertson, I. H. (2007). Measuring motor imagery ability: a review. *Eur. J. Cogn. Psychol.* 20, 232–251.

McEwan, S. E., Hujbregts, M. P. J., Ryan, J. D., and Polatajko, H. J. (2009). Cognitive strategy use to enhance motor skill acquisition post-stroke: a critical review. *Brain Injury* 23, 263–277.

Meister, I. G., Krings, T., Foltys, H., Boroojerdi, B., Muller, M., Topper, R., et al. (2004). Playing piano in the mind: an fMRI study on music imagery and performance in pianists. *Brain Res. Cogn. Brain Res.* 19, 219–228.

Milton, J., Solodkin, A., Hlustik, P., and Small, S. L. (2007). The mind of expert performance is cool and focused. *Neuroimage* 35, 804–813.

Moran, A. (2002). In the mind's eye. *Psychologist* 15, 414–415.

Moran, A. (2009). Cognitive psychology in sport: progress and prospects. *Psychol. Sport Exerc.* 10, 420–426.

Moran, A., Guillot, A., MacIntyre, T., and Collet, C. (2012). Re-imagining motor imagery: building bridges between cognitive neuroscience and sport psychology. *Br. J. Psychol.* 103, 224–247.

Moran, A. P. (2012). *Sport and Exercise Psychology: A Critical Introduction*. London: Routledge.

Moran, A. P., and MacIntyre, T. (1998). "There's more to an image than meets the eye": a qualitative study of kinaesthetic imagery among elite canoe-slalomists. *Ir. J. Psychol.* 19, 406–423.

Morris, A., and Spittle, A. (2012). Target article: a default hypothesis of the development of internal and external imagery perspectives. *J. Mental Imagery* 36, 1–30.

Morris, T., Spittle, M., and Watt, A. P. (2005). *Imagery in Sport.* Champaign, IL: Human Kinetics.

Munroe, K., Giacobbi, P., Hall, C., and Weinberg, R. (2000). The four W's of imagery use: where, when, why, and what. *Sport Psychol.* 14, 119–137.

Munzert, J., Lorey, J., and Zentgraf, J. (2009). Cognitive motor processes: the role of motor imagery in the study of motor representations. *Brain Res. Rev.* 60, 306–326.

Munzert, J., and Zentgraf, J. (2009). Motor imagery and its implications for understanding the motor system. *Prog. Brain Res.* 174, 219–229.

Neisser, U. (1976). *Cognition and Reality: Principles and Implications of Cognitive Psychology*. New York, NY: WH Freeman.

Neisser, U. (1978). "Memory: What are the important questions?" in *Practical Aspects of Memory*, eds M. M. Gruneburg, P. E. Morris, and R. N. Sykes (London: Academic Press), 3–24.

Nielsen, B. J., and Cohen, L. G. (2008). The Olympic brain. Does corticospinal plasticity play a role in the acquisition of skills required for high performance sports? *J. Physiol.* 586, 65–70.

Nikulin, V. V., Hohlefeld, F. U., Jacobs, A. M., and Curio, G. (2008). Quasi-movements: a novel motor-cognitive phenomena. *Neuropsychologia* 46, 727–742.

Nitschke, J. B., Dixon, G. E., Sarinopoulos, I., Short, S. J., Cohen, J. D., Smith, E. E., et al. (2006). Altering expectancy dampens neural response to aversive taste in primary taste cortex. *Nat. Neurosci.* 9, 435–442.

Nordin, S. M., and Cumming, J. (2005). Professional dances describe their imagery: where when, what, why, and how. *Sport Psychol.* 19, 395–416.

Olson, C. J., and Nyberg, L. (2011). Brain simulation of action may be grounded in physical experience. *Neurocase* 17, 501–505.

Olsson, C. J., Johnsson, B., Larsson, A., and Nyberg, L. (2008). Motor representations and practice affect brain systems underlying imagery: an fMRI study of internal imagery in novices and active high jumpers. *Open Neuroimag. J.* 2, 5–13.

Paivio, A., Yuille, J. C., and Madigan, S. A. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *J. Exp. Psychol.* 76, 1–15.

Pascual-Leone, A., Nguyet, D., Cohen, L. G., Brasil-Neto, J. P., Cammarota, A., and Hallett, M. (1995). Modulation of muscle responses evoked by transcranial magnetic stimulation during the acquisition of new fine motor skills. *J. Neurophysiol.* 74, 1037–1045.

Pearson, D. G., Deeprose, C., Wallace-Hadrill, S. M. A., Heyes, S. B., and Holmes, E. A. (2012). Assessing mental imagery in clinical psychology: a review of imagery measures and a guiding framework. *Clin. Psychol. Rev.* 33, 1–23.

Pearson, J., Clifford, C. W. G., and Tong, F. (2008). The functional impact of mental imagery on conscious perception. *Curr. Biol.* 18, 982–986.

Pearson, J., Rademaker, R., and Tong, F. (2011). Evaluating the mind's eye: the metacognition of visual imagery. *Psychol. Sci.* 22, 1535–1542.

Perky, C. W. (1910). An experimental study of imagination. *Am. J. Psychol.* 21, 422–452.

Pylyshyn, Z. (2003). Explaining mental imagery: now you see it, now you don't. *Trends Cogn. Sci.* 7, 111–112.

Rademaker, R. L., and Pearson, J. (2012). Training visual imagery: improvements of metacognition, but not imagery strength. *Front. Psychol.* 3:224. doi: 10.3389/fpsyg.2012.00224

Richardson, A. (1967a). Mental practice: a review and discussion, Part 1. *Res. Q.* 38, 95–107.

Richardson, A. (1967b). Mental practice: a review and discussion, Part II. *Res. Q.* 38, 263–273.

Roekelein, J. E. (2004). *Imagery and Psychology: A Reference Guide.* Westport, CT: Praeger.

Rosenbaum, D. A. (2005). The Cinderella of psychology: the neglect of motor control in the science of mental life and behavior. *Am. Psychol.* 60, 308–317.

Ross, J. S., Thach, J., Ruggieri, P. M., Lieber, M., and Lapresto, E. (2003). The mind's eye: functional MR imaging evaluation of golf motor imagery. *Am. J. Neuroradiol.* 24, 1036–1044.

Saarliluoma, P., Karlsson, H., Lyytinen, H., Teras, M., and Geisler, F. (2004). Visuospatial representations used by chess experts: a preliminary study. *Eur. J. Cogn. Psychol.* 16, 753–766.

Sarter, M., Berntson, G. G., and Cacioppo, J. T. (1996). Brain imaging and cognitive neuroscience: toward strong inference in attributing function to structure. *Am. Psychol.* 51, 13–21.

Schmuckler, M. A. (2001). What is ecological validity? A dimensional analysis. *Infancy* 2, 419–436.

Senior, C. T., Russell, T., and Gazzaniga, M. S. (2006). *Methods in Mind.* Cambridge, MA: MIT Press.

Sevdalis, N., Moran, A., and Arora, S. (2013). "Mental imagery and mental practice applications in surgery," in *Multisensory Imagery,* eds S. Lacey and R. Lawson (New York, NY: Springer), 343–363.

Shepard, R. N., and Metzler, J. (1971). Mental rotation of three dimensional objects. *Science* 171, 701–703.

Slotnick, S. D., Thompson, W. L., and Kosslyn, S. M. (2005). Visual mental imagery induces retinotopically organized activation of early visual areas. *Cereb. Cortex* 15, 1570–1583.

Smith, E. E., and Kosslyn, S. M. (2007). "Motor cognition and mental simulation," in *Cognitive Psychology: Mind and Brain*, eds E. E. Smith and S. M. Kosslyn (Upper Saddle River, NJ: Prentice Hall), 451–481.

Smyth, M. M., and Waller, A. (1998). Movement imagery in rock climbing: patterns of interference from visual, spatial and kinaesthetic secondary tasks. *Appl. Cogn. Psychol.* 12, 145–157.

Steiger, J. H., and Yuille, J. C. (1983). Long term memory and mental rotation. *Can. J. Psychol.* 37, 367–389.

Stewart, L., and Walsh, V. (2006). "Transcranial magnetic stimulation in cognition," in *Methods in Mind*, eds C. Senior, T. Russell, and M. S. Gazzaniga, (Cambridge, MA: MIT Press), 1–26.

Tartaglia, E. M., Bamert, L., Mast, F. W., and Herzog, M. H. (2009). Human perceptual learning by mental imagery. *Curr. Biol.* 19, 2081–2085.

ter Horst, A. C., Cole, J., van Lier, R., and Steenbergen, B. (2012). The effect of chronic deafferentation on mental imagery: a case study. *PLoS ONE* 7:e42742. doi: 10.1371/journal.pone.0042742

Van den Auweele, Y., Depreeuw, E., Rzewnicki, R., and Ballon, F. (1993). "Elite performance and personality: from description and prediction to diagnosis and intervention," in *Handbook of Research on Sport Psychology,* eds R. N. Singer, M. Murphey, and L. K. Tennant (New York, NY: Macmillan), 257–289.

Vandell, R. A., Davis, R. A., and Clugston, H. A. (1943). The function of mental practice in the acquisition of motor skills. *J. Gen. Psychol.* 29, 243–250.

Wakefield, C. J., Smith, D., Moran, A., and Holmes, P. (2013). Functional equivalence or behavioural matching? A critical reflection on 15 years of research using the PETTLEP model of motor imagery. *Int. Rev. Sport Exerc. Psychol.* 6, 105–121.

Walsh, V., and Cowey, A. (2000). TMS and cognitive neuroscience. *Nat. Rev. Neurosci.* 1, 73–79.

Weinberg, R. S. (2008). Does imagery work? Effects on performance and mental skills. *J. Imagery Res. Sport Phys. Act.* 3, 1–21.

Welfringer, A., Leifert-Fiebach, G., Babinsky, R., and Brandt, T. (2011). Visuomotor imagery as a new tool in the rehabilitation of neglect: a randomised controlled study of feasibility and efficacy. *Disabil. Rehabil.* 33, 2033–2043.

Williams, S. E., and Cumming, J. (2011). Measuring athlete imagery ability: the sport imagery ability questionnaire. *J. Sport Exerc. Psychol.* 33, 416–440.

Wraga, M., and Kosslyn, S. M. (2002). "Imagery," in *Encyclopedia of Cognitive Science, Vol. 2* (London: Nature Publishing Group), 466–470.

Yarrow, K., Brown, P., and Krakauer, J. W. (2009). Inside the brain of an elite athlete: the neural processes that support high achievement in sports. *Nat. Rev. Neurosci.* 10, 585–596.

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: 01 May 2012; accepted: 11 March 2013; published online: 01 April 2013.

Citation: MacIntyre TE, Moran AP, Collet C and Guillot A (2013) An emerging paradigm: a strength-based approach to exploring mental imagery. Front. Hum. Neurosci. 7:104. doi: 10.3389/fnhum.2013.00104

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 04 February 2013
doi: 10.3389/fpsyg.2013.00001

Vividness of visual imagery and incidental recall of verbal cues, when phenomenological availability reflects long-term memory accessibility

Amedeo D'Angiulli*, Matthew Runge, Andrew Faulkner, Jila Zakizadeh, Aldrich Chan and Selvana Morcos

Neuroscience of Imagery, Cognition, and Emotion Research Lab, Carleton University, Ottawa, ON, Canada

Edited by:
Joel Pearson, The University of New South Wales, Australia

Reviewed by:
Joel Pearson, The University of New South Wales, Australia
Giorgio Ganis, Plymouth University, UK

****Correspondence:***
Amedeo D'Angiulli, Neuroscience of Imagery, Cognition, and Emotion Research Lab, Carleton University, 1125 Colonel By Drive, Room 1316 Dunton Tower, Ottawa, ON K1S 5B6, Canada.
e-mail: amedeo@connect.carleton.ca

The relationship between vivid visual mental images and unexpected recall (*incidental recall*) was replicated, refined, and extended. In Experiment 1, participants were asked to generate mental images from imagery-evoking verbal cues (controlled on several verbal properties) and then, on a trial-by-trial basis, rate the vividness of their images; 30 min later, participants were surprised with a task requiring free recall of the cues. Higher vividness ratings predicted better incidental recall of the cues than individual differences (whose effect was modest). Distributional analysis of image latencies through ex-Gaussian modeling showed an inverse relation between vividness and latency. However, recall was unrelated to image latency. The follow-up Experiment 2 showed that the processes underlying trial-by-trial vividness ratings are unrelated to the Vividness of Visual Imagery Questionnaire (VVIQ), as further supported by a meta-analysis of a randomly selected sample of relevant literature. The present findings suggest that vividness may act as an index of availability of long-term sensory traces, playing a non-epiphenomenal role in facilitating the access of those memories.

Keywords: episodic memory, incidental recall, multi-trace theory, visual imagery, vividness, VVIQ

INTRODUCTION

People often report they experience vivid spontaneous visual mental images in situations in which they have to recall something they did not expect to recall (*incidental recall*). Early imagery studies revealed that the spontaneous and involuntary appearance of a vivid visual mental image consistently occurred in response to certain memory conditions and tasks involving incidental recall. For example, upon asking subjects to remember the type of breakfast one had in the morning (Galton, 1880), the number of windows in one's house (Shepard, 1966) or to verify a property of an experienced event with no aid of a current percept (Goldenberg et al., 1992) individuals often report vivid images. In such context, *vividness* is traditionally defined as a construct expressing the self-rated degree of richness, amount of detail (resolution), and clarity of a mental image, as compared to the experience of actual seeing (D'Angiulli and Reeves, 2007). Although vividness correlates with performance on certain memory tasks (Baddeley and Andrade, 2000), with arousal level (Barrowcliff et al., 2004; Bywaters et al., 2004), with positive emotional valence toward a stimulus (Alter and Balcetis, 2010), and with increased visual cortex activity (Farah and Peronnet, 1989; Farah et al., 1989; Sparing et al., 2002; Cui et al., 2007; Cattaneo et al., 2011, 2012), any attempt to clarify its function and its relationship to underlying processes still presents numerous challenges.

Manipulating vividness directly is difficult, and the lack of converging analyses has generally led to the use of correlational approaches that examine vividness predominantly as an index of individual differences in the ability to generate mental images. Furthermore, many preceding studies either confounded vividness with other variables, or did not appropriately interpret the validity criteria by anchoring the vividness construct to models of memory and verbal report underlying processes. This is a situation analogous to the one denounced years ago by Ericsson and Simon (1980) in the context of models of verbal reports, instruments such as vividness ratings/scale/questionnaires seem to be used in a brute empirical fashion, without considering a satisfactory *a priori* theory of the processes involved in the measurement instruments themselves. For the latter reason, it has been argued that there has also been confusion between issues of validity (e.g., discriminant or construct) and issues of reliability (e.g., specificity and precision). In the context of these challenges, the measurement of vividness has been hotly debated. As Pearson (1995) points out, vividness is usually measured using the Vividness of Visual Imagery Questionnaire (VVIQ) or its updated version, the VVIQ2 (Marks, 1995). However, these are not ideal measures for the experimental study of vividness *per se*, as they only measure the overall individual's ability to generate vivid mental images ("trait vividness"), not differences between single experiences of mental imagery ("state vividness"). To study specific processes behind the phenomenon of vividness itself, it is more appropriate to use *trial-by-trial* self-reports in which the vividness of each individual mental image is rated immediately after its generation by the subject (Begg, 1988; Hertzog and Dunlosky, 2006; D'Angiulli, 2009; Pearson et al., 2011). The self reports were successfully employed in several previous studies, where the findings were consistent with both VVIQ research and new results outside the VVIQ's realm of individual differences, which demonstrates that it is a reasonably robust measure (D'Angiulli, 2002, 2008, 2009; D'Angiulli and

Reeves, 2002, 2007; Alter and Balcetis, 2010; Rabin et al., 2010; Pearson et al., 2011). Despite these successes, so far there has been no clear empirical evidence showing exactly why trial-by-trial vividness reports should be considered more informative and reliable than the VVIQ. Do these sets of verbal reports reflect different or overlapping processes?

Many of the mentioned challenges could be mitigated by developing a model of processes underlying trial-by-trial vividness self-reports in visual mental image generation tasks, as opposed to just VVIQ measurement. One of the goals of the model should be to clarify the non-epiphenomenal role of the subjective vividness experience, a fundamental and difficult issue that continues to elude research efforts. An opportunity to gain some upper hand may be offered by conditions in which vivid imagery influences incidental recall in example situations such as the one mentioned earlier. The link between vividness and incidental recall was first suggested long ago (Richardson, 1969; Paivio, 1971) but the best evidence comes from studies showing that self-reported vividness is related with incidental recall of imagery-evoking verbal cues (Sheehan and Neisser, 1969; Sheehan, 1971, 1972b, 1973). In a typical paradigm devised by Sheehan (1972a), "vivid imagers" and "non-vivd imagers," as defined by the VVIQ, were either intentionally or incidentally instructed to recall concrete (high imagery-evoking) and abstract (low imagery-evoking) words. Results showed that vivid imagers recalled concrete words significantly better in the incidental than in the intentional recall condition; whereas recall of abstract words was similarly poor in both conditions.

In another line of research, Neisser and Kerr used objective methods of mnemonic effectiveness and response time to study the spatial properties of visual imagery (Neisser and Kerr, 1973; Keenan and Moore, 1979; Kerr and Neisser, 1983). They asked the subjects to construct images in three different conditions according to presented sentences describing two objects in a given reciprocal spatial relation (concealed, next to/"pictorial," far from/"separate") and measured incidental recall rates of target verbal cues. Visual images acted as mnemonics in the concealed condition as well as the "pictorial" condition. If the procedure changed subtly and intentional learning was used instead, the objects in the concealed condition were recalled no better than the separate condition. The data from these experiments also showed that concealed images were less vivid than pictorial images, and response time was longer for less vivid images. Although instruction for imagery/recall had an effect on imagery vividness, incidental recall was invariably found to predict vividness even in studies that attempted to falsify Neisser and Kerr's findings (Keenan, 1983).

The association between vividness and incidental recall is a relatively consistent finding across several different conditions and manipulations, and suggests that incidental recall could be used as the benchmark variable against which alternative hypotheses on the nature of imagery vividness and its function could be compared. Because older research had several shortcomings, Experiment 1 was designed to replicate, generalize, and extend said relationship. Most of those studies used global or delayed self-report of vividness. In addition, image generation time was confounded with vividness, and most paradigms did not clearly show whether the observed effects were discriminatively and specifically linked to recall processes (refer to Sheehan, 1973, for one exception). Furthermore, individual differences were often globally defined by the VVIQ, such that "good" versus "poor" imagers determined "high" versus "low" vividness, respectively. Finally, the lack of control for factors relating to the cued words themselves was a consistent problem in previous research. In the present research, a direct imagery and incidental recall paradigm were used, and several verbal properties were controlled for (age of word acquisition, word familiarity/frequency, imageability, and concreteness).

We compared two hypothetical cognitive components of mental image generation from verbal descriptions, which possibly could account for the outcomes of Experiment 1. If the relationship between vividness and unexpected recall were contingent upon shared processing while encoding the cues in the study phase (image generation), a possible relationship may be explained by *depth of elaboration* (Craik and Lockhart, 1972; Eysenck and Eysenck, 1980). The more time spent elaborating the imagined material, the more subjectively vivid the material should be. Subsequently, this should lead to better retention and recall in the test phase (free incidental recall). The main predictions derived from this hypothesis were that: (1) a direct relationship between image latency and incidental recall should exist, as should a relationship between incidental recall and self-rated vividness; (2) however, the correlation between vividness and incidental recall should be accounted for by image latency. Therefore, the correlation between vividness and incidental recall should be non-significant and/or correspond to a small effect size when image latency would be controlled for.

A possible alternative based on neurocognitive considerations is that vividness ratings rely on an index of the availability of multiple sensory traces in long-term memory (Hintzman and Block, 1971). Thus, because the strength of vividness would reflect the magnitude of the networks of sensory traces consolidated from episodic memory (Morris and Hampson, 1983; Rabbitt and Winthorpe, 1988), higher vividness ratings should be associated with better incidental recall performance (higher likelihood of accessing long-term traces). This model would also predict that the relationship between vividness and incidental recall can be partly explained by individual differences in participants' ability to access long-term memory sensory information based on the prior estimate of availability supported by vividness judgments. The latter aspect could be conceived as a "meta-imagery" contribution, where the vividness judgment may reflect "a judgment of the richness of the current image combined with an estimate of the additional sensory information that could be incorporated, should the task requirements change." (Baddeley and Andrade, 2000; p. 141). Consequently, individuals with greater metacognitive ability should experience more vivid images, be more efficient and faster in generating images, and yield higher incidental recall than the individuals who possess a reduced metacognitive ability. If greater vividness were related to greater incidental recall accuracy, and the relationship was not simply due to longer image latencies, then this would support the hypothesis that vividness acts as an index of stored memory trace availability, and plays a non-epiphenomenal role in determining the likelihood of accessing such memories in long-term memory.

In all the following experiments, explicit instructions to generate mental images was adopted as this manipulation has proven to be perhaps the most reliable and most direct way to ensure that participants are actually generating mental images, as shown by converging evidence from hundreds of studies showing that the report of having an image at request is associated with behavioral, neural, or clinical neuropsychological indices. In addition, while direct interference of imagery on low-level perception is an established phenomenon (Craver-Lemley and Reeves, 1987), the opposite effect, direct interference of low-level perception on imagery, is either weak and ubiquitous (see D'Angiulli, 2002) or is based again on introspective reports (as in Baddeley and Andrade, 2000). Therefore, the latter manipulations are no better or different than the ones we used for verifying the employment of imagery.

EXPERIMENT 1

METHOD AND MATERIALS

Participants

Participants were 26 first-year university students *age range* = 17–25; 14 female and 12 male). None had participated in an imagery study before (Campos et al., 2007). Participants signed up through a subject pool within 3 weeks of beginning introductory psychology courses, with 2% credit toward their final grade used as incentive. No significance was found for gender or age against any factors, so these variables were dropped from further consideration.

Stimuli

A body of 60 verbal description-cues from previous research (D'Angiulli and Reeves, 2002; available in D'Angiulli, 2001a) were matched with regards to noun or compound word frequency, imageability, concreteness, and reading time. These cues included single-noun and double-noun descriptions comprising both animate (e.g., dog, cat) and inanimate objects (e.g., car, bottle). The present data showed no significant differences between the two subsets of stimuli in terms of vividness or latency of elicited imagery. Secondary analyses indicated that these descriptions were rated as emotionally neutral, with negligible inter-item variability along a simple emotional rating scale (D'Angiulli, 2001b). In addition, the 10 noun-cues were selected from earlier research (Paivio et al., 1968) to use as buffer items during the incidental recall phase of the experiment (i.e., to filter out recency and primacy effects during recall). The 60 cues were presented in random order, preceded by five buffer noun-cues and followed by five other buffer noun-cues (which were presented in a fixed order).

Stimuli properties previously shown to intercorrelate were controlled for. Verbal cues with higher concreteness levels were shown to be recalled at significantly higher rates (Paivio, 1971), as were high frequency words (e.g., Miller and Roodenrys, 2009). Imageability, which refers to how easily a mental image can be generated from a word, has been correlated with concreteness (Tse and Altarriba, 2007). Age of acquisition, which refers to the average age a word enters a subject's lexicon was indirectly controlled for, as it is highly correlated with both imageability (Ma et al., 2009) and concreteness (Barry and Gerhand, 2003). The well-validated MRC Psycholinguistic Database (Clark, 1997) was used to ensure the words used for cuing had approximately the same scores on these factors. Because it was assumed that vividness is an image-specific process, and it could not be rated if an image does not reach to conscious awareness, all cases rated "no image" were eliminated from our analysis.

PROCEDURE

The protocol for Experiment 1 was approved by the Carleton University Research Ethics Board.

Image generation phase

Participants were seated facing a computer monitor and pressed the right mouse button to begin each trial. Upon clicking the mouse, an alerting beep was sounded, followed 250 ms later by the display of a noun-cue at the center of the screen. Participants were instructed to read the cue silently and as quickly as possible. They were immediately asked to generate an image that corresponded to the noun-cue. Participants were required to press the right mouse button again when they considered their image to be complete, and at its most vivid.

Upon pressing the button, another alerting beep was sounded, followed 250 ms later by a horizontal array of seven choices appearing near the bottom of the screen. From left to right, each button was labeled with one of seven vividness level descriptions in a seven-point scale format [(1), "no image"; (2), "very vague/dim"; (3), "vague/dim"; (4), "not vivid"; (5), "moderately vivid"; (6), "very vivid"; and (7), "perfectly vivid"], as in Marks (1995). Time was taken to familiarize participants with the rating system during pre-test practice sessions. Participants used the mouse to click on one of these seven buttons, and were instructed to rate any failure to generate an image as a "no image." There was no deadline for their response.

Following the vividness response, the array of buttons disappeared and the display reverted back to a screen instructing the participant to click the mouse when they were ready to begin the next trial. In an effort to minimize imagery persistence between trials, stimuli were presented in random order with a minimum inter-trial interval of 5 s (Craver-Lemley and Reeves, 1987). Participants were not informed that latency times were covertly measured. Button presses were justified as a means to signal a complete image, which was ready to be rated, and prompt the appearance of the vividness scale buttons.

Free incidental recall phase

After completing the image generation phase, participants took a 20 min break. Afterward, they were asked to return to the lab to fill out additional paperwork, to receive course credit, and complete the debriefing process. Prior to the image generation phase, participants had not been informed that they would be required to recall any of the stimuli. Upon their return, precisely 30 min from the end of the image generation phase, they were asked to complete the incidental recall task, wherein they were required to recall and record as many of the previously read descriptions as possible.

Each phase of the experiment was exclusively conducted by one of two paid undergraduate research assistants. Both research assistants received training in their module, yet remained naïve to the purposes and hypotheses of the study. Final debriefing was conducted through an exit interview with the principal investigator.

RESULTS AND DISCUSSION

Preliminary analyses were conducted on the empirical distributions of raw response times (RTs) for each level of vividness (except level 1 = "no image"). A total of 1490 valid observations were available after all cases with a rating of "no image" (5% of total trials) were removed. Data were binned using the smallest increment that did not make the histograms appear too irregular. From the initial binning it became apparent that our RT data could be fitted by an ex-Gaussian – that is, the convolution of an exponential with a Gaussian. This ex-Gaussian model has been used successfully in several experimental paradigms (for reviews, see Ratcliff, 1979, 1993; McNicol and Stewart, 1980; Luce, 1986) to fit explicit theoretical distribution functions and to give convenient summary of empirical RT distributions. The assumption of the ex-Gaussian model is that RT is the sum of two other random variables, one distributed as a Gaussian and one distributed as an exponential (Luce, 1986). Previous work (D'Angiulli and Reeves, 2002) has supported the hypothesis that the ex-Gaussian model reflects the time to retrieve images from memory so that "image generation" can be essentially reduced to "retrieving images from memory." Therefore, variations in each of ex-Gaussian parameters across vividness levels could be assumed to describe the core underlying generative processes common to both imagery and incidental recall. The ex-Gaussian model was fitted using a robust regression method due to Hoaglin et al. (1983).

To ensure the ex-Gaussian reflected the shape of the group data, and the shape of the individual data, the model was first vincentized for individual data, and then averaged over vividness levels. Histograms were constructed by pooling the raw RTs from each vividness level over subjects, irrespective of the individual source of the RTs. This method has been used in situations where there are too few trials for single subjects (see Ratcliff, 1979). We verified whether the related observations were serially independent and not autocorrelated for each subject, if so we could assume independence of collective observations (see Neter et al., 1996). In our case, the Durbin–Watson autocorrelation test statistic D clearly exceeded the upper bound in the assessment of each subject [$d_u > 1.62$; $\alpha = 0.05$; $n = 60$; lag = 1] as well as for each vividness level submitted to fitting, thereby showing no autocorrelation.

Table 1 shows the ex-Gaussian fit to the distribution histograms of RTs obtained for each vividness level. For each distribution, the ex-Gaussian fit explained at least 68% of the variance associated with RTs. The general distribution of the vividness data showed the median rating was a value of 4 ("non-vivid"). Examination of each vividness level regressed onto RTs showed both distributions were best summarized by piecewise linear regressions of opposite slope. These data supported a clear split between *vivid* (rating values 5–7) and *non-vivid* (2–4) observations.

The Gaussian of both vivid images (levels 5–7) and less vivid images (levels < 4) are reported in **Figure 1**. Both distributions have comparable standard deviation, as evidenced by the left tail of the distributions. However, the distribution of less vivid images is delayed >500 ms, as evidenced by the shift on the time axis. Consistent with previous findings (D'Angiulli and Reeves, 2002), more vivid images were typically associated with shorter Gaussian latency components than were less vivid images. It is important to point out the enormous variability in the response latencies, and that the relationship between vividness could not be easily guessed by naïve participants. Therefore, it is rather implausible that the observed pattern might be due to response-bias based on an explicit or conscious criterion-shift, or set of decisions, since this would have required the participants to first tacitly simulate the ex-Gaussian model, and then retrofit their responses coherently to the model to produce the observed pattern. Because this would have to be done uniformly by all participants, the variability should have been much more contained than what we observed.

Table 1 | Results of the ex-Gaussian fit to empirical image latency distributions in unconstrained image generation phase of Experiment 1 (see text for details).

Vividness	λ	μ	r^2	MRT	SDRT	N
2	18046.0	1500.0	0.68	19547.2	26055.7	17
3	8807.0	2500.0	0.95	11307.2	14520.5	31
4	18641.0	5000.0	0.95	23641.2	28188.5	64
5	12027.0	2500.0	0.96	14527.3	16280.8	174
6	11612.0	2500.0	0.96	14112.3	15416.4	328
7	8162.0	5000.0	0.99	12162.0	18938.2	449

All values reported in the table – except the ones corresponding to n's – are in ms; $500 < \sigma < 1000$. Robust regression with Ramsay's weighting function.

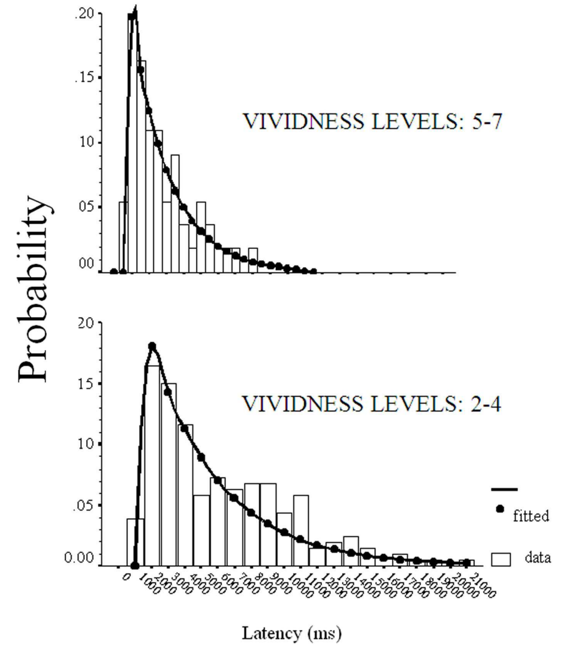


FIGURE 1 | Ex-Gaussian model-fit of RT distributions for images rated with vividness 2–4 and 5–7 in Experiment 1.

The key analysis examined the predictability of recall and RTs from vividness rating category (non-vivid versus vivid). In an effort to meet assumptions for parametric procedures and augment robustness to violations, the distribution of RTs was normalized with a logarithmic transformation, after which no

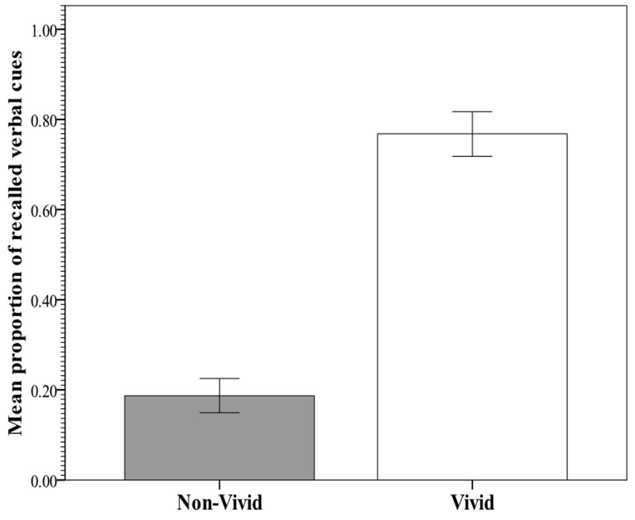


FIGURE 2 | Mean proportion of incidental recall for verbal cues corresponding to vivid (rating: 5–7) and non-vivid (rating: 2–4) images in Experiment 1.

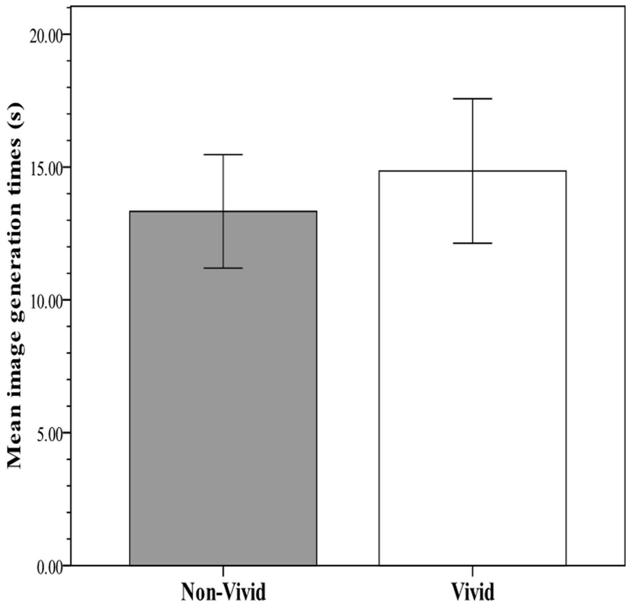


FIGURE 3 | Mean image generation time for verbal cues corresponding to vivid (rating: 5–7) and non-vivid (rating: 2–4) images in Experiment 1.

multivariate outliers were detected. **Figure 2** shows the within-subject mean proportion of incidentally recalled imagery-evoking verbal cues presented during the image generation phase against the rated vividness level. **Figure 3** shows the within-subjects mean RTs of image generation against the rated vividness level (for presentation, RT data are expressed as *seconds*, derived from antilog transformation). The proportion of recalled cues corresponding to vivid images was 0.77 (SE = 0.05), whereas the proportion of recalled cues corresponding to non-vivid images was 0.19 (SE = 0.04). A paired samples test showed the difference to be significant [$t(25) = 6.69; p < 0.0001$], explaining 74% of the variance. In contrast, the mean RTs for vivid (14.8 s, SE = 2.71) and non-vivid (13.33, SE = 2.14) cues did not differ [$t(25) < 1$, $p = 0.34$; $R^2 < 0.01$].

A linear regression analysis examining the effect of individual differences on the total number of images recalled showed that 14% of the variance in incidental recall accuracy was explained by participants' average vividness rating [$F(1, 25) = 4.05, \mathrm{MS}_e = 0.38$, $p = 0.05$]. Therefore, the role of individual differences was modest and its effect size (r) was significantly smaller than that of vividness described earlier (0.86 versus 0.37, $z = 3.07$, $p = 0.002$).

A two-predictor model (stimulus and vividness) was fitted to the data to test the hypothesis regarding relationship between vividness and recall. Stimulus was plotted as a nominal factor, in which each category was a noun-cue. It was included as a predictor to ensure vividness effects were not due to the tendency for some words to produce more vivid images than others. The resulting model [*Predicted logit of (Recall)* $= 0.664 + \beta 1^{*\mathrm{Vividness}} + \beta 2^{*\mathrm{Stimuli}}$] was statistically reliable, $\chi^2(62, 1441) = 340.969, p < 0.001$ (see Appendix A for analysis details). According to the model, greater vividness ratings for noun-cues predicted recall with an overall success rate of 72.2%. The model correctly classified 83.7% of unrecalled cues and 54.7% of recalled cues. Stimulus and vividness generate a statistically significant predictive model for recall (see Appendix) that accounted for 28.3% of the variance in incidental recall. No change was observed if the model was fit to predict recall when response time was added as a predictor [$\chi^2(65, 1490) = 389.437$, $p < 0.001$]. RT did not exert an influence on the model ($B = -0.002$, $p = 0.587$), which further confirmed the null effect of RT on recall. Therefore, vividness could not account for recall accuracy simply because participants spent more time imagining the items corresponding to the verbal cues.

A linear mixed model was fit to the data to assess the contribution of stimulus and RTs to linear change in vividness of imagery. The variables in the model were evaluated by a Type III test. Since the sample size was not large, Restricted Iterative Generalized Least -squares (RIGLS) was used (Goldstein, 1986). Stimuli and RT had a significant effect on vividness [$F(59, 1000) = 1.59$, $\eta^2 = 0.086$, $p < 0.05$], and $F(1, 1103) = 5.17, \eta^2 = 0.005, p < 0.05$, respectively. Therefore, because the effects were small, RTs and stimuli influenced vividness only minimally. There was no interaction between stimuli and RTs ($F < 1$).

To determine if recall and vividness ratings were affected by the verbal properties of the word stimuli that were not kept constant during stimulus selection, correlation analyses were conducted on age of acquisition, and familiarity versus recall. No significant relationship was found between the percentage of participants that recalled a cue, and either age of acquisition ($r = 0.213$, $p = 0.317$) or familiarity scores ($r = 0.118$, $p = 0.445$). In addition, effects of stimuli regressed onto vividness, recall and RTs all explained less than 0.5% of the variance.

The results of Experiment 1 implicate vividness ratings as a predictor of incidental recall for imagery-evoking cues. The effect of individual differences in imaging ability on incidental recall was much smaller than the effect of vividness. Because image latency was unrelated to incidental recall, and inversely related to vividness, these data were incompatible with the depth of elaboration account. Because the effect of vividness on incidental recall for verbal cues was tested, the influence of expectancy and demand characteristics were minimized. These results support the validity of vividness as a measurable construct, and as an entity which may represent real underlying memory processes. Vividness ratings likely reflect a process which provides a natural mnemonic for unexpected retrieval of implicitly coded information (see Kosslyn et al., 2006).

EXPERIMENT 2

Although Experiment 1 did not include a measure of VVIQ, incidental and intentional recall has traditionally shown a modest correlation with the VVIQ and VVIQ2, with average effect sizes generally of about $r = 0.13$ (see McKelvie, 1995; Dean and Morris, 2003). More recent evidence suggests the relationship between VVIQ2 and trial-by-trial vividness ratings is weak to moderate ($r < 0.20$) (D'Angiulli, 2001a; D'Angiulli and Reeves, 2007). Also, the patterns of results from Sheehan (1971, 1972b) suggest the quality of imagery is contingent upon properties of the stimuli within the setting of each trial, and predicts incidental free recall and recognition performance. Lastly, other studies found the modest correlation between trial-by-trial ratings and VVIQ holds only for female participants (Sheehan, 1971, 1973).

In contrast with these findings, Pearson et al. (2011) reported large predictive effects of both trial-by-trial vividness ratings and VVIQ2 scores when related to bias in reporting a dominant pattern during a binocular rivalry task. The underlying assumption was that similar metacognitive processes (i.e., knowing how and what the observer knows about his/her own processes of visual mental imagery) would be used in trial-by-trial vividness ratings and in VVIQ2. If this assumption is correct, the overlapping processes could shed some light on the results of our Experiment 1. One interpretation of the results of Experiment 1 is that trial-by-trial vividness ratings may be accounted for by the same metacognitive judgment processes involved in responding to the VVIQ2. Experiment 2 was designed to examine the putative relationship between vividness ratings and VVIQ2. If the association between the VVIQ2 and vividness ratings were confirmed in Experiment 2, then one may also explain the basis through which vividness ratings could predict incidental recall in terms of the overlapping metacognitive processes involved in the VVIQ2.

The design of Experiment 2 was a variation of the paradigm used by Baddeley and Andrade (2000). Upon completing the VVIQ2, female participants were asked to read a short description of a static or dynamic scene, and press a key upon generating complete visual mental image. Participants then rated the vividness and the subjectively perceived latency of the image on a trial-by-trial basis. If, as the results of Experiment 1 would suggest, vividness ratings are based on an index of multiple sensory traces available in long-term memory, this account would predict: (1) higher trial-by-trial vividness ratings for dynamic scenes than static scenes, and (2) a negative (i.e., inverse) relationship between trial-by-trial vividness and perceived imagery latency. The VVIQ2 should correlate with trial-by-trial vividness ratings from both dynamic and static scenes, but should not relate to perceived imagery latency when the effects of vividness are removed.

Conversely, if the VVIQ2 accounts for most of the relationship between trial-by-trial vividness ratings and perceived imagery latency, then vividness judgments could be attributed to similar individual metacognitive skill differences involved in the two types of vividness measures (Baddeley and Andrade, 2000; Pearson et al., 2011). However, because dynamic mental imagery capacitates working memory more than static mental imagery, fewer resources are available for concurrent metacognitive processes. Then, under such circumstances one would expect less vivid images for dynamic scenes than static ones.

METHODS AND MATERIALS

Participants

Participants were 44 female undergraduate students (*age range*: 18–25). Participants signed up through a subject pool, with 2% credit toward their final grade used as incentive. All participants had normal or corrected-to-normal vision, and no reported or documented learning disabilities. Participation required the attendance of two appointments. The first appointment was a preliminary screening session, where participants filled out the VVIQ2 and individual data. The second appointment was the experimental session. Five potential participants were excluded from the experiment, as they were unable to evoke the images as required.

Materials

An adaptation of 17 static and 17 dynamic scene descriptions were used (Baddeley and Andrade, 2000; Experiment 4, see Appendix A, p. 144). The scenes were adapted such that words including British content (e.g., Big Ben) were substituted with equally long words describing North American content (e.g., CNN Tower) which were validated through pilot experiments. During the screening phase, a question from the visual portion of the *procedure for assessing expectations on the vividness of imagery* was asked (Baddeley and Andrade, 2000; see Appendix C, Q2, Question 2, p. 145). After the experimental phase, a tacit knowledge assessment procedure was administered.

PROCEDURE

The protocol for Experiment 2 was approved by the Carleton University Research Ethics Board.

Participants were given instructions, and 10 min of practice with five dynamic and five static imagery scenes. Between each practice trial, participants were required to report how well they could control each image. Only participant ratings with vividness greater than "extremely slow" (1) for 80% of the practice trials qualified for the entire experiment. One participant was eliminated from the initial pool under such criteria. Upon completing the practice session, participants verbally repeated the instructions to the experimenter to ensure the instructions were understood.

Participants were instructed to silently read a description of a dynamic or static scene displayed on a computer screen, which occurred 250 ms after an alerting beep. The experiment consisted of 17 dynamic, and 17 static descriptions. Participants were tested individually, and the procedure lasted approximately 40 min. Upon reading each description, participants were required to press a key to indicate the description was understood. Participants were instructed to imagine the description with their eyes open, and as seen from the front. Outline drawings were shown as examples before the experiment began. Each description was presented in random order with an inter-trial interval of 5 s (Craver-Lemley and Reeves, 1987). Upon forming a complete mental image, participants were required to press a button on a mouse. Four seconds after the button press, participants were shown buttons to rate perceived vividness, and perceived latency of the images. Participants were asked to rate their image as "complete" or "finished" when the image was maximally clear and detailed (see Cocude and Denis, 1988). Participants were required to rate their mental image as they had experienced it at the time of the key press. There was no deadline for the rating responses.

The presentation order of the scales was randomized, such that vividness could follow or precede perceived imagery rating. The second rating task followed immediately after the first rating response. The vividness scale consisted of a horizontal array of seven buttons appearing at the center of the screen. From left to right, each button was labeled with a short description corresponding to one of seven levels of the vividness scale used in Experiment 1. The imagery latency (speed) scale consisted of a horizontal array of seven buttons appearing at the center of the screen. From left to right, each button was labeled with a short description corresponding to one of the seven levels: from "extremely fast" (7), to "extremely slow" (1). Valid trials were defined by vividness greater than 1. Subjects were instructed to give a "1" response if they were unable to form a mental image. Upon completing the experiment, participants underwent a post-experimental interview, wherein they quickly described what they had imagined for seven randomly probed descriptions from both dynamic and static condition. Post-experimental interviews were concluded with the tacit knowledge assessment procedure (Baddeley and Andrade, 2000), and included the following question:

> "We are interested in knowing if you think that there was a relationship between how vivid your images were and other factors. Please just tell us what you expect or think, please do not use images to answer the question, we are just interested in what you predict or think about things that may be related or may determine the vividness of your images."

RESULTS AND DISCUSSION

To eliminate effects of discrepant scales, total scores for the VVIQ2 were converted to mean vividness values through a simple linear transformation. The transformation resulted in a seven-point scale; henceforth, referred to as *mean vviq2*. As in Experiment 1, we considered only valid responses. The rate of excluded invalid trials was approximately 3% (level 1 = "no image"), a proportion similar to Experiment 1. On average, images were reported as moderately vivid, and were produced at a relatively fast perceived latency in both static ($M_{\text{viv.}} = 5.31$, $\text{SD}_{\text{viv.}} = 0.55$; $M_{\text{speed}} = 5.40$, $\text{SD}_{\text{speed}} = 0.39$) and dynamic ($M_{\text{viv.}} = 5.29$, $\text{SD} = 0.77$; $M_{\text{speed}} = 5.59$, $\text{SD}_{\text{speed}} = 0.62$) conditions. Paired samples contrasts showed dynamic imagery was perceived as significantly faster than static imagery [$t(38) = 2.52$, $p < 0.025$]. However, mean vividness ratings did not differ between the two conditions [$\text{t}(38) < 1, p = 0.797$]. The latter result differed from Baddeley and Andrade's findings (Experiment 4). (They found dynamic imagery was significantly less vivid than static imagery). Images produced for the VVIQ2 were significantly more vivid ($M_{\text{vviq2}} = 5.68$, $\text{SD}_{\text{vviq2}} = 0.62$) than vividness for static images [$t(38) = 3.88$, $p < 0.0001$], and dynamic images [$t(38) = 2.83$, $p < 0.01$]. These data may be interpreted as evidence that participants were generally much more confident in their imagery abilities than what they were capable of demonstrating during the experimental procedure. The discrepancy between trial-by-trial vividness level and VVIQ2 imply a lack of agreement between metacognitive judgment as measured through the VVIQ2, and verbal reports specific to the actual imagery task.

Table 2 shows correlations among all measures. VVIQ2 was significantly correlated with vividness of static imagery, but was not related to vividness of dynamic imagery, nor perceived latency in both static and dynamic imagery conditions. A very strong inverse relationship between trial-by-trial vividness ratings and perceived imagery latency was observed in both static and dynamic imagery conditions, with strong to marginal evidence of the same trends in crossed conditions.

Whereas vividness ratings correlated with perceived latency, the VVIQ2 did not. These data provide very weak evidence validating the VVIQ2, when the criterion is a self-report, subjective third variable. Logically, one would not expect any predictive success of VVIQ2 in relation to a behavioral variable such as incidental recall. The observed patterns were analyzed to determine if they could be predicted by expectations or tacit knowledge (Pylyshyn, 2003). There was no significant difference in the number of participants expecting vivid imagery to be less or more vivid than static imagery ($\chi^2 < 1$). **Figure 4** describes participant responses concerning self-rated predictions about the type of relationship they expect to exist between perceived vividness and perceived imagery latency, as documented during the preliminary screening session. Most participants predicted a positive relationship, or no relationship between vividness and imagery latency. One participant correctly predicted the inverse relationship. Upon removing the data of this participant from the analysis, there were no significant differences between results.

In conclusion, the association between the VVIQ2 and vividness ratings was not observed consistently in both the conditions of Experiment 2, and if collapsed across conditions (static and dynamic) the effect becomes modest and not significant. VVIQ2

Table 2 | Correlation matrix among VVIQ2 and self-reported image vividness ratings and perceived generation speed in dynamic and static imagery conditions of Experiment 2.

	Dyn. vividness	Stat. vividness	Dyn. speed	Stat. speed
VVIQ2	0.259	0.505^{**}	0.044	−0.219
Dyn. vividness	–	0.679^{**}	-0.626^{**}	-0.652^{**}
Stat. vividness	–	–	$-0.282^{\dagger}$	-0.531^{**}
Dyn. speed	–	–	–	0.622^{**}

Dyn., dynamic imagery condition; Stat., static imagery condition. N = 39.
$^{\dagger}p < 0.10$, $^{**}p < 0.01$.

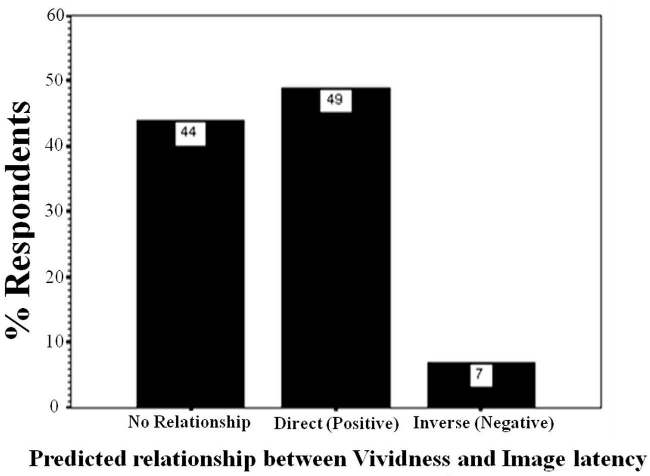


FIGURE 4 | Percentages of participants predicting what type of relationship they tacitly think there should be between vividness ratings and speed of imagery in Experiment 2.

also failed to validate against a third self-report criterion variable (perceived image latency). If the VVIQ2 assesses individual differences in metacognitive ability, it seems implausible that such abilities would predict incidental recall. Because trial-by-trial vividness predicted incidental recall, the metacognitive aspects assumed to be reflected by VVIQ2 do not appear to influence vividness and the mental imagery process to a significant degree.

GENERAL DISCUSSION

Despite controlling for imageability, concreteness, age of acquisition, and verbal frequency/familiarity, the results from Experiment 1 showed a positive relationship between vividness ratings and incidental recall of imagery-evoking cues. These results are not consistent with depth of elaboration, as faster image generation latencies accompanied higher vividness ratings, a pattern opposite to what depth of elaboration would predict. Furthermore, because depth of elaboration predicts a positive correlation between incidental recall and image generation time, it again fails to account for the data from Experiment 1.

Our findings are compatible with an alternative model of vividness processes based on multi-trace memory theory (MMT; Moscovitch et al., 2005). This model proposes that vividness ratings are based on an index of the availability of multiple sensory traces in long-term memory, the strength of vividness reflecting the magnitude of the networks of sensory traces that have been consolidated from episodic memory. This is described by the inverse relationship between vividness ratings and image latency (the "vivid-is-fast" relation). Thus, higher vividness ratings are associated with higher likelihood of incidental recall, as shown by the data of Experiment 1.

The follow-up results observed in Experiment 2 showed that individual differences, as measured by the VVIQ2, are not a viable account for the relationship in Experiment 1 between vividness and incidental recall. Most important, the results of Experiment 2 also suggest that if there were metacognitive aspects involved in trial-by-trial vividness ratings, they would not likely be the same ones underlying VVIQ measures. Taken together the results of Experiment 1 and Experiment 2 are consistent with those observed in a meta-analysis we conducted, representing 5% of the literature pertaining to "vividness" and "VVIQ" (reported in Appendix B). The proportion of significant and non-significant experimental outcomes for trial-by-trial vividness ratings and VVIQ factor effects were calculated. For behavioral, cognitive, and neural measures, a greater number of significant experimental outcomes accompanied trial-by-trial vividness ratings than the VVIQ. Furthermore, the correlation between VVIQ scores and trial-by-trial vividness ratings for 21 entries showed an average correlation of 0.15, and variability in these values ranged from $r = -0.27$, to $r = 0.64$. Consistent with the results of experiment 2, these additional results support the contention that trial-by-trial vividness self-reports and VVIQ scores share some descriptive properties of visual imagery. However, trial-by-trial vividness ratings seem to resolve the construct of mental imagery with much greater reliability. Although metacognitive processes may be occurring in single trial judgment, it is perhaps more parsimonious to assume that vividness ratings are mostly a form of *Level 2 retrospective verbal reports* (Ericsson and Simon, 1993).

Considered as retrospective verbal reports, vividness ratings may be based on a direct translation of residual top-down sensory traces available in long-term memory (D'Angiulli and Reeves, 2002), wherein vividness intensity is proportional to the magnitude of sensory traces available. This statement agrees with a number of neurocognitive considerations borne out of MMT research. According to that theoretical framework, each sensory trace is distributed across the cortex, such that various distributive patterns are unique to a specific sensory input, and is distinct from all other distributive patterns (Hintzman, 1976). Sensory traces are thought to be indexed by the hippocampus (Ryan et al., 2001), and integrated into a mental image by the cuneus, precuneus, and occipital lobes (Svoboda et al., 2006; Svoboda and Levine, 2009; Cabeza and St. Jacques, 2007). However, hippocampal indexing becomes less influential as each individual sensory trace is integrated into cortical networks through successive (re)presentations (Takashima et al., 2009). Mental images are consolidated neural patterns that correspond to these "synthetic" sensory long-term traces, whose levels of interconnectedness are correlated to their perceived reportable vividness (Rabin et al., 2010).

Our study also indicates that although the VVIQ or VVIQ2 may very well measure an individual's ability to generate vivid mental images ("trait vividness"), it likely lacks the resolution to measure an individual's ability to experience vivid mental images in specific situational contexts ("state vividness"). To study specific processes behind the phenomenon of vividness itself (rather than "trait vividness"), it is perhaps more appropriate to use trial-by-trial self-reports, wherein vividness is rated immediately after its generation (Begg, 1988; Hertzog and Dunlosky, 2006; D'Angiulli, 2009; Pearson et al., 2011). Such self-reports have met with compounding success progressing beyond the VVIQ's realm of individual differences, while remaining generally consistent with it. Vividness ratings demonstrate the reasonably robust nature of self-reports as a measure of "state" and "trait" vividness (D'Angiulli, 2002; D'Angiulli and Reeves, 2002, 2007;

D'Angiulli, 2009; Alter and Balcetis, 2010; Rabin et al., 2010; Pearson et al., 2011). This particular issue is critical given the recent resurgence of use of the VVIQ in cognitive neuroscience – especially in the realm of neuroimaging (Amedi et al., 2005; Palmiero et al., 2010).

In summary, we found that trial-by-trial vividness ratings predict incidental recall, and the relationship cannot be attributed to depth of elaboration or metacognitive processes related to self-appraisal of individual imagery ability, as measured by the VVIQ2. Our results suggest that vividness of imagery makes implicit information available to consciousness, and to some extent, is linked with the associative processes through which phenomenal availability translates into access of incidental episodic memories. Therefore, we conclude, in certain conditions conscious phenomenological experience associated with imagery does not have a trivial role as it can have a critical influence on recall performance.

ACKNOWLEDGMENTS

We thank Adam Reeves and the late Ulric Neisser for comments on earlier ideas and bits of manuscripts before the complete draft even existed. Many thanks also go to John M. Kennedy, Robert Kunzendorf, J. T. E. Richardson, Stuart McKelvie for comments and pre-reviews of the ideas and data presented here, Judy Hall for meta-analytic advice. Portions of this research were presented at ICOM (International Congress on Memory), University of York, UK, 2012.

REFERENCES

Alter, A. L., and Balcetis, E. (2010). Fondness makes the distance grow shorter: desired locations seem closer because they seem more vivid. *J. Exp. Soc. Psychol.* 47, 16–21.

Amedi, A., Malach, R., and Pascual-Leone, A. (2005). Negative BOLD differentiates visual imagery and perception. *Neuron* 48, 859–872.

Baddeley, A. D., and Andrade, J. (2000). Working memory and the vividness of imagery. *J. Exp. Psychol. Gen.* 129, 126–145.

Barrowcliff, A. L., Gray, N. S., Freeman, T. C. A., and MacCulloch, M. J. (2004). Eye-movements reduce the vividness, emotional valence and electrodermal arousal associated with negative autobiographical memories. *J. Forens. Psychiatry Psychol.* 15, 325–345.

Barry, C., and Gerhand, S. (2003). Both concreteness and age-of-acquisition affect reading accuracy but only concreteness affects comprehension in a deep dyslexic patient. *Brain Lang.* 84, 84–104.

Begg, I. (1988). What does the vividness of an image tell us about the value of imagery? *J. Ment. Imagery* 12, 45–56.

Bywaters, M., Andrade, J., and Turpin, G. (2004). Determinants of the vividness of visual imagery: the effects of delayed recall, stimulus affect and individual differences. *Memory* 12, 479–488.

Cabeza, R., and St. Jacques, P. (2007). Functional neuroimaging of autobiographical memory. *Trends Cogn. Sci. (Regul. Ed.)* 11, 219–227.

Campos, A., Gómez-Juncal, R., and Pérez-Fabello, M. J. (2007). Experience in imagery and imagery vividness. *Imagin. Cogn. Pers.* 27, 337–348.

Cattaneo, Z., Bona, S., and Silvanto, J. (2012). Cross-adaptation combined with TMS reveals a functional overlap between vision and imagery in the early visual cortex. *Neuroimage* 59, 3015–3020.

Cattaneo, Z., Pisoni, A., Papagno, C., and Silvanto, J. (2011). Modulation of visual cortical excitability by working memory: effect of luminance contrast of mental imagery. *Front. Psychol.* 2:29. doi:10.3389/fpsyg.2011.00029

Clark, C. (1997). *MRC Psycholinguistic Database.* The University of Western Australia School of Psychology. Available at: http://websites.psychology.uwa.edu.au/school/MRCDatabase/mrc2.html

Cocude, M., and Denis, M. (1988). Measuring the temporal characteristics of visual images. *J. Ment. Imagery* 12, 89–101.

Craik, F., and Lockhart, R. (1972). Levels of processing: a framework for memory research. *J. Verb. Learn. Verb. Behav.* 11, 671–684.

Craver-Lemley, C., and Reeves, A. (1987). Visual imagery selectively reduces vernier acuity. *Perception* 16, 599–614.

Cui, X., Jeter, C. B., Yang, D., Montague, P. R., and Eagleman, D. M. (2007). Vividness of mental imagery: individual variability can be measured objectively. *Vision Res.* 47, 474–478.

D'Angiulli, A. (2001a). Phenomenal and temporal aspects of visual mental image generation: validating verbal reports on vividness through latency analysis. *Diss. Abstr. Int.* 61, 5015.

D'Angiulli, A. (2001b). Vividness vs. imageability: effects on word frequency and image latency. *Northwest Cognition and Memory, 3rd Annual Meeting*, May 26, Vancouver, BC.

D'Angiulli, A. (2002). Mental image generation and the contrast sensitivity function. *Cognition* 85, B11–B19.

D'Angiulli, A. (2008). Is the spotlight an obsolete metaphor of "seeing with the mind's eye"? A constructive naturalistic approach to the inspection of visual mental images. *Imagin. Cogn. Pers.* 28, 117–135.

D'Angiulli, A. (2009). "Vividness and behavioral-specificity in imagery: not what you'd expect," in *Proceedings of the 31st Annual Meeting of The Cognitive Science Society: Symmetry & Vividness in Vision*, eds N. Taatgen and H. van Rijn (Amsterdam, NL: The Cognitive Science Society), 49.

D'Angiulli, A., and Reeves, A. (2002). Generating visual mental images: latency and vividness are inversely related. *Mem. Cognit.* 30, 1179–1188.

D'Angiulli, A., and Reeves, A. (2007). The relationship between self-reported vividness and latency during mental size scaling of everyday items: phenomenological evidence of different types of imagery. *Am. J. Psychol.* 120, 521–551.

Dean, G. M., and Morris, P. E. (2003). The relationship between self-reports of imagery and spatial ability. *Br. J. Psychol.* 94, 245–273.

Ericsson, K. A., and Simon, H. A. (1980). Verbal reports as data. *Psychol. Rev.* 87, 215–251.

Ericsson, K. A., and Simon, H. A. (1993). *Protocol Analysis: Verbal Reports as Data.* Cambridge, MA: The MIT Press.

Eysenck, M. W., and Eysenck, M. C. (1980). Effects of processing depth, distinctiveness, and word frequency on retention. *Br. J. Psychol.* 71, 263–274.

Farah, M. J., and Peronnet, F. (1989). Event-related potentials in the study of mental imagery. *J. Psychophysiol.* 3, 99–109.

Farah, M. J., Weisberg, L. L., Monheit, M. A., and Peronnet, F. (1989). Brain activity underlying mental imagery: event-related potentials during mental image generation. *J. Cogn. Neurosci.* 1, 302–316.

Galton, F. (1880). Statistics of mental imagery. *Mind* 5, 301–318.

Goldenberg, G., Steiner, M., Podreka, I., and Deecke, L. (1992). Regional cerebral blood flow patterns related to verification of high and low imagery sentences. *Neuropsychologia* 30, 1081–1092.

Goldstein, H. (1986). Multilevel mixed linear model analysis using iterative generalized least square. *Biometrika* 73, 43–56.

Hertzog, C., and Dunlosky, J. (2006). "Using visual imagery as a mnemonic for verbal associative learning," in *Imagery and Spatial Cognition: Methods, Models, and Cognitive Assessment*, eds T. Vecchi and G. Bottini (Philadelphia, PA: John Benjamins), 259–278.

Hintzman, D. (1976). "Repetition and memory," in *The Psychology of Learning and Motivation*, ed. G. Bower (New York, NY: Academic Press), 47–91.

Hintzman, D. L., and Block, R. A. (1971). Repetition and memory: evidence for a multi-trace hypothesis. *J. Exp. Psychol.* 88, 297–306.

Hoaglin, D. C., Mosteller, F., and Tukey, J. W. (1983). *Understanding Robust and Exploratory Data Analysis.* New York: Wiley.

Keenan, J. M. (1983). Qualifications and clarifications of images of concealed objects: a reply to Kerr and Neisser. *J. Exp. Psychol. Learn. Mem. Cogn.* 9, 222–230.

Keenan, J. M., and Moore, R. E. (1979). Memory for images of concealed objects: a reexamination of Neisser and Kerr. *J. Exp. Psychol. Learn. Mem. Cogn.* 5, 374–385.

Kerr, N. H., and Neisser, U. (1983). Mental images of concealed objects-new evidence. *J. Exp. Psychol. Learn. Mem. Cogn.* 9. 212–221.

Kosslyn, S. M., Thompson, W., and Ganis, G. (2006). *The Case for Mental Imagery*. New York, NY: Oxford University Press.

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

Ma, W., Golinkoff, R., Hirsh-Pasek, K., McDonough, C., and Tardif, T. (2009). Imageability predicts the age of acquisition of verbs in Chinese children. *J. Child Lang.* 36, 405–423.

Marks, D. F. (1995). New directions for mental imagery research. *J. Ment. Imagery* 19, 153–167.

McKelvie, S. J. (1995). *Vividness of Visual Imagery: Measurement, Nature, Function & Dynamics*. New York, NY: Brandon House.

McNicol, D., and Stewart, G. W. (1980). "Reaction time and the study of memory," in *Reaction Times*, ed. A. T. Welford (London: Academic Press), 253–307.

Miller, L., and Roodenrys, S. (2009). The interaction of word frequency and concreteness in immediate serial recall. *Mem. Cognit.* 37, 850–865.

Morris, P. E., and Hampson, P. J. (1983). *Imagery and Consciousness*. London: Academic Press.

Moscovitch, M., Rosenbaum, R. S., Gilboa, A., Addis, D. R., Westmacott, R., Grady, C., et al. (2005). Functional neuroanatomy of remote episodic, semantic and spatial memory: a unified account based on multiple trace theory. *J. Anat.* 207, 35–66.

Mosteller, F., and Bush, R. R. (1954). "Selected quantitative techniques," in *Handbook of Social Psychology*, ed. G. Lindzey (Cambridge, MA: Addison-Wesley), 289–334.

Neisser, U., and Kerr, N. (1973). Spatial and mnemonic properties of visual images. *Cogn. Psychol.* 5, 138–150.

Neter, J., Kutner, M. H., Nachtsheim, C. J., and Wasserman, W. (1996). *Applied Linear Statistical Models*, 4th Edn. New York: McGraw-Hill/Irwin Series: Operations and Decision Sciences.

Paivio, A. (1971). *Imagery and Verbal Processes*. New York: Holt, Rinehart and Winston.

Paivio, A., Yuille, J. C., and Madigan, S. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *J. Exp. Psychol.* 76, 1–25.

Palmiero, M., Nakatani, C., Raver, D., Belardinelli, M., and Leeuwen, C. (2010). Abilities within and across visual and verbal domains: how specific is their influence on creativity? *Creat. Res. J.* 22, 369–377.

Pearson, D. G. (1995). The VVIQ and cognitive models of imagery: future directions for research. *J. Ment. Imagery* 19, 167–170.

Pearson, J., Rademaker, R. L., and Tong, F. (2011). Evaluating the mind's eye: the metacognition of visual imagery. *Psychol. Sci.* 22, 1535–1542.

Pylyshyn, Z. W. (2003). *Seeing and Visualizing: It's not What you Think*. Cambridge, MA: MIT Press.

Rabbitt, P., and Winthorpe, C. (1988). "What do old people remember? The Galton paradigm reconsidered," in *Practical Aspects of Memory: Current Research and Issues*, eds M. M. Gruneberg, P. E. Morris, and R. N. Sykes (New York: Wiley), 301–307.

Rabin, J. S., Gilboa, A., Stuss, D. T., Mar, R. A., and Rosenbaum, R. S. (2010). Common and unique neural correlates of autobiographical memory and theory of mind. *J. Cogn. Neurosci.* 22, 1095–1111.

Ratcliff, R. (1979). Group reaction time distributions and an analysis of distribution statistics. *Psychol. Bull.* 86, 446–461.

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychol. Bull.* 114, 510–532.

Richardson, A. (1969). *Mental Imagery*. New York: Spring Publishing Company.

Rosenthal, R. (1991). *Meta-Analytic Procedures for Social Research (rev. ed.)*. Newbury Park, CA: Sage.

Ryan, L., Nadel, L., Keil, K., Putnam, K., Schnyer, D., Trouard, T., et al. (2001). Hippocampal complex and retrieval of recent and very remote autobiographical memories: evidence from functional magnetic resonance imaging in neurologically intact people. *Hippocampus* 11, 707–714.

Sheehan, P. W. (1971). The role of imagery in incidental learning. *Br. J. Psychol.* 62, 235–243.

Sheehan, P. W. (1972a). *The Function and Nature of Imagery*. Oxford: Academic Press.

Sheehan, P. W. (1972b). Role of imagery in incidental learning: replication and extension of an effect. *J. Exp. Psychol.* 95, 226–228.

Sheehan, P. W. (1973). Stimulus imagery effect and the role of imagery in incidental learning. *Aust. J. Psychol.* 25, 93–102.

Sheehan, P. W., and Neisser, U. (1969). Some variables affecting the vividness of imagery in recall. *Br. J. Psychol.* 60, 71–80.

Shepard, R. N. (1966). Learning and recall as organization and search. *J. Verb. Learn. Verb. Behav.* 5, 201–204.

Sparing, R., Mottaghy, F. M., Ganis, G., Thompson, W. L., Topper, R., Kosslyn, S. M., et al. (2002). Visual cortex excitability increases during visual mental imagery – a TMS study in healthy human subjects. *Brain Res.* 938, 92–97.

Svoboda, E., and Levine, B. (2009). The effects of rehearsal on the functional neuroanatomy of episodic autobiographical and semantic remembering: a functional magnetic resonance imaging study. *J. Neurosci.* 29, 3073–3082.

Svoboda, E., McKinnon, M., and Levine, B. (2006). The functional neuroanatomy of autobiographical memory: a meta-analysis. *Neuropsychologia* 44, 2189–2208.

Takashima, A., Nieuwenhuis, I. L. C., Jensen, O., Talamini, L. M., Rijpkema, M., and Fernández, G. (2009). Shift from hippocampal to neocortical centered retrieval network with consolidation. *J. Neurosci.* 29, 10087–10093.

Tse, C., and Altarriba, J. (2007). Testing the associative-link hypothesis in immediate serial recall: evidence from word frequency and word image ability effects. *Memory* 15, 675–690.

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: 02 April 2012; accepted: 02 January 2013; published online: 04 February 2013.

*Citation: D'Angiulli A, Runge M, Faulkner A, Zakizadeh J, Chan A and Morcos S (2013) Vividness of visual imagery and incidental recall of verbal cues, when phenomenological availability reflects long-term memory accessibility. Front. Psychology **4**:1. doi: 10.3389/fpsyg.2013.00001*

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

APPENDICES

APPENDIX A

A linear mixed model was fit to the data to assess the contribution of the variables to linear change in vividness of imagery. The analysis was carried out using SPSS17. The variables in the model were evaluated by a Type III test. Since the sample size was not large, Restricted Iterative Generalized Least Square (RIGLS) estimation method was used. **Table A1** shows the type III tests of fixed effects. (Cases with rating of "no image" (vividness rating value 1) were excluded from all analysis).

To test how well the factors in the dataset predict recall, a logistic regression analysis was performed. The response for recall was recorded as 1 for recalled verbal descriptions and 0 for not recalled verbal descriptions. Explanatory variables included vividness, stimuli, and reaction time (RT). **Table A2** displays model specifications, including the specified distribution and link function. **Table A3** summarizes the results. **Table A4** shows the full model results by predictor.

All 1490 valid observations were entered in the logistic regression model as the preliminary linear mixed modeling fitting analysis indicated that residual errors were only modestly correlated within each subject and were independent across subjects. Robustness to the violation of the assumption of independence was demonstrated by replicating the results with the following confirmatory repeated measure logistic regression model.

The dichotomous outcome for recall was further modeled with a repeated measure logistic regression analysis. The model was based on the probability of the largest value of response variable, which was 1. Two stimuli, which caused singularity of Hessian matrix, were removed from the dataset, resulting in 1441 observations (and no difference in the results). Models specifications, including the specified distribution and link function were same as the initial logistic regression model.

Type III test evaluated the effect of explanatory variables on recall accuracy in the purposed model. The test result showed that Vividness and stimuli were significant predictors for recall, $\chi^2 = 14.77$ and $\chi^2 = 4276$, $p < 0.05$ respectively (see **Table A5**). No other significance was found. **Table A6** shows estimation for parameters in the model. **Table A7** shows validity of predicted probabilities. The prediction for descriptions which were not recalled was more accurate than that for the verbal descriptions which were, 50.5% of non-recalled descriptions and 22% of recalled descriptions were correctly predicted. This confirmed that the model had overall 72.5% accuracy.

Table A1 | Type III tests of fixed effects in linear mixed model analysis testing the influences of stimuli and image generation time (RTs) on vividness ratings.

Source	*df* numerator	*df* denominator	*F*	*P*
Intercept	1	632.643	3002.619	0.000
Stimuli	59	1000.202	1.588	0.004
RTs	1	1103.724	5.172	0.023
Stimuli × RTs	59	1046.513	0.956	0.572

Dependent variable: vividness.

Table A2 | Basic repeated measure logistic regression model information.

Dependent variable	Recall
Probability distribution	Binomial
Link function	Logit
Observation used	1490

Table A3 | Evaluation result for logistic regression predictive model for incidental recall using stimuli and vividness as predictor.

		Predicted		
		Cues of imagined objects		Percentage correct
Observed		Non-recalled	Recalled	
Cues	Non-recalled	779	142	83.7
	Recalled	259	310	54.7
Overall percentage				72.2

Table A4 | Logistic regression analysis of vividness and incidental recall data.

Predictor	β	SE	Wald	*df*	*p*	Exp (β)
Vividness			17.426	5	0.004	
Vividness (1)	−0.091	0.374	0.059	1	0.808	0.913
Vividness (2)	0.315	0.362	0.760	1	0.383	1.371
Vividness (3)	0.684	0.317	4.647	1	0.031	1.982
Vividness (4)	0.537	0.305	3.100	1	0.078	1.712
Vividness (5)	0.850	0.307	7.684	1	0.006	2.341
Stimuli			246.836	57	0.000	
Stimuli (1)	−2.776	0.751	13.658	1	0.000	0.062
Stimuli (2)	−2.511	0.690	13.256	1	0.000	0.081
Stimuli (3)	−2.485	0.690	12.955	1	0.000	0.083
Stimuli (4)	−0.873	0.659	1.755	1	0.185	0.418
Stimuli (5)	1.055	0.894	1.393	1	0.238	2.873
Stimuli (6)	−2.783	0.710	15.374	1	0.000	0.062
Stimuli (7)	−3.751	0.895	17.561	1	0.000	0.023
Stimuli (8)	−1.146	0.658	3.035	1	0.082	0.318
Stimuli (9)	−2.282	0.696	10.739	1	0.001	0.102
Stimuli (10)	−3.082	0.743	17.185	1	0.000	0.046
Stimuli (11)	−0.587	0.662	0.788	1	0.375	0.556
Stimuli (12)	−3.852	0.893	18.601	1	0.000	0.021
Stimuli (13)	−0.987	0.647	2.330	1	0.127	0.373
Stimuli (14)	−0.684	0.663	1.064	1	0.302	0.505
Stimuli (15)	−0.008	0.696	0.000	1	0.991	0.992
Stimuli (16)	−2.110	0.680	9.636	1	0.002	0.121
Stimuli (17)	−2.882	0.747	14.874	1	0.000	0.056
Stimuli (18)	−2.528	0.689	13.458	1	0.000	0.080
Stimuli (19)	0.254	0.744	0.117	1	0.733	1.289
Stimuli (20)	−2.955	0.745	15.724	1	0.000	0.052
Stimuli (21)	−3.342	0.797	17.586	1	0.000	0.035
Stimuli (22)	−1.610	0.647	6.186	1	0.013	0.200
Stimuli (23)	−2.172	0.679	10.230	1	0.001	0.114
stimuli(24)	−0.780	0.661	1.391	1	0.238	0.458
Stimuli (25)	−2.358	0.695	11.524	1	0.001	0.095
Stimuli (26)	−1.756	0.658	7.112	1	0.008	0.173
Stimuli (27)	−2.515	0.689	13.305	1	0.000	0.081
Stimuli (28)	−3.760	0.895	17.656	1	0.000	0.023
Stimuli (29)	−1.680	0.650	6.682	1	0.010	0.186
Stimuli (30)	−2.244	0.676	11.026	1	0.001	0.106
Stimuli (31)	−2.098	0.664	10.001	1	0.002	0.123
Stimuli (32)	−2.052	0.666	9.500	1	0.002	0.128
Stimuli (33)	−0.850	0.655	1.685	1	0.194	0.427
Stimuli (34)	−1.538	0.651	5.577	1	0.018	0.215
Stimuli (35)	−3.255	0.798	16.655	1	0.000	0.039
Stimuli (36)	−0.493	0.664	0.551	1	0.458	0.611
Stimuli (37)	−1.570	0.656	5.724	1	0.017	0.208
Stimuli (38)	−0.527	0.658	0.641	1	0.423	0.590
Stimuli (39)	−1.746	0.649	7.245	1	0.007	0.174
Stimuli (40)	−1.934	0.672	8.296	1	0.004	0.145
Stimuli (41)	−0.115	0.709	0.026	1	0.871	0.891
Stimuli (42)	0.262	0.723	0.132	1	0.716	1.300
Stimuli (43)	−0.490	0.667	0.539	1	0.463	0.613
Stimuli (44)	−0.764	0.651	1.376	1	0.241	0.466
Stimuli (45)	−2.352	0.674	12.196	1	0.000	0.095

(Continued)

Table A4 | Continued

Predictor	β	SE	Wald	*df*	*p*	Exp (β)
Stimuli (46)	−1.479	0.645	5.259	1	0.022	0.228
Stimuli (47)	−3.562	0.901	15.633	1	0.000	0.028
Stimuli (48)	−2.927	0.747	15.370	1	0.000	0.054
Stimuli (49)	−1.717	0.650	6.989	1	0.008	0.180
Stimuli (50)	−1.421	0.643	4.883	1	0.027	0.241
Stimuli (51)	−2.051	0.665	9.501	1	0.002	0.129
Stimuli (52)	−1.712	0.682	6.299	1	0.012	0.181
Stimuli (53)	−0.028	0.691	0.002	1	0.967	0.972
Stimuli (54)	−2.172	0.679	10.238	1	0.001	0.114
Stimuli (55)	−2.407	0.691	12.132	1	0.000	0.090
Stimuli (56)	−1.138	0.735	2.395	1	0.122	0.321
Stimuli (57)	−2.455	0.691	12.612	1	0.000	0.086
Constant	0.664	0.576	1.329	1	0.249	1.943

Variable(s) entered: vividness, stimuli.

Table A5 | Type III test of model effects for repeated measure logistic regression model.

Source	Wald χ^2	*df*	*p*
(Intercept)	22.744	1	0.000
Vividness	14.766	5	0.011
RT	0.050	1	0.824
Vividness × RT	6.972	5	0.223
Stimuli	4276.081	25	0.000

Dependent variable: recall.

Table A6 | Parameter estimate for repeated measure logistic regression.

			95% Wald confidence interval			Hypothesis test	
Parameter	**β**	**SD**	**Lower**	**Upper**	**Wald**	***df***	***P***
(Intercept)	0.905	0.4538	–	1.794	3.976	1	0.046
[vividness = 2.00]	1.136	0.3524	0.445	1.826	10.383	1	0.001
[vividness = 3.00]	0.748	0.3818	0.000	1.496	3.837	1	0.050
[vividness = 4.00]	0.608	0.4688	−0.311	1.526	1.680	1	0.195
[vividness = 5.00]	0.162	0.2989	−0.424	0.748	0.295	1	0.587
[vividness = 6.00]	0.356	0.2540	−0.142	0.854	1.965	1	0.161
[vividness = 7.00]	0[a]	–	–	–	–	–	–
RT	0.004	0.0054	−0.007	0.015	0.540	1	0.462
[vividness = 2.00] × s	−0.025	0.0188	−0.062	0.012	1.723	1	0.189
[vividness = 3.00] × s	0.018	0.0130	−0.008	0.043	1.892	1	0.169
[vividness = 4.00] × s	−0.005	0.0132	−0.031	0.021	0.147	1	0.701
[vividness = 5.00] × s	0.000	0.0079	−0.016	0.015	0.005	1	0.942
[vividness = 6.00] × s	−0.004	0.0100	−0.023	0.016	0.146	1	0.702
[vividness = 7.00] × s	0[a]	–	–	–	–	–	–
Stimuli							
[stimuli = 1.00]	−2.481	0.6790	−3.811	−1.150	13.345	1	0.000
[stimuli = 2.00]	0.293	0.8285	−1.331	1.917	0.125	1	0.723
[stimuli = 3.00]	0.036	0.7209	−1.377	1.449	0.002	1	0.960
[stimuli = 4.00]	0.023	0.5302	−1.016	1.062	0.002	1	0.965
[stimuli = 5.00]	−1.597	0.4804	−2.538	−0.655	11.046	1	0.001
[stimuli = 6.00]	−3.531	0.7974	−5.094	−1.968	19.609	1	0.000
[stimuli = 7.00]	0.318	0.6115	−0.880	1.517	0.271	1	0.603
[stimuli = 8.00]	1.291	0.9455	−0.563	3.144	1.863	1	0.172
[stimuli = 9.00]	−1.313	0.6087	−2.506	−0.120	4.650	1	0.031
[stimuli = 10.00]	−0.152	0.6943	−1.513	1.209	0.048	1	0.827
[stimuli = 11.00]	0.624	0.6256	−0.603	1.850	0.993	1	0.319
[stimuli = 12.00]	−1.936	0.5493	−3.013	−0.860	12.428	1	0.000
[stimuli = 13.00]	1.380	0.6474	0.111	2.649	4.542	1	0.033
[stimuli = 14.00]	−1.511	0.5988	−2.685	−0.337	6.368	1	0.012
[stimuli = 15.00]	−1.784	0.5815	−2.924	−0.644	9.410	1	0.002
[stimuli = 16.00]	−2.427	0.5536	−3.512	−1.342	19.224	1	0.000
[stimuli = 17.00]	−0.305	0.5736	−1.429	0.819	0.282	1	0.595
[stimuli = 18.00]	0.417	0.7336	−1.020	1.855	0.324	1	0.569
[stimuli = 19.00]	0.066	0.7188	−1.343	1.475	0.008	1	0.927
[stimuli = 20.00]	−2.730	0.7174	−4.136	−1.324	14.485	1	0.000
[stimuli = 21.00]	0.466	0.5001	−0.514	1.446	0.869	1	0.351
[stimuli = 22.00]	0.873	0.6202	−0.342	2.089	1.983	1	0.159
[stimuli = 23.00]	−0.877	0.6083	−2.069	0.315	2.079	1	0.149
[stimuli = 24.00]	−0.297	0.7287	−1.725	1.132	0.166	1	0.684
[stimuli = 25.00]	−1.656	0.6106	−2.853	−0.460	7.359	1	0.007
[stimuli = 26.00]	−0.097	0.7548	−1.576	1.382	0.016	1	0.898
[stimuli = 27.00]	−0.711	0.6595	−2.004	0.581	1.164	1	0.281
[stimuli = 28.00]	0.054	0.6819	−1.282	1.391	0.006	1	0.937
[stimuli = 29.00]	1.313	0.9485	−0.546	3.172	1.915	1	0.166
[stimuli = 30.00]	−0.786	0.6830	−2.125	0.552	1.326	1	0.249
[stimuli = 31.00]	−0.227	0.6441	−1.489	1.036	0.124	1	0.725
[stimuli = 32.00]	−0.366	0.6257	−1.592	0.860	0.342	1	0.559
[stimuli = 33.00]	−0.410	0.7467	−1.874	1.053	0.302	1	0.583
[stimuli = 34.00]	−1.615	0.6717	−2.931	−0.298	5.778	1	0.016
[stimuli = 36.00]	−0.931	0.6474	−2.200	0.338	2.069	1	0.150

(Continued)

Table A6 | Continued

			95% Wald confidence interval			Hypothesis test	
Parameter	**β**	**SD**	**Lower**	**Upper**	**Wald**	***df***	***P***
[stimuli = 37.00]	0.881	0.7813	−0.650	2.412	1.272	1	0.259
[stimuli = 38.00]	−1.978	0.5974	−3.149	−0.807	10.958	1	0.001
[stimuli = 39.00]	−0.934	0.6869	−2.280	0.412	1.849	1	0.174
[stimuli = 40.00]	−1.942	0.5491	−3.018	−0.866	12.513	1	0.000
[stimuli = 41.00]	−0.726	0.5983	−1.898	0.447	1.471	1	0.225
[stimuli = 42.00]	−0.527	0.7072	−1.913	0.860	0.554	1	0.457
[stimuli = 43.00]	−2.346	0.6794	−3.678	−1.014	11.922	1	0.001
[stimuli = 44.00]	−2.760	0.6862	−4.104	−1.415	16.175	1	0.000
[stimuli = 45.00]	−1.974	0.6386	−3.226	−0.723	9.558	1	0.002
[stimuli = 46.00]	−1.686	0.6688	−2.997	−0.375	6.355	1	0.012
[stimuli = 48.00]	−0.120	0.5964	−1.289	1.049	0.041	1	0.840
[stimuli = 49.00]	−0.994	0.5766	−2.124	0.136	2.971	1	0.085
[stimuli = 50.00]	1.164	0.7937	−0.392	2.719	2.150	1	0.143
[stimuli = 51.00]	0.488	0.7849	−1.050	2.027	0.387	1	0.534
[stimuli = 52.00]	−0.759	0.6465	−2.026	0.508	1.379	1	0.240
[stimuli = 53.00]	−1.054	0.6764	−2.380	0.271	2.430	1	0.119
[stimuli = 54.00]	−0.415	0.5582	−1.509	0.679	0.553	1	0.457
[stimuli = 55.00]	−0.753	0.6128	−1.954	0.448	1.509	1	0.219
[stimuli = 56.00]	−2.425	0.7889	−3.971	−0.879	9.448	1	0.002
[stimuli = 57.00]	−0.292	0.6341	−1.535	0.950	0.213	1	0.645
[stimuli = 58.00]	−0.102	0.6570	−1.390	1.185	0.024	1	0.876
[stimuli = 59.00]	−1.387	0.7481	−2.853	0.079	3.439	1	0.064
[stimuli = 60.00]	0[a]	–	–	–	–	–	0.015

Dependent variable: recall.

[a] Set to zero because this parameter is redundant.

Table A7 | Predicted recall value from repeated measure logistic regression model.

		Predicted category value		
Recalled		**0.00**	**1.00**	**Total**
0.0 (no)	Count	728	144	872
	% of Total	50.5	10.0	60.5
1.00 (yes)	Count	252	317	569
	% of Total	17.5	22.0	39.5
Total	Count	980	461	1441
	% of Total	68.0	32.0	100.0

APPENDIX B

A corpus of 66 peer-reviewed experimental journal articles representing 4.32% of the literature available through PsycINFO, and containing the keyword "vividness" was randomly compiled by a research assistant naive to the purposes of the study (see Appendix references). Random selection of a relevant representative sample can be defended as a sound, reasonable meta-analytic tactic, provided the selected sources are analyzed according to a set of predefined, *a priori* criteria (Rosenthal, 1991). As a prerequisite for inclusivity, any statistical outcome directly pertaining to the measures VVIQ and trial-by-trial vividness ratings were to be utilized in the analysis, except those pertaining to *post hoc* comparisons.

The analysis consisted of two phases, a preliminary non-parametric analysis, and a secondary parametric analysis. Data for the preliminary analysis was obtained by partitioning individual experimental outcomes into two 2 × 2 contingency tables. Upon partitioning each experimental outcome as either a significant or non-significant experimental outcome, and as either a VVIQ or trial-by-trial vividness subjective report, each datum was further categorized as either a neural or behavioral/cognitive objective measure.

The same dataset from the preliminary non-parametric analysis was utilized in the secondary parametric analysis. However, each binomial outcome was transformed into an exact probability value. Analytic accuracy was maintained by calculating probabilities from reported test statistics and degrees of freedom. If required, raw data was statistically analyzed anew from means and variance. This rule was strictly adhered to unless otherwise unavoidable, in which case probability signifiers were rounded to the reported cut-off (i.e., $p < 0.05$ was approximated as 0.05); however, it should be noted that rounding was required six times over the course of 863 entries. The resultant entries were then categorized as either VVIQ or trial-by-trial vividness subjective report, and as either a neural or behavioral/cognitive objective measure. All values within each category were summated, and divided by the square root of the number of entries within each category.

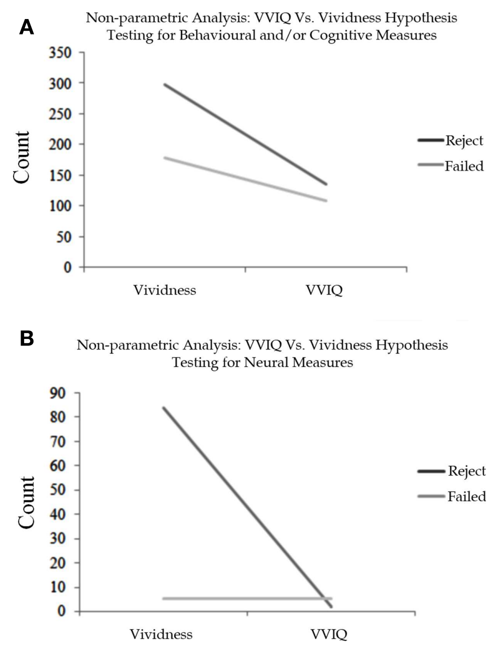


FIGURE A1 | (A) Experimental outcomes for trial-by-trial vividness ratings and VVIQ with respect to behavioral and/or cognitive measures. **(B)** Experimental outcomes for trial-by-trial vividness ratings and VVIQ with respect to neural measures. The dark line refers to results which reject the null hypothesis, and the light line refers to results which fail to reject the null hypothesis.

A non-parametric analysis examining experimental outcome between VVIQ and trial-by-trial vividness ratings is presented in **Figures A1A,B**. The data in **Figure A1A** represent the number of significant versus non-significant experimental outcomes for VVIQ and trial-by-trial vividness ratings for behavioral/cognitive objective measures. The data in **Figure A1B** represent the number of significant versus non-significant experimental outcomes for VVIQ and trial-by-trial vividness ratings for neural objective measures. A higher proportion of successes accompany trial-by-trial vividness ratings for both behavioral/cognitive and neural objective measures. This relationship is especially true for studies underlying the neural origin of vividness.

The trends observed in the preliminary analysis prompted the use of a more sensitive statistical procedure. Because the directionality of each statistical outcome was not immediately apparent, and degrees of freedom often exceed one for F-tests and Chi-square tests of significance, standard meta-analytic methodology was decidedly insufficient for such purposes (Rosenthal, 1991). Under these circumstances, Stouffer's method of adding Z's provides a straightforward and reasonable estimate (Mosteller and Bush, 1954; Rosenthal, 1991). Upon determining exact probability values for each entry introduced, the values were transformed into their standard normal deviates. These values were summated, and divided by the square root of the number of entries within each category. Data for the parametric analysis is shown in **Figure A2.** These data show the summated Z-scores for VVIQ and trial-by-trial vividness ratings for behavioral/cognitive and neural objective measures.

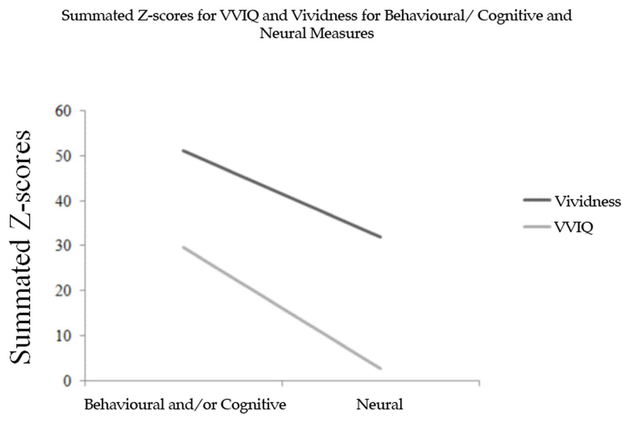


FIGURE A2 | Summed Z-scores for vividness and VVIQ for neural and behavioral and/or cognitive measures.

As evidenced by **Figure A2**, two trends remain especially salient. Firstly, trial-by-trial vividness ratings are consistently greater for behavioral/cognitive and neural measures. Secondly, behavioral/cognitive measures yield significantly greater values than those which are neural. These results suggest that trial-by-trial vividness ratings are a more effective means by which to measure the subjective experience of mental imagery. Furthermore, Fisher's Z-transformation for experimental outcomes concerning the correlation between VVIQ scores and trial-by-trial vividness ratings for 21 entries retrieved from six of the peer-reviewed journal articles showed an average Z_r of 0.154, and variability in these values ranged from $r = -0.27$, to $r = 0.64$. Consistent with the results of experiment 2, these results support the contention that trial-by-trial vividness self-reports and VVIQ scores share some descriptive properties of visual imagery; however, trial-by-trial vividness ratings seem to be much more resolved.

REFERENCES

Adeyemo, S. A. (1998). Imagery unvividness and social-psychological implications. *J. Ment. Imagery* 22, 79–98.

Aleman, A., Böcker, K. B. E., and de Haan, E. H. F. (1999). Disposition towards hallucination and subjective versus objective vividness of imagery in normal subjects. *Pers. Individ. Differ.* 27, 707–714.

Aleman, A., and de Haan, E. H. F. (2001). Do people with better mental imagery ability make more reality monitoring errors? *Curr. Psychol. Lett.* 6, 7–15.

Aleman, A., and de Haan, E. H. F. (2004). Fantasy proneness, mental imagery and reality monitoring. *Pers. Individ. Differ.* 36, 1747–1754.

Aleman, A., Nieuwenstein, M. R., Böcker, K. B. E., and de Haan, E. H. F. (2000). Mental imagery and perception in hallucination-prone individuals. *J. Nerv. Ment. Dis.* 188, 830–836.

Allbutt, J., Ling, J., Heffernan, T. M., and Shafiullah, M. (2008). Self-report imagery questionnaire scores and subtypes of social-desirable responding: auditory imagery, visual imagery, and thinking style. *J. Individ. Differ.* 29, 181–188.

Baddeley, A. D., and Andrade, J. (2000). Working memory and the vividness of imagery. *J. Exp. Psychol. Gen.* 129, 126–145.

Barlett, C. P., and Brannon, L. A. (2007). "If only...": the role of visual imagery in counterfactual thinking. *Imagin. Cogn. Pers.* 26, 87–100.

Belardinelli, M. O., Palmiero, M., Sestieri, C., Nardo, D., Di Matteo, R., Londei, A., et al. (2009). An fMRI investigation on image generation in different sensory modalities: the influence of vividness. *Acta Psychol. (Amst.)* 132, 190–200.

Bent, N. A., and Wick, E. (2006). Beyond vividness: parental filters as moderators in mental imagery and measured anxiety level. *J. Ment. Imagery* 30, 21–38.

Bird, C. M., and Burgess, N. (2008). The hippocampus and memory: insights from spatial processing. *Nat. Rev. Neurosci.* 9, 182–194.

Borkovec, T. D., and Sides, J. K. (1979). The contribution of relaxation and expectancy to fear reduction via graded, imaginal exposure to feared stimuli. *Behav. Res. Ther.* 17, 529–540.

Burton, L. J. (2003). Examining the relation between visual imagery and spatial ability tests. *Int. J. Test.* 3, 277–291.

Campos, A., Amor, Á., and González, M. Á. (2002). Presentation of keywords by means of interactive drawings. *Span. J. Psychol.* 5, 102–109.

Campos, A., Gómez-Juncal, R., and Pérez-Fabello, M. J. (2007). Experience in imagery and imagery vividness. *Imagin. Cogn. Pers.* 27, 337–348.

Campos, A., López, A., and González, M. A. (1999a). Effects of eyes open/closed and order of rating on VVIQ scores. *J. Ment. Imagery.* 23, 35–43.

Campos, A., Marcos, J. L., and Gonzáles, M. Á. (1999b). Emotionality of words as related to vividness of imagery and concreteness. *Percept. Mot. Skills* 88, 1135–1140.

Cartwright, D., Jenkins, J. L., Chavez, R., and Peckar, H. (1983). Studies in imagery and identity. *J. Pers. Soc. Psychol.* 44, 376–384.

Cocude, M., and Denis, M. (1988). Measuring the temporal characteristics of visual images. *J. Ment. Imagery* 12, 89–101.

Conduit, R., Crewther, S. G., and Coleman, G. (2004). Spontaneous eyelid movements (ELMS) during sleep are related to dream recall on awakening. *J. Sleep Res.* 13, 137–144.

Cornoldi, C., de Beni, R., Cavedon, A., and Mazzoni, G. (1992). How can a vivid image be described? Characteristics influencing vividness judgments and the relationship between vividness and memory. *J. Ment. Imagery* 16, 89–107.

Cui, X., Jeter, C. B., Yang, D., Montague, P. R., and Eagleman, D. M. (2007). Vividness of mental imagery: individual variability can be measured objectively. *Vision Res.* 47, 474–478.

D'Angiulli, A. (2002). Mental image generation and the contrast sensitivity function. *Cognition* 85, B11–B19.

D'Angiulli, A., and Reeves, A. (2002). Generating visual mental images: latency and vividness are inversely related. *Mem. Cognit.* 30, 1179–1188.

D'Angiulli, A., and Reeves, A. (2005). "Picture theory, tacit knowledge or vividness-core? Three hypotheses on the mind's eye and its elusive size," in *Proceedings of the 27th Cognitive Science Society Annual Meeting*, eds B. Bara, L. Barsalou, and M. Bucciarelli (Stresa: The Cognitive Science Society), 536–541.

D'Angiulli, A., and Reeves, A. (2007). The relationship between self-reported vividness and latency during mental size scaling of everyday items: phenomenological evidence of different types of imagery. *Am. J. Psychol.* 120, 521–551.

De Pascalis, V., Marucci, F. S., Penna, P. M., and Pessa, E. (1987). Hemispheric activity of 40 Hz EEG during recall of emotional events: differences between low and high hypnotizables. *Int. J. Psychophysiol.* 5, 167–180.

Dean, G. M., and Morris, P. E. (2003). The relationship between self-reports of imagery and spatial ability. *Br. J. Psychol.* 94, 245–273.

Dyckman, J. M., and Cowan, P. A. (1978). Imaging vividness and the outcome of in vivo and imagined scene desensitization. *J. Consult. Clin. Psychol.* 46, 1155–1156.

Eysenck, M. (1974). Age differences in incidental learning. *Dev. Psychol.* 10, 936–941.

Gale, A., Morris, P. E., Lucas, B., and Richardson, A. (1972). Types of imagery and imagery types: an EEG study. *Br. J. Psychol.* 63, 523–531.

Giusberti, F., Cornoldi, C., De Beni, R., and Massaroni, M. (1992). Differences in vividness ratings of perceived and imagined patterns. *Br. J. Psychol.* 83, 533–547.

Hale, S. M., and Simpson, H. M. (1971). Effects of eye movements on the rate of discovery and the vividness of visual images. *Percept. Psychophys.* 9, 242–264.

Holmes, E. A., Lang, T. J., Moulds, M. L., and Steele, A. M. (2008). Prospective and positive mental imagery deficits in dysphoria. *Behav. Res. Ther.* 46, 976–981.

Hunt, M., Bylsma, L., Brock, J., Fenton, M., Goldberg, A., Miller, R., et al. (2006). The role of imagery in the maintenance and treatment of snake fear. *J. Behav. Ther. Exp. Psychiatry* 37, 283–298.

Iaccino, J., and Byrne, J. (1989). Mental layouts of concealed objects as a function of imagery type and experimental conditions. *Bull. Psychon. Soc.* 27, 402–404.

Jenness, A., and Jorgensen, A. P. (1941). Ratings of vividness of imagery in the waking state compared with reports of somnambulism. *Am. J. Psychol.* 54, 253–259.

Keenan, J. M., and Moore, R. E. (1979). Memory for images of concealed objects: a reexamination of Neisser and Kerr. *J. Exp. Psychol. Learn. Mem. Cogn.* 5, 374–385.

Keogh, L., and Markham, R. (1998). Judgements of other people's memory reports: differences in reports as a function of imagery vividness. *Appl. Cogn. Psychol.* 12, 159–171.

Kerr, N. H., and Neisser, U. (1983). Mental images of concealed objects-new evidence. *J. Exp. Psychol. Learn. Mem. Cogn.* 9. 212–221.

Kidd, L. K., and Workman, J. E. (1999). Assessment of creativity in apparel design. *Cloth. Text. Res. J.* 17, 58–64.

Kosslyn, S., and Alper, S. (1977). On the pictorial properties of visual images: effects of image size on memory for words. *Can. J. Psychol.* 31, 32–40.

Lorey, B., Pilgramm, S., Bischoff, M., Stark, R., Vaitl, D., Kindermann S., et al. (2011). Activation of the parieto-premotor network is associated with vivid motor imagery–a parametric fMRI Study. *PLoS ONE* 6:e20368. doi:10.1371/journal.pone.0020368

Mantani, T., Okamoto, Y., Shirao, N., Okada, G., and Yamawaki, S. (2005). Reduced activation of posterior cingulate cortex during imagery in subjects with high degrees of alexithymia: a functional magnetic resonance imaging study. *Biol. Psychiatry* 57, 982–990.

Markham, R., and Hynes, L. (1993). The effect of vividness of imagery on reality monitoring. *J. Ment. Imagery* 17, 159–170.

Marks, D. F., and Isaac, A. R. (1995). Topographical distribution of EEG activity accompanying visual and motor imagery in vivid and non-vivid imagers. *Br. J. Psychol.* 86, 271–282.

McClelland, A., Kemps, E., and Tiggemann, M. (2006). Reduction of vividness and associated craving in personalized food imagery. *J. Clin. Psychol.* 62, 355–365.

McKelvie, S. J. (1998). Effects of gender on reported vividness of visual imagery for parents. *J. Ment. Imagery* 22, 99–112.

Morrison, R. G., and Wallace, B. (2001). Imagery vividness, creativity and the visual arts. *J. Ment. Imagery* 25, 135–152.

Neisser, U., and Kerr, N. (1973). Spatial and mnemonic properties of visual images. *Cogn. Psychol.* 5, 138–150.

Overton, J. A. G. (2004). Correlation of EEG activity with subjective performance on a guided imagery test: an exploratory study. *J. Ment. Imagery* 28, 17–60.

Pearson, J., Rademaker, R. L., and Tong, F. (2011). Evaluating the mind's eye: the metacognition of visual imagery. *Psychol. Sci.* 22, 1535–1542.

Rauch, S. A. M., Foa, E. B., Furr, J. M., and Filip, J. C. (2004). Imagery vividness and perceived anxious arousal in prolonged exposure treatment for PTSD. *J. Trauma. Stress* 17, 461–465.

Rehm, L. P. (1973). Relationships among measures of visual imagery. *Behav. Res. Ther.* 11, 265–270.

Riske, M. L., Wallace, B., and Allen, P. A. (2000). Imaging ability and eye witness accuracy. *J. Ment. Imagery* 24, 137–148.

Ritchey, H., and Beal, C. (1980). Image detail and recall: evidence for within-item elaboration. *J. Exp. Psychol.* 6, 66–76.

Rodway, P., Gillies, K., and Schepman, A. (2006). Vivid imagers are better at detecting salient changes. *J. Individ. Differ.* 27, 218–228.

Sadler, P., and Woody, E. Z. (2006). Does the more vivid imagery of high hypnotizables depend on greater cognitive effort? A test of dissociation and social-cognitive theories of hypnosis. *Int. J. Clin. Exp. Hypn.* 54, 372–391.

Sheehan, P. W. (1966). Accuracy and vividness of visual images. *Percept. Mot. Skills* 23, 391–398.

Slee, J. A. (1980). Individual differences in visual imagery ability and the retrieval of visual appearances. *J. Ment. Imagery* 4, 93–113.

Sutherland, M. E., Harrell, J. P., and Isaacs, C. (1987). The stability of individual differences in imagery ability. *J. Ment. Imagery* 11, 97–104.

Tracy, R. J., Roesner, L. S., and Kovac, R. N. (1988). The effect of visual versus auditory imagery on vividness and memory. *J. Ment. Imagery* 12, 145–161.

Van Diest, I., De Peuter, S., Devriese, S., Wellens, E., Van, d. W., and Van, d. B. (2005). Imagined risk of suffocation as a trigger for hyperventilation. *Psychosom. Med.* 67, 813–819.

Vergeer, I., and Roberts, J. (2006). Movement and stretching imagery during flexibility training. *J. Sports Sci.* 24, 197–208.

Walczyk, J. J., and Hall, V. C. (1988). The relationship between imagery vividness ratings and imagery accuracy. *J. Ment. Imagery* 12, 163–171.

Zhang, M., Weisser, V. D., Stilla, R., Prather, S. C., and Sathian, K. (2004). Multisensory cortical processing of object shape and its relation to mental imagery. *Cogn. Affect. Behav. Neurosci.* 4, 251–259.

frontiers in
HUMAN NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 08 November 2012
doi: 10.3389/fnhum.2012.00305

The effects of visual imagery on face identification: an ERP study

Jianhui Wu[1]*, Hongxia Duan[1,2], Xing Tian[3], Peipei Wang[4] and Kan Zhang[1]

[1] Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, Beijing, China
[2] University of the Chinese Academy of Sciences, Beijing, China
[3] Department of Psychology, New York University, New York, NY, USA
[4] Center for Higher Brain Functions, Capital Medical University, Beijing, China

Edited by:
Joel Pearson, The University of New South Wales, Australia

Reviewed by:
Ruthger Righart, Institute for Stroke and Dementia Research, Germany
Michael Wibral, Goethe University, Germany

***Correspondence:**
Jianhui Wu, Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, 4A Datun Road, Chaoyang District, Beijing 100101, China.
e-mail: wujh@psych.ac.cn

The present study tested the hypothesis that the effects of mental imagery on subsequent perception occur at a later matching stage in perceptual identification, but not in the early perceptual stage as in perceptual detection. The behavioral results suggested that the effect of visual imagery on visual identification is content-specific, i.e., imagining a congruent face facilitates face identification, whereas a mismatch between imagery and perception leads to an interference effect. More importantly, the ERP results revealed that a more negative N2 response to the subsequent visual face stimuli was elicited over fronto-central sites in the mismatch and no-imagery conditions as compared to that in the match condition, with the early P1 and N170 components independent of manipulations. The latency and distribution of the neural effects demonstrate that the matching step, but not the earlier perceptual process, is affected by the preceding visual imagery in the context of face identification. We discuss these results in a broader context that the imagery-perception interaction may depend on task demand.

Keywords: visual imagery, face identification, event-related potentials, N170, N2, matching

INTRODUCTION

Our visual perception is shaped by our previous visual experience in the real world (Gilbert, 1996; Webster and MacLeod, 2011). The visual imagery process simulates perceptual representations on the basis of past experience and provides a mental template that can influence the subsequent perception (Moulton and Kosslyn, 2009). Evidence supporting the modulation effects of mental imagery, however, demonstrates distinct directional influence on perception. Some studies show that imagery interferes with perception (Perky, 1910; Segal and Fusella, 1970; Reeves, 1981; Craver-Lemley and Reeves, 1987; Craver-Lemley et al., 1997; Ishai and Sagi, 1997b; Craver-Lemley and Arterberry, 2001); whereas others show facilitation effects (Freyd and Finke, 1984; Farah, 1985; Ishai and Sagi, 1995, 1997a; Pearson et al., 2008).

Several factors have been proposed to reconcile the conflicting results about the different directional effects of mental imagery. The direction of imagery modulation has been hypothesized to be content-specific. That is, whether preceding imagery facilitates or interferes with perception depends on how similar the imagined and presented patterns are. Subjects were more likely to perceive a stimulus when the imagined content matched the presented stimulus than when the two were mismatched (Peterson and Graham, 1974; Farah, 1985, 1989; Finke, 1986; Farah et al., 1988; Djordjevic et al., 2004a,b).

Task demands can be another factor that determines the direction of interaction between imagery and perception. The interference effect was found when the subsequent perceptual task was a simple detection task, i.e., to give a single response to any stimuli presented while not necessary to identify them; whereas the facilitation effect was observed during an identification task, i.e., to give the appropriate response to different stimuli (Finke, 1986). Detection is modeled as a task that only needs to register the presence of a stimulus in an all-or-none fashion; further processing of any specific features of stimuli are not required, and may even slow down the detection performance. In contrast, identification is modeled as a task that necessitates the processing of detailed featural information for the following matching processing, in which the comparison between the bottom-up sensory information and the top-down formed template is carried out to classify the sensory input. Most of the previous studies focused on the effects of imagery on subsequent perceptual detection, in which both directional modulations were found (e.g., Segal and Fusella, 1970; Farah, 1985). With only a few studies investigating the aftereffect on identification (Finke, 1986; McDermott and Roediger, 1994; Cabeza et al., 1997; Michelon and Koenig, 2002), consistent facilitation effects were reported, i.e., mental imagery of a visual pattern in advance will facilitate the identification of the same pattern.

The neural mechanisms mediating the interaction between imagery and perception are also unclear. One of the questions is that at which level the top-down and bottom-up processes interact. The dominant interpretation is the *perceptual level hypothesis* in which this facilitation/interference effect occurs at the early perceptual level, where the visual processing of visual features in external stimuli is directly manipulated by preceding imagery (Peterson and Graham, 1974; Neisser, 1976; Freyd and Finke, 1984; Farah, 1985; Craver-Lemley and Reeves, 1987; Ishai and Sagi, 1995; Pearson et al., 2008). However, the differential observations of interference and facilitation effects during detection and identification tasks (Finke, 1986), lead us to propose, along the line of a similar theory (Finke, 1986), that perceptual task demand may influence the occurrence of imagery-perception interaction at distinct stages along the visual information processing stream. Specifically, the effects of imagery on perceptual identification occur at the later stage where integrated features of an object are matched with those stored in memory to achieve recognition (referred to as the *matching level hypothesis* henceforth in this paper); whereas the effects of imagery on perceptual detection are presented in early perceptual processes where the representation of object features in establishing and spotting the existence of features in an all-or-none fashion would be sufficient in the task of detection *(perceptual level hypothesis)*.

To distinguish the *perceptual level hypothesis* and the *matching level hypothesis*, we need to investigate the dynamics of cognitive functions. It is hard, if not impossible, to separate different cognitive stages using behavioral experiments as in most previous studies, because the behavioral performance is the cumulative result of processing at multiple levels. Moreover, the behavioral measures suffer from methodological limitations, such as confounds from experimenter expectancy effects and subjects' tacit knowledge (Farah et al., 1988; Pylyshyn, 2002). The event-related potentials (ERP) technique, on the other hand, is an objective measure that may be relatively less confounded by these strategic effects. More importantly, with its high temporal resolution, ERP can be used to determine the time course of neural activity, making it possible to determine the cognitive stage at which mental imagery has effects on perception.

To our knowledge, only two ERP studies have been carried out to investigate the effects of visual imagery on perception, using either a detection task paradigm (Farah et al., 1988) or without an active task (Ganis and Schendan, 2008). Farah et al. (1988) found larger early negativity to the subsequent visual stimuli which peaked at 160 ms over temporo-occipital sites when the imagery matched perception compared with the mismatch condition. Consistently, stimuli were detected more often for the match than for the mismatch condition. Ganis and Schendan (2008) observed that both perception and imagery affect the N170 response to the subsequently presented test faces. Specifically, the amplitude of N170 was enhanced when they were preceded by face imagery rather than object imagery, and similar effects were found for the non-face objects. Interestingly, Sreenivasan et al.'s study (2007) also found that noise probes presented during the delay interval of a delayed-recognition task elicited a larger N170 during face relative to house working memory. These ERP experiments provided neural evidence for the *perceptual level hypothesis* for visual detection. In these two ERP experiments, however, no identification task was implemented; it is still unclear at which stage mental imagery affects the identification task.

The goal of the present ERP study was to test the proposed *matching level hypothesis* in visual identification by examining the cognitive stage(s) at which the interaction occurs between face imagery and face identification. We designed a face imagery-face identification paradigm, during which participants were required to imagine one of two faces or without any imagery followed by an identification task to visually presented face pictures. We predicted that, behaviorally, matching between imagined and presented faces will lead to a facilitation effect as compared with the no imagery condition, and mismatch will lead to an interference effect. For the electrophysiological recording, at least two processing components were proposed in the context of face recognition (Bruce and Young, 1986): the early pre-categorical structural encoding of faces reflected in the early N170 component (e.g., Sagiv and Bentin, 2001) and a later matching process between encoded facial representation and stored structural codes (templates) presumably mediated by a fronto-central distributed N2 component that is usually elicited by a perceptual mismatch from a template and detection of novelty (for a review, see Folstein and Van Petten, 2008). Some studies reported that the earlier P1 component is also related to the earlier stage of face perception (Itier and Taylor, 2004; Thierry et al., 2007). According to the *matching level hypothesis*, this N2 component is predicted to be modulated by preceding imagery; whereas the modulation of the P1 and N170 responses is predicted by the *perceptual level hypothesis* and any changes observed in these early components will suggest the formation of perceptual features in the identification task is equally affected by preceding mental imagery.

MATERIALS AND METHODS

PARTICIPANTS

Data from 24 participants (mean age 22 ± 1.7 years, 12 men, all right-handed) were collected and analyzed. All were undergraduates from China Agricultural University and Beijing Forestry University who gave informed consent and were paid for their participation. None of them had a history of neurological or psychiatric disorders. All reported normal hearing and normal or corrected-to-normal vision. The experiment had been approved by the Ethics Committee of Human Experimentation in the Institute of Psychology, Chinese Academy of Sciences.

STIMULI

Two female faces with neutral expression from the Revision of the Chinese Facial Affective Picture System (Gong et al., 2011) were chosen as visual stimuli for both imagery and real presentation. These two pictures and one visual noise mask (**Figure 1**) were presented on a computer screen placed 75 cm away from the participants' eyes and subtended at an angle of approximately 7° horizontally and 7.7° vertically. Three auditory vowels (a, o, and u) were used as acoustic cues to indicate different tasks (imagery vs. no-imagery) as well as content in the imagery tasks. Specifically, two of them indicated participants to visualize different faces and the third one reminded participants not to imagine anything (no-imagery condition). The associations between auditory cues and imagined faces/no imagery were counterbalanced across participants. The duration of each sound was 600 ms and the intensity was adjusted to a comfortable listening level of about 70 dB SPL using Adobe Audition (version 1.0). Sounds were delivered binaurally through headphones by the Stim interface system (Neuroscan Labs, Sterling, VA).

PROCEDURE

Participants were seated in a relaxed position on a comfortable chair in a dimly lit, sound-attenuated, and electrically isolated

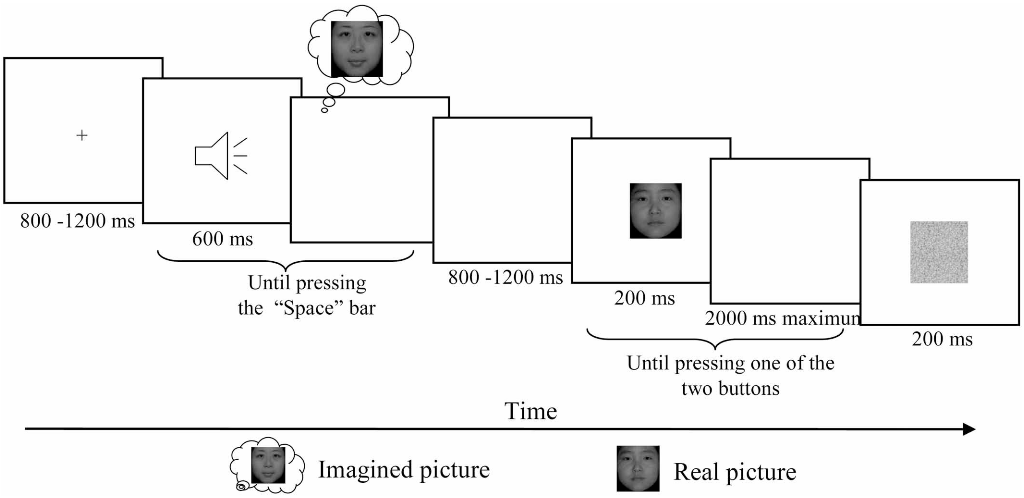


FIGURE 1 | Schematic description of the experimental paradigm. After the presentation of fixation, one of these three letter sounds was presented and participants were asked to imagine the corresponding face vividly and then press the "Space" bar in the self-paced interval. For the no imagery associated sound, participants were asked to simply press "Space" once hearing the letter. After a random delay between 800 and 1200 ms, one of these two real faces was presented for 200 ms and participants were asked to press one of the two buttons according to which of the two faces presented on the screen. After the response a visual noise mask with duration of 200 ms was presented before the next trial began.

room. Participants completed six testing blocks while EEG was recorded. Each block started with a short familiarization session to remind participants of the pairing of auditory cues and tasks. After a fixation presented with duration between 800 and 1200 ms, one of the auditory cues was presented and the participants were asked to press the "Space" bar after they vividly formed the image of the corresponding face. In the no-imagery condition, participants were asked to press "Space" once hearing the letter. A random delay between 800 and 1200 ms was inserted before a visual face stimulus was presented for 200 ms followed by a maximal 2000 ms response window, in which participants were asked to press the left button for one face and right button for another face as quickly and accurately as possible. A visual noise mask with duration of 200 ms was presented before the next trial began (**Figure 1**).

This experiment included three conditions that differed in the tasks associated with the different auditory cues. Participants were asked either to skip an auditory cue (no-imagery condition), or to vividly visualize a corresponding face that could be congruent (match condition) or incongruent (mismatch condition) with the subsequent visual stimuli. Sixty testing trials were included in each block, yielding 120 trials for each condition. The two visual face stimuli were presented equally often in each condition. A pseudo-random presentation order was used, so that participants could not predict the upcoming visual face stimuli. Sequential effects of trial-to-trial transitions were also counterbalanced within each block.

Four training sessions were run before the EEG recordings to ensure the correct pairing of auditory cue and imagery, as well as the vividness of visual imagery. The first session was a familiarization session in which the auditory cues and corresponding face pictures (or no picture) were presented simultaneously at least 35 times for each pair, until participants reported that they had learned these associations. The second session was an imagery training session in which only the auditory cues were presented and the participants were encouraged to vividly imagine the corresponding faces; the correct pairing of visual stimulus and auditory cue was then presented, and participants were required to adjust their imagination. For the no imagery condition, only a vowel sound was presented and no imagination or subsequent adjustment was required. Each pair was repeated 20 times to ensure that participants were able to imagine the corresponding faces vividly. The third session was a face identification training session in which one of these two faces was presented and the participants were asked to identify the presented face by pressing the "left" or "right" key as quickly and accurately as possible. The association of visual stimuli and response keys was counterbalanced across participants. Feedback was provided following responses. The fourth session was an imagery-perception training session with identical procedure to the testing blocks, except feedback was provided after their responses instead of a white noise mask. For all the four training sessions, each participant received equal exposure to both faces.

Participants completed a brief questionnaire by rating the vividness of their visual imagery on a 7-point scale (1 = very vivid imagery, 7 = no imagery at all) at the end of this experiment.

EEG RECORDING AND ANALYSIS

EEG data were continuously recorded from 64 cap-mounted Ag/AgCl electrodes arranged according to the 10–20 international placement system (Neuroscan Inc.) with an on-line reference to the left mastoid and off-line algebraic re-reference to the average of the left and right mastoids. The EEG data were amplified with a bandpass filter of 0.05–100 Hz and digitized at 500 Hz. The vertical and horizontal electrooculogram (VEOG and HEOG) were recorded from two pairs of electrodes: one pair placed 1 cm above and below the left eye, and another pair placed 1 cm lateral from the outer canthi of both eyes. Interelectrode impedances were maintained below 5 kΩ.

The EEG data were processed offline using the Neuroscan 4.3 software. Ocular artifacts were removed using a regression procedure implemented in the Neuroscan software (Semlitsch et al., 1986). Data were lowpass filtered with cutoff frequency at 30 Hz and epochs of 400 ms in duration (including 100 ms of pre-stimulus time as a baseline) were extracted, time-locked to the onset of visual stimuli. Epochs exceeding ±70 μV were considered artifacts and rejected from further analysis. Average responses were obtained for each condition.

The peak amplitude and latency of P1 were measured at electrodes PO7, CB1, O1 PO8, CB2, and O2 and were subjected to a repeated measures Three-Way ANOVA with factors of matching (mismatch, match and no-imagery) × laterality (left and right) × sites. The mean amplitude of N170 was measured in the time window of 140–180 ms over 12 parieto-occipital sites (P7, PO3, PO5, PO7, O1, CB1, P8, PO4, PO6, PO8, O2, and CB2) and was subjected to a repeated measures Three-Way ANOVA with factors of matching (mismatch, match and no-imagery) × laterality (left and right) × sites. Given that the N170 originates in temporal regions (e.g., Henson et al., 2003), analysis in posterior and inferior temporal channels seems necessary. We thus ran separate ANOVAs for each of the four channels (TP7, TP8, T7, and T8), though the N170 amplitude was small or the polarity was reversed in these channels. The mean amplitude of the N2 component was measured in the time window of 250–350 ms at the following 21 sites: Fz, FCz, Cz, CPz, Pz, POz, Oz, F3, FC3, C3, CP3, P3, PO3, O1, F4, FC4, C4, CP4, P4, PO4, and O2. The N2 amplitudes were subjected to a repeated measures Three-Way ANOVA with factors of matching (mismatch, match and no imagery) × anterior-posterior scalp location (F, FC, C, CP, P, PO, and O) × laterality (left, midline and right). Three additional ANOVAs were carried out to directly test the distinct neural correlates of facilitation/interference with paired matching conditions (match vs. mismatch, match vs. no-imagery, and mismatch vs. no-imagery). Behaviorally incorrect trials were excluded from analysis. The Greenhouse–Geisser correction was used to adjust for sphericity violations. The Bonferroni correction was applied for multiple comparisons. We also performed *post-hoc* Pearson correlation analyses in order to assess the relationship between ERP components and actual behavior (two-tailed).

RESULTS

BEHAVIOR AND POST-EXPERIMENTAL QUESTIONNAIRE

The mean reaction time (RT) in the face identification task was significantly different among the three matching conditions [$F_{(2, 46)} = 26.07$, $p < 0.001$] (**Figure 2**). Pairwise comparisons indicated that participants reliably responded faster on match trials than both no imagery trials (456 vs. 482 ms, $p < 0.001$) and mismatch trials (456 vs. 497 ms, $p < 0.001$), but slower on mismatch trials than on no imagery trials (497 vs. 482 ms, $p = 0.01$). The response accuracy in the face identification task was not significantly different among the three matching conditions [$F_{(2, 46)} = 1.88$, $p = 0.17$; mismatch: 95.58%, match: 97.02%, no-imagery: 96.44%].

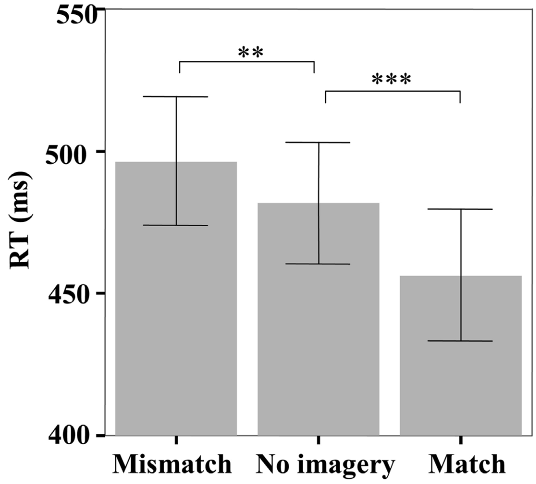


FIGURE 2 | The mean reaction times in the face identification task. Error bars represent two standard errors of the mean (SEM). **$p < 0.01$; ***$p < 0.001$.

The mean RT of "Space" bar presses after the onset of the auditory cue was significantly different among the three cues [$F_{(2, 46)} = 8.50$, $p = 0.001$]. Pairwise comparisons indicated that participants spent less time after the no-imagery cue than after the two face-imagery cues (976 ± 265 ms vs. 1358 ± 698 ms, $p < 0.01$ and 976 ± 265 ms vs. 1359 ± 806 ms, $p < 0.05$), but there was no significant difference between the two face-imagery cues ($p = 0.98$). The post-experimental questionnaire revealed that participants had experienced vivid face imagery (2.54 ± 0.72) when the imagery cues were presented.

ERP RESULTS

An occipital P1 was observed in all three conditions (**Figure 3**). ANOVA did not reveal any significant difference among conditions for either the peak amplitude or latency of the P1 component [$F_{(2, 46)} = 0.42$, $p = 0.66$ and $F_{(2, 46)} = 1.62$, $p = 0.21$, respectively], and also no interaction effect between matching conditions and laterality [$F_{(2, 46)} = 0.57$, $p = 0.57$ and $F_{(2, 46)} = 1.18$, $p = 0.31$, $\varepsilon = 0.72$, respectively]. Additional correlation analysis revealed no significant correlation between RT change and the P1 amplitude/latency differences between the paired comparisons of the three conditions ($Ps > 0.10$).

Typical N170 potentials were observed in responses to visual face stimuli over parieto-occipital sites (**Figure 3**). ANOVA did not reveal any significant difference among conditions for the mean amplitude of the N170 component [$F_{(2, 46)} = 0.091$, $p = 0.91$], and also no interaction effect between matching conditions and laterality or sites [$F_{(2, 46)} = 0.70$, $p = 0.50$ and $F_{(10, 230)} = 1.01$, $p = 0.42$, $\varepsilon = 0.60$, respectively]. Additional correlation analysis revealed no significant correlation between RT change and the N170 amplitude differences across all the 12 channels between the paired comparisons of the three conditions ($Ps > 0.10$). ANOVAs also did not reveal any significant differences among conditions in TP7, TP8, T7, and T8 ($Ps > 0.10$).

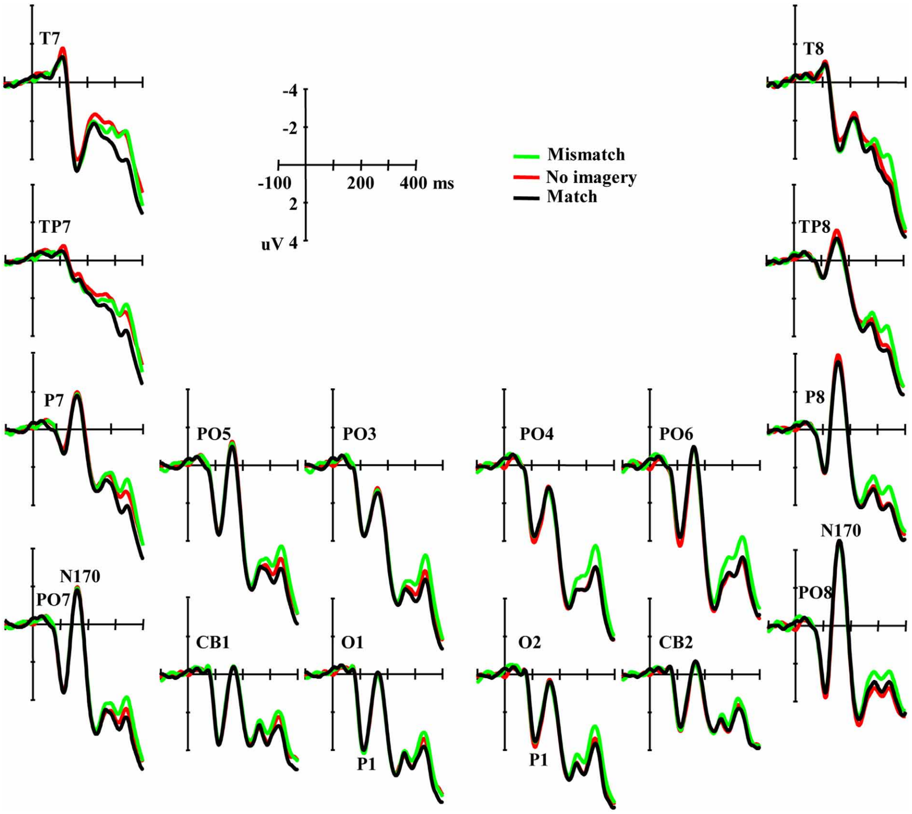


FIGURE 3 | Grand averaged ERPs illustrating the P1 and the N170 elicited by visual stimuli in three conditions.

For the amplitude of the N2 component elicited by real faces (**Figure 4**), ANOVA revealed a significant main effect of matching (mismatch, match and no-imagery) [$F_{(2,\ 46)} = 5.56$, $p < 0.01$], and a marginally significant interaction effect between matching factors and anterior-posterior scalp location [$F_{(12,\ 276)} = 2.50$, $p < 0.10$, $\varepsilon = 0.21$]. One additional ANOVA indicated a significant main effect of matching (mismatch vs. match) [$F_{(1,\ 23)} = 13.92$, $p < 0.01$] and an interaction effect between matching (mismatch vs. match) and anterior-posterior electrodes [$F_{(6,\ 138)} = 3.39$, $p < 0.05$, $\varepsilon = 0.29$]. *Post-hoc* analyses revealed that the mismatch condition was more negative than the match condition and this effect was broadly distributed along the anterior–posterior dimension ($Ps < 0.01$), but was maximal at the fronto-central areas. Correlation analysis revealed a marginally significant negative correlation between RT change (mismatch vs. match) and the N2 amplitude change (mismatch-match) over anterior–posterior sensors ($r = -0.347$, $p = 0.096$) (i.e., the longer the RT change, the larger the N2 amplitude change). Another additional ANOVA indicated a significant main effect of matching (no-imagery vs. match) [$F_{(1,\ 23)} = 4.60$, $p < 0.05$] and an interaction effect between matching (no-imagery vs. match) and anterior-posterior electrodes [$F_{(6,\ 138)} = 4.47$, $p < 0.05$, $\varepsilon = 0.27$]. *Post-hoc* analyses revealed that the no-imagery condition was more negative than the match condition and this effect was broadly distributed along the fronto–parietal areas ($Ps < 0.05$), and was maximal at the fronto-central areas, but not at the parieto-occipital and occipital areas ($p = 0.36$ and 0.52, respectively). Correlation analysis revealed a significant negative correlation between RT change (no imagery vs. match) and the N2 amplitude change (no imagery-match) over fronto–parietal sensors ($r = -0.422$, $p < 0.05$). There was no significant difference between mismatch and no imagery conditions ($p = 0.39$) and no interaction effect between imagery (mismatch and no imagery) and anterior-posterior electrodes ($p = 0.55$).

DISCUSSION

The present study investigated the effects of mental imagery on subsequent face identification. Behaviorally, participants responded faster in the match condition than in the no-imagery condition, with the slowest RT in the mismatch condition, which provides further evidence supporting the hypothesis that the imagery-perception interaction is content-specific in an identification task. In our experiment, we demonstrated the content-specific effects extending to complex visual stimuli, human faces in this case, compared to the simple visual objects and features used previously (Finke, 1986). The results of the imagery vividness questionnaire suggest that the subjects successfully executed the visual imagery. Participants pressed the "Space" bar more rapidly to the no-imagery cue than to face-imagery cues, providing further evidence for the execution of the visual imagery task as required.

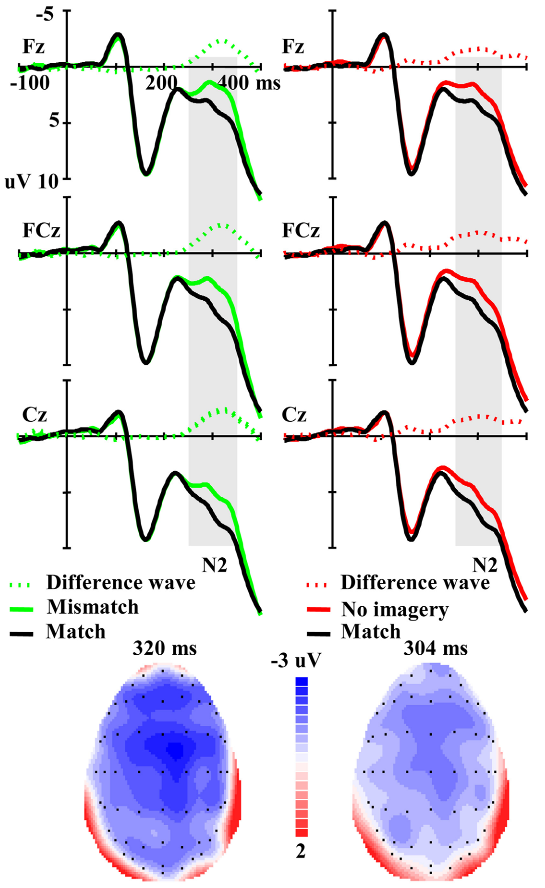


FIGURE 4 | The left panel illustrates the grand averaged ERPs in mismatch and match conditions and their difference wave. The right panel illustrates the grand averaged ERPs in no imagery and match conditions and their difference wave. The gray areas highlight the time windows of N2 (250–350 ms) used for statistical analysis. The scalp distributions are time-locked to the peak amplitude of the difference wave (Left: Mismatch minus Match; Right: No imagery minus Match).

Consistent with these behavioral measures, ERP results showed that, compared with the match condition, both mismatch and no-imagery conditions elicited a higher amplitude of the fronto-central maximal N2 (250–350 ms window). No significant difference was observed for the P1 and N170 components. The observation of changes in the later N2 component but not in early visual processing components provides neural evidence supporting the *matching level hypothesis*: imagined visual stimuli affected the matching stage, rather than the earlier perceptual processing stages in the visual identification task. Pearson correlation analyses revealed that the longer the RT change, the larger the N2 amplitude change from match to mismatch/no imagery condition, but not for P1 or N170 amplitude change, providing further evidence for the *matching level hypothesis*.

The mismatch condition elicited higher amplitude of the fronto-central maximal N2 during the 250–350 ms window, and this result is consistent with previous findings using repetition paradigms. The fronto-central N2 component at this time range is usually elicited under the S1–S2 matching task, where the S1 can be presented physically or formed from imagery (Wu et al., 2010). This component has been generally considered to be sensitive to mismatch from visual templates (for a review, see Folstein and Van Petten, 2008). Wang et al. (2001) have revealed that the mismatch-related N2 was elicited independently of task demands. Results from studies using repetition paradigms also revealed that, compared with the primed target, the unprimed target (mismatch) elicited a more negative anterior-distributed negativity around 250–350 ms after stimulus onset (Eddy et al., 2006; Eddy and Holcomb, 2011). According to the model of face recognition (Bruce and Young, 1986), a sense of familiarity is achieved, if the bottom-up sensory process matches the stored structure code retrieved from memory. The mental imagery of a face, in this study, is a top-down process that forms an internal perceptual template of a face representation (Kosslyn et al., 2006; Wu et al., 2011), and this perceptual template provides a recent context for the subsequent visual target. When the features of visual stimuli match the pre-activated face template from preceding mental imagery, the identification can be achieved faster, resulting in the behavioral benefits in the matching condition. Whereas when the internal image is incompatible with the target, as in the mismatch condition, more time and higher neural cost is required to resolve the conflict, manifested in a behavioral interference effect and larger N2 amplitude.

Our ERP results also revealed that the fronto-central distributed N2 was observed in the no-imagery condition. There might be two cognitive mechanisms for this result, and both of them can be considered occurring at the matching stage of face identification. The first one is the mismatch interpretation of this N2 (for a review, see Folstein and Van Petten, 2008). Theories have been proposed that our perception of external objects and events is the result of bottom-up process meeting the consistent top-down prediction (Hochstein and Ahissar, 2002). That is, the automatic predictive process is also available in the no-imagery condition but provides less specific predictions for the coming stimuli compared with the matching imagery. Such under-specified predictions would induce greater mismatches with the bottom-up process and hence elicit larger fronto-central N2 and longer RT. The second one is to explain this N2 as the effect of immediate context. Besides the mismatch interpretation, the fronto-central N2 has also been interpreted as a mechanism of novelty detection (for a review, see Folstein and Van Petten, 2008). Previous studies suggested that both long-term context and immediate context contribute to the novelty response (Daffner et al., 2000a,b). In the current study, the preceding imagery process of match trials provided immediate context for the subsequently presented face of the same trial, thus the presented face is relatively more novel in no-imagery trials than in the match trials; this novelty effect within the trial or in a short-term/immediate context elicited the N2 in the no-imagery condition. The matching stage of face identification in no-imagery trials has to depend on the mechanism of the long-term memory system (Grützner et al., 2010), then more time and higher neural cost is required as compared with that in the match condition during which the short-term memory trace can be accessed from preceding mental imagery.

It might be argued that the difference between conditions can also be interpreted as an Late Positive Complex (LPC) effect but not the N2 effect, i.e., the match condition elicited more positive potentials than both the mismatch and no-imagery conditions. The fronto-central LPC has been related to the old/new effect where old stimuli elicit more positive LPC than do new stimuli, and further studies suggested that this fronto-central old-new effect was an index of familiarity (for a review, see Rugg and Curran, 2007). In the current study, we interpreted that the no imagery condition was relatively more novel in a short-term context, and thus elicited more negative fronto-central N2 than the match condition and we interpreted it as a mechanism of novelty detection, i.e., novel faces elicit more negative N2. Thus, the trend between conditions was the same for both components/interpretations. More importantly, both components have the same scalp distribution and similar time window. Thus, we argue that both components and interpretations are not contrary and may in fact reflect similar cognitive processing. For the difference between match and mismatch conditions, we focus on the mismatch interpretation because the current experimental design is quite different from that used in research on the old/new effect.

The distinct temporal profiles of neural correlates suggest that task demand may be an important factor determining the imagery-perception interaction in the hierarchical visual process. The present study revealed that the effects of mental imagery on visual identification occur at a later matching stage, where feature information from bottom-up processes is matched with that from top-down processes to achieve recognition. Thus the *matching level hypothesis* is supported in a visual identification task. Previous ERP studies provided neural evidence that, during a detection task or without an active task, imagery-perception interaction occurs at the earlier perceptual stage of visual processing, thus supporting the *perceptual level hypothesis* (Farah et al., 1988; Ganis and Schendan, 2008). Task demand modulates not only the direction that imagery affects perception (Finke, 1986) but also the cognitive stage(s) at which this influence occurs.

Some limitations should be noted. The first limitation is that only face stimuli were used in this study. A previous study by Ganis and Schendan (2008) demonstrated that face imagery affected the early perceptual processing of subsequent test faces (no task on the test faces) and such effects could generalize to other object categories. The choice of only including face stimuli in our study is because of the emphasis on testing the perceptual and matching level hypotheses in the context of face identification. But without stimuli from other categories as comparison conditions, we cannot generalize the observed effects of visual imagery to other perceptual categories. The second limitation is that only two tokens of face stimuli were used in this study. This choice is because, arguably, fewer tokens to imagine would result in more vivid visual reconstructions after equal amounts of training, and hence increase the effect sizes of mental imagery on subsequent perception. However, participants could adapt strategies by visualizing either specific parts/features or the global configuration of a face, thus it is not clear which processing mechanisms would account for the observed effects of mental imagery on visual face perception. The third limitation is that the 100 ms pre-face stimulus baseline might bring in inequity between the imagery and no-imagery conditions. Although in both the current study and previous studies (Farah et al., 1988; Ganis and Schendan, 2008) participants were asked to press a button to separate the mental imagery processing period and the following perception period, participants could inertially perform visual imagery and the visual trace was still available during the pre-face stimulus interval for the imagery conditions, but not for the no-imagery condition. Such potential overlap in the imagery conditions may influence the ERP baseline and result in an amplitude shift in the ERPs to the subsequent face stimuli.

In conclusion, the effect of mental imagery on subsequent face identification is content-specific, i.e., mismatch between the presented and imagined face leads to an interference and match leads to a facilitation effect. The ERP results suggested that both facilitation and interference effects in a face identification task occur at a later matching stage, but not in the early perceptual processing.

ACKNOWLEDGMENTS

This work was supported by the NSF China (30900442) and the Project for Young Scientists Fund, IP, CAS (O9CX042004). We acknowledge Richard Carciofo for his editorial assistance.

REFERENCES

Bruce, V., and Young, A. (1986). Understanding face recognition. *Br. J. Psychol.* 77, 305–327.

Cabeza, R., Burton, A. M., Kelly, S. W., and Akamatsu, S. (1997). Investigating the relation between imagery and perception: evidence from face priming. *Q. J. Exp. Psychol. A* 50, 274–289.

Craver-Lemley, C., and Arterberry, M. E. (2001). Imagery-induced interference on a visual detection task. *Spat. Vis.* 14, 101–119.

Craver-Lemley, C., Arterberry, M. E., and Reeves, A. (1997). Effects of imagery on vernier acuity under conditions of induced depth. *J. Exp. Psychol. Hum. Percept. Perform.* 23, 3–13.

Craver-Lemley, C., and Reeves, A. (1987). Visual imagery selectively reduces vernier acuity. *Perception* 16, 599–614.

Daffner, K. R., Mesulam, M. M., Scinto, L. F. M., Calvo, V., Faust, R., and Holcomb, P. J. (2000a). An electrophysiological index of stimulus unfamiliarity. *Psychophysiology* 37, 737–747.

Daffner, K. R., Scinto, L. F. M., Calvo, V., Faust, R., Mesulam, M. M., West, W. C., et al. (2000b). The influence of stimulus deviance on electrophysiologic and behavioral responses to novel events. *J. Cogn. Neurosci.*12, 393–406.

Djordjevic, J., Zatorre, R. J., and Jones-Gotman, M. (2004a). Effects of perceived and imagined odors on taste detection. *Chem. Senses* 29, 199–208.

Djordjevic, J., Zatorre, R. J., Petrides, M., and Jones-Gotman, M. (2004b). The mind's nose: effects of odor and visual imagery on odor detection. *Psychol. Sci.* 15, 143–148.

Eddy, M. D., and Holcomb, P. J. (2011). Invariance to rotation in depth measured by masked repetition priming is dependent on prime duration. *Brain Res.* 1424, 38–52.

Eddy, M., Schmid, A., and Holcomb, P. J. (2006). Masked repetition priming and event-related brain potentials: a new approach for tracking the time-course of object perception. *Psychophysiology* 43, 564–568.

Farah, M. J. (1985). Psychophysical evidence for a shared representational medium for mental images and percepts. *J. Exp. Psychol. Gen.* 114, 91–103.

Farah, M. J. (1989). Mechanisms of imagery-perception interaction. *J. Exp. Psychol. Hum. Percept. Perform.* 15, 203–211.

Farah, M. J., Peronnet, F., Gonon, M. A., and Giard, M. H. (1988). Electrophysiological evidence for a shared representational medium for visual images and visual percepts. *J. Exp. Psychol. Gen.* 117, 248–257.

Finke, R. A. (1986). Some consequences of visualization in pattern identification and detection. *Am. J. Psychol.* 99, 257–274.

Folstein, J. R., and Van Petten, C. (2008). Influence of cognitive control and mismatch on the N2 component of the ERP: a review. *Psychophysiology* 45, 152–170.

Freyd, J. J., and Finke, R. A. (1984). Facilitation of length discrimination using real and imagined context frames. *Am. J. Psychol.* 97, 323–341.

Ganis, G., and Schendan, H. E. (2008). Visual mental imagery and perception produce opposite adaptation effects on early brain potentials. *Neuroimage* 42, 1714–1727.

Gilbert, C. D. (1996). Plasticity in visual perception and physiology. *Curr. Opin. Neurobiol.* 6, 269–274.

Gong, X., Huang, Y. X., Wang, Y., and Luo, Y. J. (2011). Revision of the Chinese facial affective picture system. *Chin. Ment. Heal. J.* 25, 40–46.

Grützner, C., Uhlhaas, P. J., Genc, E., Kohler, A., Singer, W., and Wibral, M. (2010). Neuroelectromagnetic correlates of perceptual closure processes. *J. Neurosci.* 30, 8342–8352.

Henson, R. N., Goshen-Gottstein, Y., Ganel, T., Otten, L. J., Quayle, A., and Rugg, M. D. (2003). Electrophysiological and haemodynamic correlates of face perception, recognition and priming. *Cereb. Cortex* 13, 793–805.

Hochstein, S., and Ahissar, M. (2002). View from the top: hierarchies and reverse hierarchies in the visual system. *Neuron* 36, 791–804.

Ishai, A., and Sagi, D. (1995). Common mechanisms of visual imagery and perception. *Science* 268, 1772–1774.

Ishai, A., and Sagi, D. (1997a). Visual imagery facilitates visual perception: psychophysical evidence. *J. Cogn. Neurosci.* 9, 476–489.

Ishai, A., and Sagi, D. (1997b). Visual imagery: effects of short- and long-term memory. *J. Cogn. Neurosci.* 9, 734–742.

Itier, R. J., and Taylor, M. J. (2004). N170 or N1? Spatiotemporal differences between object and face processing using ERPs. *Cereb. Cortex* 14, 132–142.

Kosslyn, S. M., Thompson, W. L., and Ganis, G. (2006). *The Case for Mental Imagery*, New York, NY: Oxford University Press.

McDermott, K. B., and Roediger, H. L. 3rd. (1994). Effects of imagery on perceptual implicit memory tests. *J. Exp. Psychol. Learn. Mem. Cogn.* 20, 1379–1390.

Michelon, P., and Koenig, O. (2002). On the relationship between visual imagery and visual perception: evidence from priming studies. *Eur. J. Cogn. Psychol.* 14, 161–184.

Moulton, S. T., and Kosslyn, S. M. (2009). Imagining predictions: mental imagery as mental emulation. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 364, 1273–1280.

Neisser, U. (1976). *Cognition and Reality*. San Francisco, CA: W. H. Freeman.

Pearson, J., Clifford, C. W., and Tong, F. (2008). The functional impact of mental imagery on conscious perception. *Curr. Biol.* 18, 982–986.

Perky, C. W. (1910). An experimental study of imagination. *Am. J. Psychol.* 21, 422–452.

Peterson, M. J., and Graham, S. E. (1974). Visual detection and visual imagery. *J. Exp. Psychol.* 103, 509–514.

Pylyshyn, Z. W. (2002). Mental imagery: in search of a theory. *Behav. Brain Sci.* 25, 157–182.

Reeves, A. (1981). Visual-imagery lowers sensitivity to hue-varying, but not to luminance-varying, visual stimuli. *Percept. Psychophys.* 29, 247–250.

Rugg, M. D., and Curran, T. (2007). Event-related potentials and recognition memory. *Trends Cogn. Sci.* 11, 251–257.

Sagiv, N., and Bentin, S. (2001). Structural encoding of human and schematic faces: holistic and part-based processes. *J. Cogn. Neurosci.* 13, 937–951.

Segal, S. J., and Fusella, V. (1970). Influence of imaged pictures and sounds on detection of visual and auditory signals. *J. Exp. Psychol.* 83, 458–464.

Semlitsch, H. V., Anderer, P., Schuster, P., and Presslich, O. (1986). A solution for reliable and valid reduction of ocular artifacts applied to the P300 ERP. *Psychophysiology* 23, 695–703.

Sreenivasan, K. K., Katz, J., and Jha, A. P. (2007). Temporal characteristics of top-down modulations during working memory maintenance: an event-related potential study of the N170 component. *J. Cogn. Neurosci.* 19, 1836–1844.

Thierry, G., Martin, C. D., Downing, P., and Pegna, A. J. (2007). Controlling for interstimulus perceptual variance abolishes N170 face selectivity. *Nat. Neurosci.*10, 505–511.

Wang, H., Wang, Y., Kong, J., Cui, L., and Tian, S. (2001). Enhancement of conflict processing activity in human brain under task relevant condition. *Neurosci. Lett.* 298, 155–158.

Webster, M. A., and MacLeod, D. I. (2011). Visual adaptation and face perception. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 366, 1702–1725.

Wu, J., Mai, X., Yu, Z., Qin, S., and Luo, Y. J. (2010). Effects of discrepancies between imagined and perceived sound on the N2 component of the event-related potential. *Psychophysiology* 47, 289–298.

Wu, J., Yu, Z., Mai, X., Wei, J., and Luo, Y. J. (2011). Pitch and loudness information encoded in auditory imagery as revealed by event-related potentials. *Psychophysiology* 48, 415–419.

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

Received: 29 April 2012; accepted: 23 October 2012; published online: 08 November 2012.

Citation: Wu J, Duan H, Tian X, Wang P and Zhang K (2012) The effects of visual imagery on face identification: an ERP study. Front. Hum. Neurosci. ***6****:305. doi: 10.3389/fnhum.2012.00305*

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 14 September 2012
doi: 10.3389/fpsyg.2012.00329

Electrophysiological potentials reveal cortical mechanisms for mental imagery, mental simulation, and grounded (embodied) cognition

Haline E. Schendan[1,2]* **and Giorgio Ganis**[1,2,3]

[1] School of Psychology, University of Plymouth, Plymouth, Devon, UK
[2] Massachusetts General Hospital, Athinoula A. Martinos Center for Biomedical Imaging, Charlestown, MA, USA
[3] Department of Radiology, Harvard Medical School, Boston, MA, USA

Edited by:
Joel Pearson, The University of New South Wales, Australia

Reviewed by:
Fred Mast, Universität Bern, Switzerland
Karla K. Evans, Harvard Medical School, USA
Adam Reeves, Northeastern University, USA

***Correspondence:**
Haline E. Schendan, School of Psychology, University of Plymouth, Drake Circus, Plymouth, Devon, PL4 8AA, UK.
e-mail: haline.schendan@plymouth.ac.uk

Grounded cognition theory proposes that cognition, including meaning, is grounded in sensorimotor processing. The mechanism for grounding cognition is mental simulation, which is a type of mental imagery that re-enacts modal processing. To reveal top-down, cortical mechanisms for mental simulation of shape, event-related potentials were recorded to face and object pictures preceded by mental imagery. Mental imagery of the identical face or object picture (congruous condition) facilitated not only categorical perception (VPP/N170) but also later visual knowledge [N3(00) complex] and linguistic knowledge (N400) for faces more than objects, and strategic semantic analysis (late positive complex) between 200 and 700 ms. The later effects resembled semantic congruity effects with pictures. Mental imagery also facilitated category decisions, as a P3 peaked earlier for congruous than incongruous (other category) pictures, resembling the case when identical pictures repeat immediately. Thus mental imagery mimics semantic congruity and immediate repetition priming processes with pictures. Perception control results showed the opposite for faces and were in the same direction for objects: Perceptual repetition adapts (and so impairs) processing of perceived faces from categorical perception onward, but primes processing of objects during categorical perception, visual knowledge processes, and strategic semantic analysis. For both imagery and perception, differences between faces and objects support domain-specificity and indicate that cognition is grounded in modal processing. Altogether, this direct neural evidence reveals that top-down processes of mental imagery sustain an imagistic representation that mimics perception well enough to prime subsequent perception and cognition. Findings also suggest that automatic mental simulation of the visual shape of faces and objects operates between 200 and 400 ms, and strategic mental simulation operates between 400 and 700 ms.

Keywords: mental imagery, visual shape perception, object categorization, face identification, semantic memory priming, visual knowledge, embodiment and grounded cognition, event-related potential

INTRODUCTION

Mental imagery is the ability to reactivate and manipulate modality-specific mental representations when current sensory stimulation or overt motor action is absent, and this ability can be associated with the subjective experience of perceiving or acting within one's mental world (e.g., "seeing with the mind's eye"). Mechanisms of mental imagery have been proposed in large scale network theories (Kosslyn et al., 2006). At an abstract level, such imagistic theories propose that, during mental imagery, modality-specific, long-term memory is reactivated in a top-down manner and maintained via working memory processes so that they can be inspected and manipulated to achieve a task goal (Kosslyn, 1994; Kosslyn et al., 2001; Ganis et al., 2003; Ganis and Schendan, 2011). Notably, the mechanisms proposed in these theories of mental imagery resemble those in grounded (embodied) cognition theory, which proposes that cognition is grounded in modal processing of sensorimotor information and introspective states (e.g., emotion, motivation, intention; Pulvermuller, 1999; Barsalou, 2008). Like imagistic theories of mental imagery, grounded cognition theory challenges the dominant symbol systems paradigm inspired by formal theories of logic, language, and computation that proposes that amodal symbol representations, which are independent from the sensorimotor processes, support language, thinking, attention, memory, and meaning (Fodor, 1983; Johnson-Laird, 1983; Pylyshyn, 2003). Recently, however, tests of grounded cognition theory have yielded compelling evidence that modal processing affects cognition, including meaning, even when task irrelevant, and *vice versa* (e.g., Tucker and Ellis, 1998; Wilson, 2002; Vigliocco et al., 2006; Barsalou, 2008; Fischer and Zwaan, 2008; Kemmerer et al., 2008; Chatterjee, 2010; Anderson et al., 2011; Hirschfeld and Zwitserlood, 2011; Meteyard et al., 2011). However, little is yet known about when, how, and how much cognition is grounded and about the brain mechanisms (Mahon and Caramazza, 2009; Chatterjee, 2010; Rumiati et al., 2010; Meteyard et al., 2011). The main proposal for how cognition is grounded is mental simulation, defined as the "re-enactment of perceptual, motor, and introspective states acquired during experience with the world, body, and mind" (Barsalou, 2008, 2009). The present experiment aims to reveal the cortical dynamics of mental imagery mechanisms that may ground cognition in mental simulations. Crucially, the cortical dynamics of mental imagery and mental simulations that ground cognition are almost entirely unknown because electromagnetic brain sensing methods, which reveal neural activity directly with the required high time precision [milliseconds (ms)], have not been applied, as done here.

An important distinction in grounded cognition theory to consider is that between non-conscious automatic simulations implicated, for example, in constructing meaning from language (Bub and Masson, 2010; Wassenburg and Zwaan, 2010) and conscious effortful simulations, such as mental imagery, that result from conscious representations constructed in working memory strategically (Kosslyn et al., 2006; Barsalou, 2008, 2009). Critically for the present study, both types of simulation share common representations (Barsalou, 2008; Moulton and Kosslyn, 2009). We hypothesize that, at the level of brain mechanisms, the top-down feedback mechanisms that support automatic simulation are a subset of those that support mental imagery (and conscious effortful simulations; Ganis and Schendan, 2011). Specifically, we propose that non-conscious automatic simulations unfold via reflexive top-down signals from higher to lower level areas along modal information processing pathways (e.g., the ventral stream for processing visual objects): Perceiving a stimulus triggers these processes reflexively. These same processes are triggered by effortful, strategic, task-oriented, top-down signals from the prefrontal cortex during mental imagery (Ganis and Kosslyn, 2007), which also triggers them most strongly due to the deliberate, targeted nature of the task requirements. Thus, studying mental imagery provides a powerful way to reveal the maximum set of top-down feedback mechanisms that support mental simulation, including non-conscious automatic simulation.

To ensure that mental imagery mechanisms underlie the effects, this experiment used a validated mental imagery task (Kosslyn et al., 2006). The key task elements are that subjects first memorize pictures of faces of real people or objects extensively and then practice visualizing these pictures mentally. Afterward, during the mental imagery task, the name of the person (or object) cues mental visualization of the associated picture of the face (or object). The critical and novel element of this design is that, after subjects report (by pressing a key) that they have generated a vivid mental image, a target picture appears 200 ms later. This picture is either identical to the picture that they had learned to visualize (*congruous condition*) or different from it, being from the opposite category (i.e., if a face was visualized, an object is shown, and *vice versa: incongruous condition*). The impact of the imagined picture on the target picture is used to define the cortical dynamics of mental imagery. The advantage to comparing mental imagery of these two categories is that faces and objects recruit different posterior visual processing areas (Hasson et al., 2003; Downing et al., 2006) and are associated with distinct ERP signatures (e.g., Schendan et al., 1998; Allison et al., 1999; Puce et al., 1999). This experiment thus reveals when top-down processes for mental imagery of visual shape can ground cognition of faces and other objects. Such neurophysiological markers will be crucial for future work on when, how, and how much top-down cortical processes of mental simulation ground cognition.

Findings from this experiment have been reported already for early ERPs before 200 ms (Ganis and Schendan, 2008), that is, the vertex positive potential and its occipitotemporal N170 counterpart (VPP/N170), which are associated with categorical perception. The goal here was to analyze the later ERPs that were not analyzed previously and are associated with knowledge, meaning, and category decisions. These abilities are the main targets for grounded cognition explanations. The main hypothesis is that, if mental simulation constructs cognition, including meaning, then mental simulation of modal processes induced by mental imagery (and the associated cortical dynamics revealed by the ERPs) should resemble those associated with cognitive and semantic effects. If true, then this would constitute crucial evidence that links mental simulation of modal processing (using top-down mental imagery mechanisms) with cognition and meaning, as proposed in grounded cognition theory. To assess cognitive effects, this study capitalizes on ERP markers of face and object cognition (Neumann and Schweinberger, 2008; Schendan and Maher, 2009). To assess semantic effects, in particular, this study capitalizes also on well-studied semantic congruity and priming effects. These effects are thought to reflect processing meaning in the semantic memory system (Rossell et al., 2003) and to operate via the same automatic top-down processes implicated in automatic mental simulation (Collins and Loftus, 1975; Franklin et al., 2007; Kutas and Federmeier, 2011).

To predict the specific pattern of ERP cognitive, semantic congruity, and priming effects that mental imagery could produce, this report capitalizes on the two-state interactive (2-SI) account of the brain basis of visual object cognition (Schendan and Kutas, 2003, 2007; Schendan and Maher, 2009) and extends it into a multi-state interactive (MUSI) account. This framework proposes that posterior object processing areas activate at multiple times in brain "states" serving distinct functions. State 1: Initial activity in object processing areas feeds forward from occipital to temporal cortex between ∼120 and ∼200 ms when a visual object is coarsely perceptually categorized, indexed by the VPP/N170 (Schendan et al., 1998; Schendan and Lucia, 2009). State 2: Object processing areas activate again but in a sustained, interactive manner dominated by feedback and recurrent processing among these areas and ventrolateral prefrontal cortex (VLPFC) between ∼200 and 500 ms, indexed by a frontal N3(00) complex (often labeled frontal N400). The N3 is specific to processing pictures of a face or object and non-linguistic information (e.g., shape; Barrett and Rugg, 1989, 1990; McPherson and Holcomb, 1999; Nessler et al., 2005) and is the first ERP to modulate according to visual cognitive factors that similarly affect object processing areas and VLPFC, including semantic memory (Barrett and Rugg, 1990; Zhang et al., 1995; Doniger et al., 2000, 2001; Schendan and Kutas, 2002, 2003, 2007; Philiastides and Sajda, 2006, 2007; Philiastides et al., 2006; Schendan and Lucia, 2009, 2010; Schendan and Maher, 2009). For example, the N3 becomes more negative with greater mental rotation

(Schendan and Lucia, 2009), worse speed and accuracy of category decisions, greater stimulus atypicality and impoverishment (e.g., atypical views, visual degradation; Doniger et al., 2000; Schendan and Kutas, 2002, 2003; Johnson and Olshausen, 2003), and implicit memory (repetition priming, i.e., better decisions for repeated than new objects; Henson et al., 2004; Schendan and Kutas, 2007) primarily for meaningful objects (e.g., dog; Voss et al., 2010). Such effects also localize to object processing areas (David et al., 2005, 2006; Philiastides and Sajda, 2006, 2007; Sehatpour et al., 2008; Schendan and Maher, 2009; Schendan and Lucia, 2010; Clarke et al., 2011). Later in state 2, the well-studied, centroparietal N400 between 300 and 500 ms reflects interactive activation of semantic memory in anterior temporal cortex and VLPFC, especially in response to words (i.e., linguistic stimuli; Marinkovic et al., 2003; Lau et al., 2008; Kutas and Federmeier, 2011). Both mid-latency negativities (N3 and N400) are more negative for stimuli that are incongruous (i.e., non-matching) relative to congruous (i.e., matching) with the immediately preceding semantic context based on a sentence or scene (Ganis et al., 1996). A similar account has been proposed for face cognition (Neumann and Schweinberger, 2008; Burton et al., 2011).

The MUSI account revises this story by adding State 3 that operates after about 400 to 500 ms and performs internal evaluation and verification processes, including conscious effortful mental simulations. For example, later verification of category decisions, more complex semantic processes, and episodic memory have been associated with a posterior late positive complex (LPC) after ∼500 ms (Schendan and Kutas, 2002; Rugg and Curran, 2007; Voss et al., 2010) that is, instead, more positive to incongruous than congruous semantic contexts with objects, videos, and faces (Ganis et al., 1996; Schendan and Kutas, 2002; Ganis and Kutas, 2003; Sitnikova et al., 2009, 2010). Altogether, this predicts that the N3 and N400 will be more negative and the LPC will be more positive when mental imagery does not match the current picture (incongruous) than when it does match (congruous). Finally, cortical sources of the ERPs to faces and objects and associated congruity effects should differ because these two categories recruit distinct cortical areas (Hasson et al., 2003).

Prior work on mental imagery and semantic congruity and priming would not necessarily predict such later effects, however. The only prior ERP study with a similar mental imagery task revealed no effects after 300 ms, predicting no effects here (Farah et al., 1988), but that study involved imagery of two letters, which have minimal meaning. In contrast, the present work aims to reveal mental imagery for stimuli with richer semantic content. Studies of semantic congruity and priming effects would also predict no ERP effects of mental imagery after 200 or 300 ms because the slow stimulus timing and cross modal conditions here (i.e., word cue followed by picture) violate standard methods for producing such semantic effects, as detailed in the discussion. To anticipate, results reveal effects not only on the early VPP/N170 during categorical perception, as reported previously (Ganis and Schendan, 2008), but also later ERPs (N3, N400, and LPC) linked to knowledge, meaning, and categorization, as well as a P3(00) linked to immediate repetition priming of perceived faces and objects (Bentin and McCarthy, 1994; Nielsen-Bohlman and Knight, 1994). In addition, we report results of new analyses of data from the perception control experiment conducted in the same group of participants (Ganis and Schendan, 2008) that had not been explored previously (i.e., face ERPs after 500 ms and all object ERPs). As ERPs to faces starting from the early VPP/N170 until 500 ms show perceptual adaptation (i.e., reduced for congruous; Ganis and Schendan, 2008), the N3, N400, and LPC ERPs should likewise show adaptation.

MATERIALS AND METHODS

Participants, materials, design, procedures, and electroencephalographic (EEG) recording methods for these mental imagery and perception control experiments were detailed previously (Ganis and Schendan, 2008). This section summarizes key aspects of the methods for understanding this report. The mental imagery experiment started with memorization of pictures of 11 faces of real people and 11 common real objects across 13 exposures each. Next, participants practiced mentally visualizing each memorized picture three times. For this imagery practice, the name appeared followed by a gray screen during which subjects visualized the picture of the named face (object). Once they had done so, they pressed a key to see the actual picture in order to adjust their mental image. For the mental imagery test (**Figure 1A**), the name appeared for 300 ms followed by a gray screen during which subjects visualized mentally the associated memorized picture. As soon as they had generated this mental image, they pressed a key. Two-hundred milliseconds after the key press the test picture appeared for 300 ms. The test picture was either the picture of the face (object) that was visualized mentally or a picture from the other category [i.e., an object (a face)]. The two by two design included within-subject factors of image congruity (congruous, incongruous) by category type of target picture (face, object). There were 55 trials for each of the four critical combinations (image-picture target pairings) of congruous (face–face; object–object) and incongruous (object-face; face object) conditions. Stimuli were presented on a 21″ cathode-ray tube monitor (1,024 × 768, 150 Hz refresh, Dell P1130) using custom-made *StimPres2.0* software for the *Neurocog* system that ensures precise stimulus time-locking to the EEG (Holcomb, 2003). EEG was recorded at 250 Hz (bandpass .01–100 Hz) from 32 Ag/AgCl electrodes attached to a plastic cap (**Figure 1B**) and electrodes attached via adhesive to the nose, right and left mastoids, underneath the right eye, and lateral to each eye. The perception control experiment with the same subjects was identical, except for the following. There were no memorization and imagery practice sessions, and participants were shown a picture instead of a name before the test picture: The two pictures were shown consecutively with a 200 ms interstimulus interval (ISI), and the first picture was shown until the participant pressed a key to report that they had identified the object (or face). There were 66 trials in each of the four critical conditions. The primary goal of the control experiment was to replicate perceptual adaptation of the early VPP/N170, and so, in order to focus on higher-level face (object) processing in posterior category specific cortex, instead of low-level simple features, the first and second pictures were never identical, even in face–face and object–object (congruous) trials (Ganis and Schendan, 2008).

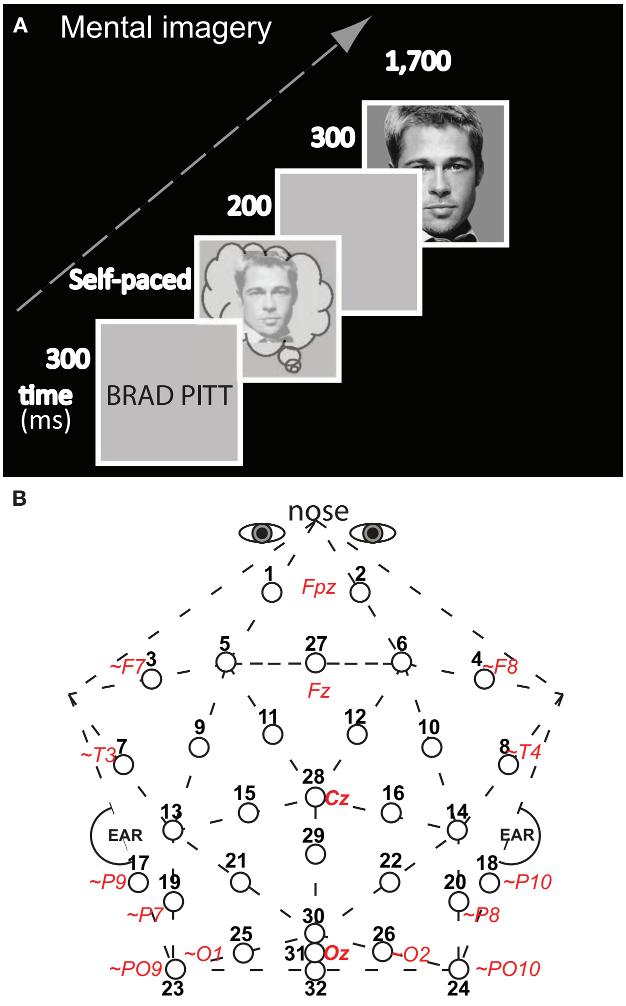


FIGURE 1 | Methods. (A) Diagram of an experimental trial for the test of mental imagery for faces and objects (only a congruous trial of face imagery – face picture is shown). An appropriate face (or object, not shown) was visualized mentally upon seeing the corresponding name (which was on the screen for 300 ms). Subjects pressed a key as soon as they had generated a vivid mental image, and 200 ms after this key press, the test picture appeared, which was either the identical picture that had been imaged mentally (congruous condition) or not (in which case the picture was from the other category: incongruous condition). There was no task on the test pictures. Perception trials (not shown) had a parallel structure to imagery, except that the trial started with perception of a picture of a face (object) was presented until subjects pressed a key as soon as they identified it. **(B)** Thirty-two-channel geodesic montage for EEG recording. Circles show each electrode location with its numerical label. Actual or approximate (∼) locations of 10–10 sites are italicized and in red; posterior sites 23, 24, and 32 lie 1 cm below the inion.

ERP ANALYSES

ERPs were calculated by averaging EEG to each condition, excluding trials with above threshold muscle activity, blinks, eye, and other movement artifacts, time-locking to test picture onset with a 100 ms pre-stimulus baseline. For analyses, ERPs were re-referenced to the mean of both mastoids, and, for visual comparison with prior work, also to the average of all electrodes, except bilateral eyes. Maps of voltage distribution across the head were produced using *EEGLab* software. Eighteen subjects were analyzed; note, due to a scripting error, 1 of the original 19 subjects was not analyzed, but performance results and visual inspection of ERP patterns showed that all results remained the same with and without this subject.

Data were submitted to analyses of variance (ANOVAs) with within-subject factors of congruity and category type (type). For ERPs, ANOVAs assessed separately the lateral pairs (1–26) and midline sites (27–32). Lateral ANOVAs included within-subject factors of Electrode (13 levels) and Hemisphere (left, right). Midline ANOVAs included within-subject factors of Electrode (3 levels) and Site [frontocentral (27–29), occipitoparietal (30–32)]. If Mauchley's test indicated violations of sphericity, the Greenhouse–Geisser correction was applied to the *p*-value. For brevity, only critical Congruity and Type effects are reported, as scalp location effects alone are not of theoretical interest.

Mean ERP amplitudes were measured within time periods after 200 ms chosen based on prior studies. (1) As N3(00) complex components can vary in functional modulation in ∼100 ms time periods between 200 and 500 ms (Schendan and Maher, 2009), analyses assessed separately (a) 200–299 ms (frontopolar P250 and related polarity inverted occipitotemporal Ncl/N250; e.g., Doniger et al., 2000; Federmeier and Kutas, 2002; Sehatpour et al., 2006; Schendan and Lucia, 2010), (b) 300–399 ms (frontal N350; e.g., Schendan and Kutas, 2002, 2003, 2007), and (c) 400–499 ms (frontopolar N450 and frontocentral N390; e.g., Barrett and Rugg, 1989; Ganis and Kutas, 2003; Schendan and Maher, 2009). The latter two times (300–500 ms) also included the N400 (Ganis et al., 1996). (2) The LPC was assessed from 500 to 699 ms (Heil, 2002; Schendan and Lucia, 2009). (3) Continuation of effects was assessed from 700 to 899 ms (Schendan and Maher, 2009).

To isolate effects, focal spatiotemporal analyses were run on sites and times for which the face (object) cognition-, congruity-, or priming-related ERP was maximal and overlapped least with others; these location choices were confirmed and adjusted by visual inspection. (1) From 200 to 299, 300 to 399, and 400 to 499 ms, respectively, (a) pair 1-2 assessed the frontopolar P250, N350, and N450, (b) pair 11-12 assessed the frontocentral N350 and N390, (c) and pair 17-18 assessed their polarity inverted, occipitotemporal counterparts (Scott et al., 2006; Schendan and Maher, 2009). (2) These times (200–499 ms) were also assessed at pair 19-20 for the centroparietal N400 (Ganis et al., 1996); note, 200–299 ms was included as visual inspection suggested an N400 onset before its typical 300 ms start. (3) Centroparietal pair 19-20 also assessed the LPC from 500 to 699 ms. (4) Visual inspection indicated that a P3 peaked earlier for congruous (∼375 ms) than incongruous stimuli (∼500 ms). Consequently, early on, the congruity effect is more positive for congruous than incongruous from 300 to 400 ms, and later, from 400 to 700 ms, the effect is in the opposite direction. To capture this, local positive peak latency (i.e., highest peak within ±20 ms to avoid spurious peaks due to high frequency noise) between 300 and 699 ms was assessed in each condition at midline occipitoparietal site 30, based on the location of similar immediate repetition effects on the P3 (Bentin and McCarthy, 1994). P3 mean amplitude was also assessed from 300 to 399, 400 to 499, and 500 to 699 ms at site 30. (5) Since visual inspection suggested frontal effects continued after 500 ms, frontopolar and frontocentral focal analyses were also run from 500 to 699 and 700 to 899 ms. For focal analyses, which are more precise (albeit less comprehensive) than omnibus analyses, congruity by category interactions were assessed further using planned simple effect tests of congruity for each category condition.

For perception control ERPs, analyses were the same as for imagery, except for the following. Face ERP analyses through 500 ms were already carried out and reported (Ganis and Schendan, 2008) and so not duplicated here. ERPs for faces after 500 ms and those for objects after 200 ms were analyzed separately, and, for brevity in reporting these control data, comparisons between faces and objects are reported only for focal spatiotemporal analyses, and omnibus analyses are not reported, but they confirmed the focal results.

SOURCE ESTIMATION

Source estimation methods evaluated whether distinct sources underlie congruity effects between categories and category differences. The inverse problem of localizing the cortical sources of electromagnetic data recorded from the scalp has no unique solution without additional constraints. Standardized low resolution brain electromagnetic tomography (s*LORETA*) estimates the sources (Pascual-Marqui, 2002) by making a maximum smoothness assumption to compute the three-dimensional (3D) distribution of current density using a standardized, discrete, 3D distributed, linear, minimum norm inverse solution. Localization is data-driven, unbiased (even with noisy data), and exact but has low spatial precision due to smoothing assumptions resulting in highly correlated adjacent cortical volume units. A realistic head model constrains the solution anatomically using the structure of cortical gray matter from the Montreal Neurological Institute (MNI) average of 152 human brains as determined using a probabilistic Talairach atlas.

s*LORETA* software computed the sources of the grand average ERP difference waves using data from all sites, except nose and eyes (Pascual-Marqui, 2002). ERP difference data are analogous to the signal changes between fMRI conditions, and, thus, limit the sources to those that could reflect fMRI differences, and difference waves can reveal weaker sources better (Luck, 2005). Data were analyzed with bandpass filter of 0.01–20 Hz, based on the validated sLORETA analyses reported previously for the VPP/N170 (Ganis and Schendan, 2008). Electrode coordinates were digitized using an infrared digitization system, and imported into *LORETA-Key* software. This coordinate file was converted using sLORETA electrode coordinate conversion tools. The transformation matrix was calculated with a regularization parameter (smoothness) corresponding to the signal-to-noise (SNR) ratio estimated for each difference wave separately at each 100 ms time period of interest from 200 to 900 ms relative to the 100 ms baseline.

RESULTS

These results cover all times and comparisons not analyzed for our prior report on this study, which focused mostly on face ERPs before 200 ms during imagery and perception (Ganis and Schendan, 2008).

MENTAL IMAGERY ERPs

Mental imagery results for faces during the first 500 ms and objects during the first 200 ms were reported previously (Ganis and Schendan, 2008). The new ERP results here reveal congruity effects to faces after 500 ms (**Figure 2**), those to objects after 200 ms (**Figure 3**), and category type effects (**Figures 6** and **7**) and interactions of congruity by type after 200 ms (**Figures 4–6**). For comparison with previous work, ERPs are also plotted with the common average reference (**Figures 8** and **9**). For brevity, (a) only congruity and category type effects, which are of theoretical interest, are reported, (b) degrees of freedom (*df*) are listed only for the first report of each effect, and (c) planned contrasts for omnibus results are not reported, except to note that they supported the corresponding focal spatiotemporal results. For the focal results [all *df*s (1, 13)], any interactions of congruity and category type were followed with corresponding contrast ANOVAs that assessed the congruity effects to objects and faces, separately.

200–500 ms: N3, N400, AND P3

Negativity on the N3 was greater for incongruous than congruous imagery, and the N400 showed this pattern only for faces (**Figures 2–4**). Negativity was greater for objects than faces for the N3, regardless of congruity, and the N400 showed this pattern for congruous imagery but showed the opposite for incongruous imagery (**Figure 6**), consistent with the N400 congruity effect for faces but not objects. Congruity effects had a frontocentral maximum for faces (**Figure 2B**) and a frontopolar maximum instead for objects (**Figure 3B**), and, accordingly, object and face categories differed mainly frontally (**Figure 7**). All effects inverted polarity over occipitotemporal sites. Accordingly, omnibus results (**Table 1**) showed congruity and category type effects, and interactions of congruity by category type from 200 to 400 ms in lateral and midline ANOVAs and also from 400 to 500 ms in lateral ANOVAs, but showed only category type and congruity by category type interactions in midline ANOVAs from 400 to 500 ms when N3 congruity effects to objects ended.

N3 complex

Focal spatiotemporal results (**Table 2**) confirmed frontopolar N3 congruity effects for objects, frontocentral N3 congruity effects for both categories, and centroparietal N400 congruity effects for faces, and occipitotemporal polarity inversion of congruity effects for objects, as well as category type effects. Specifically, results showed type effects on the entire N3 complex from 200 to 500 ms, as frontal negativity and occipitotemporal positivity were greater for objects than faces (**Figure 7**). Congruity was significant at frontopolar and frontocentral sites from 200 to 400 ms, and, at occipitotemporal sites, congruity was marginal from 200 to 300 ms [$F(1, 17) = 3.98$, $p = 0.062$]; note, while occipitotemporal congruity was also significant from 400 to 500 ms, this was due to the start of the posterior LPC. Congruity interacted with type from 200 to 300 ms at frontopolar sites and from 200 to 500 ms at frontocentral sites (**Figures 2** and **3**). Planned simple effects tests of the congruity effect for each type showed that this was because congruity effects were largest for objects at frontopolar sites and largest for faces at frontocentral sites and inverted polarity occipitotemporally for objects. Specifically, **Table 2** shows congruity was

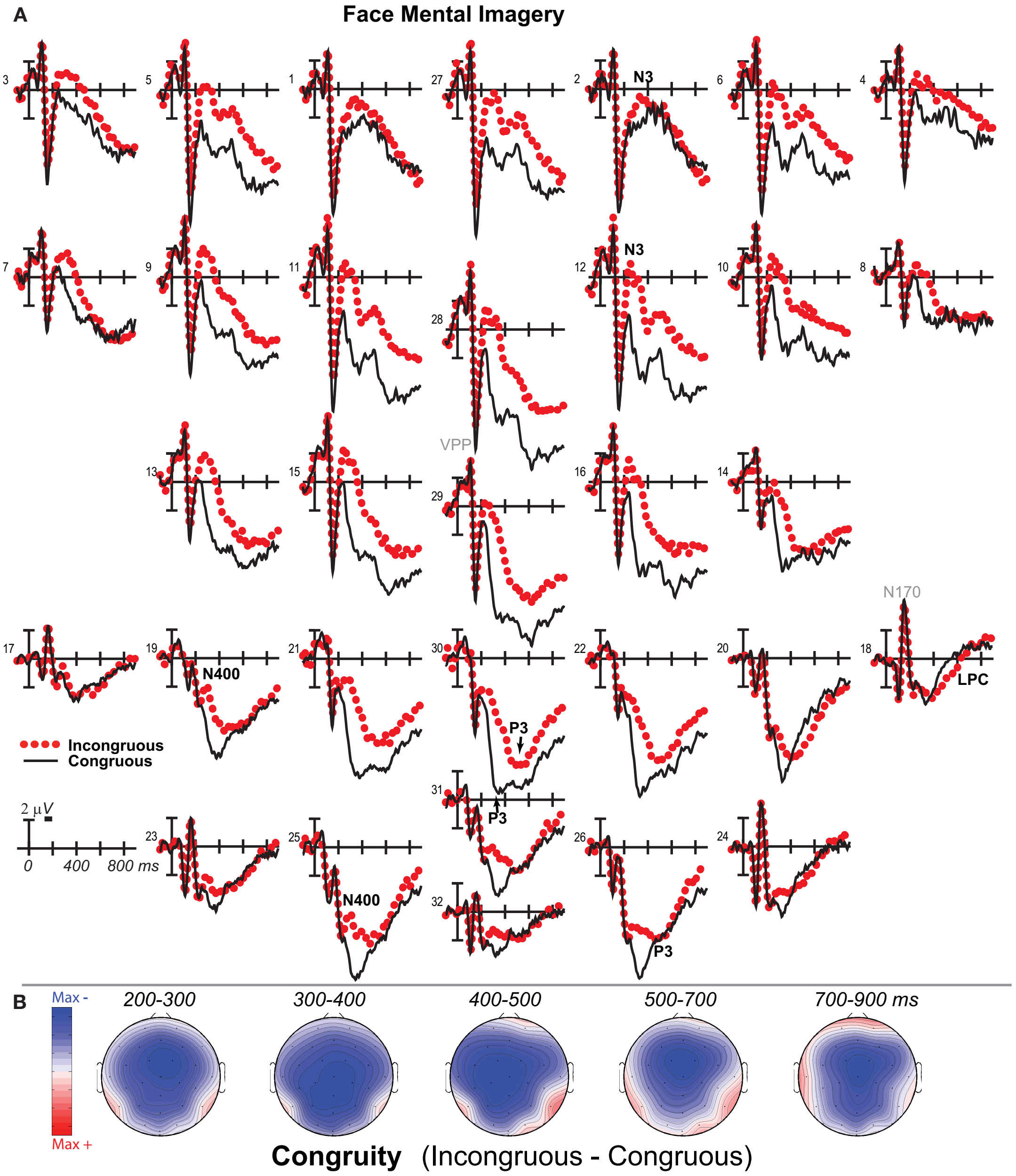


FIGURE 2 | Mental imagery congruity effects for faces. (A) Grand average ERPs between −100 and 900 ms at all 32 channels shown filtered low-pass 30 Hz. **(B)** Maps of voltage across the scalp for difference waves of the ERPs in **(A)**. VPP, vertex positive potential; LPC, late positive complex.

significant at frontopolar sites for objects from 200 to 300 ms and for both category types from 300 to 400 ms, and at frontocentral sites for both from 200 to 400 ms and then later only for faces from 400 to 500 ms. At occipitotemporal sites, congruity was significant only for objects from 200 to 300 ms during the N3 and later from 400 to 500 ms when LPC effects start.

These results and visual inspection indicated that, while frontopolar N3 congruity effects for objects ended by 400 ms, frontocentral congruity effects to faces, as well as type effects at these sites, continued after 500 ms. To assess this, focal analyses of these frontal sites were run on later times. Results (**Table 2**) at frontopolar sites confirmed no congruity effects there after 400 ms and type effects ongoing between 200 and 900 ms. Frontocentral results confirmed that type effects continued, and congruity effects remained larger for faces than objects until 900 ms, and, indeed, were significant only for faces from 400 to 900 ms. In sum, frontopolar and occipitotemporal object congruity effects occurred between 200 and 400 ms, whereas frontocentral face congruity effects occurred from 200 to 900 ms, as did type effects at these times and sites, and occipitotemporal sites also showed LPC effects after 400 ms.

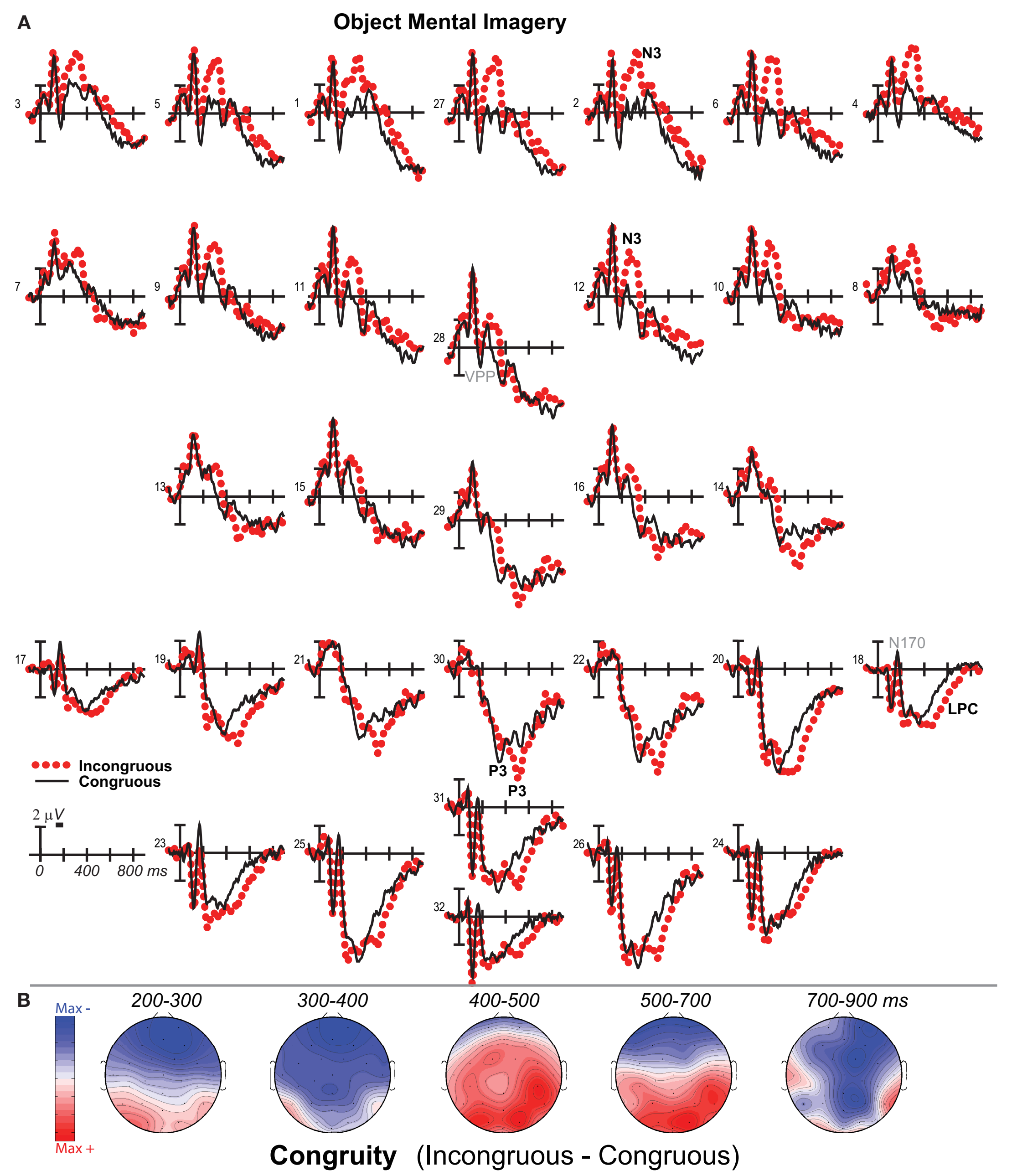


FIGURE 3 | Mental imagery congruity effects for objects. (A) Grand average ERPs between −100 and 900 ms at all 32 channels shown filtered low-pass 30 Hz. **(B)** Maps of voltage across the scalp for difference waves of the ERPs in **(A)** within each analysis time period.

N400

Focal spatiotemporal results (**Table 2**; **Figure 4C**) at centroparietal sites showed a category type effect from 400 to 500 ms, congruity effects from 300 to 500 ms, and congruity by type interactions from 200 to 500 ms. The N400 for faces was larger over the left hemisphere, as demonstrated by three-way interactions of type by congruity by hemisphere that were significant from 300 to 400 ms and marginal before and after [200–300 ms $F(1, 17) = 3.05$, $p = 0.099$; 400–500 ms $F = 3.89$, $p = 0.065$]. Planned simple effects tests (**Table 2**) from 200 to 300 ms showed no N400 effects for faces, and, instead, N3 congruity effects for objects with inverted polarity at these sites (i.e., most negative for congruous) as observed at adjacent occipitotemporal sites (**Figure 3**). Later, from 300 to 400 ms, congruity affected faces in the predicted direction (i.e., most negative for incongruous), as N400 congruity was significant for faces. From 400 to 500 ms, posterior LPC congruity effects began: Congruity was significant for objects and congruity by hemisphere was marginal for faces, as the effect was larger on the right ($F[1, 17] = 4.12$, $p = 0.058$). In sum, centroparietal sites showed occipitotemporal polarity inversion of N3 congruity effects for objects,

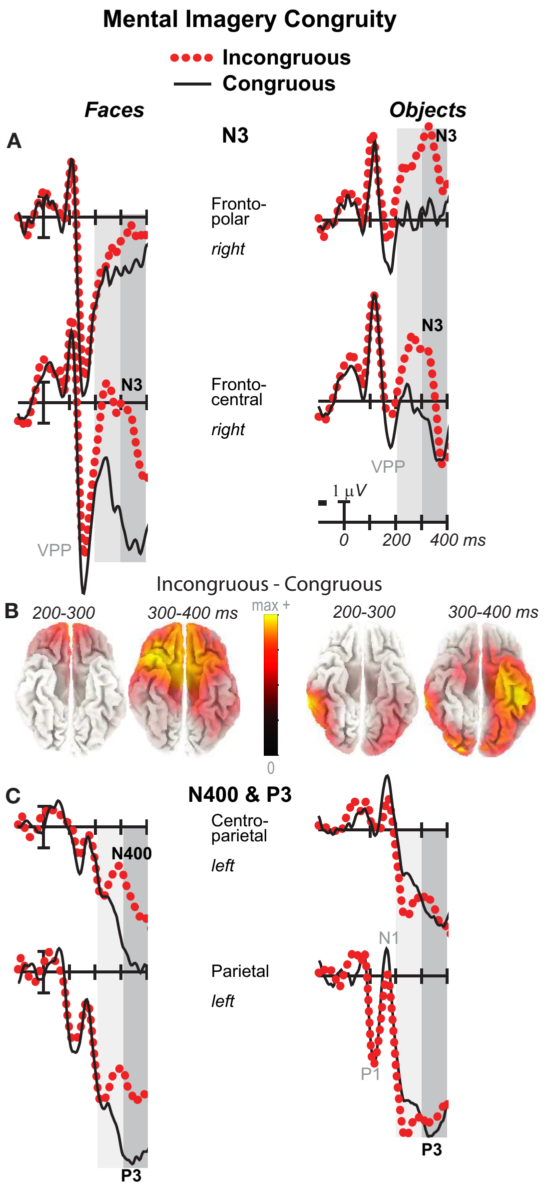


FIGURE 4 | Mental imagery congruity effects on the N3 and N400 and P3. Grand average ERPs between −100 and 400 ms at **(A)** right frontopolar (2) and right frontocentral (12) N3 sites and **(C)** left centroparietal N400 (19) and parietal P3 (25) sites. Filtered low-pass 30 Hz. VPP, vertex positive potential. **(B)** sLORETA sources of congruity difference waves at times of the N3, N400, and P3. Images plot the magnitude of the estimated current density based on the standardized electrical activity in each of 6,239 voxels of 5 mm^3 size.

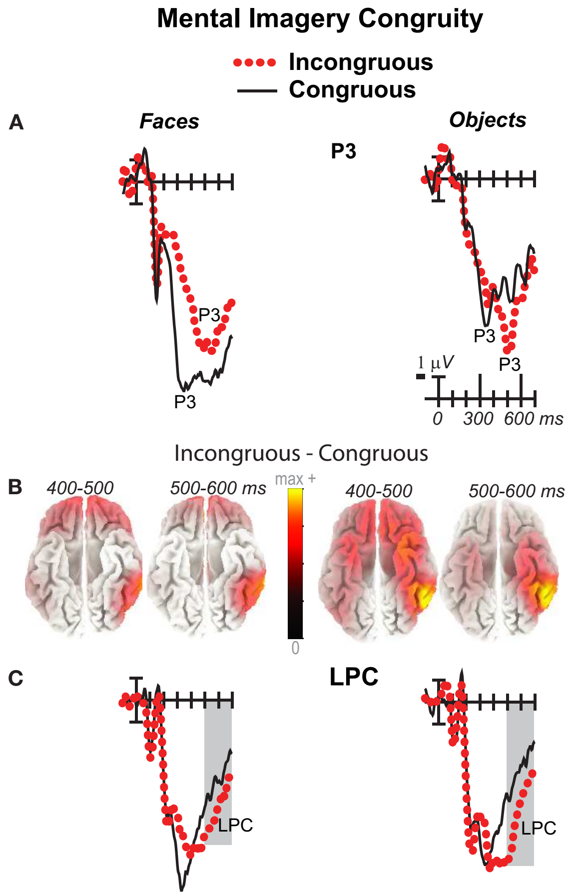


FIGURE 5 | Mental imagery congruity effects on the P3 and late positive complex (LPC). Grand average ERPs between −100 and 700 ms show congruity effects **(A)** on a P3 at a midline parietal site (30) and **(C)** on a late positive complex (LPC) at a right centroparietal site (20). Filtered low-pass 30 Hz. Shading captures the 500–700 ms time when the LPC is maximal. **(B)** sLORETA sources of congruity difference waves at times of the P3 and LPC. Images plot the magnitude of the estimated current density based on the standardized electrical activity in each of 6,239 voxels of 5 mm^3 size.

N400 congruity effects for faces, and, after 400 ms, the start of posterior LPC congruity effects (**Figures 2–4**).

500–900 ms: LPC

As the earlier centroparietal focal results indicated, LPC congruity effects began around 400 ms. Posterior positivity is greater for incongruous than congruous pictures (**Figures 2**, **3**, and **5C**) and for objects than faces (**Figure 7**). Accordingly, omnibus results showed category type and congruity effects from 500 to 900 ms and congruity by type interactions from 500 to 700 ms (**Table 1**), reflecting continuing frontal type effects and frontocentral congruity effects with faces, as reported above for frontal focal results, and continuing LPC type effects until 900 ms and congruity effects until 700 ms. Focal results at centroparietal pair 19-20 confirmed the posterior distribution of the LPC, showing category type and congruity effects (**Table 2**); as congruity and type did not interact, congruity affected both category types. Planned simple effects tests (**Table 2**) showed that congruity effects continued until 700 ms for objects and were in the same direction for faces (i.e., more positive for incongruous) but did not reach significance [500–700 ms: congruity by hemisphere, $F(1, 17) = 2.28$, $p = 0.15$] perhaps due to ongoing frontal congruity effects in the opposite direction that may partly cancel out the posterior LPC effect for faces.

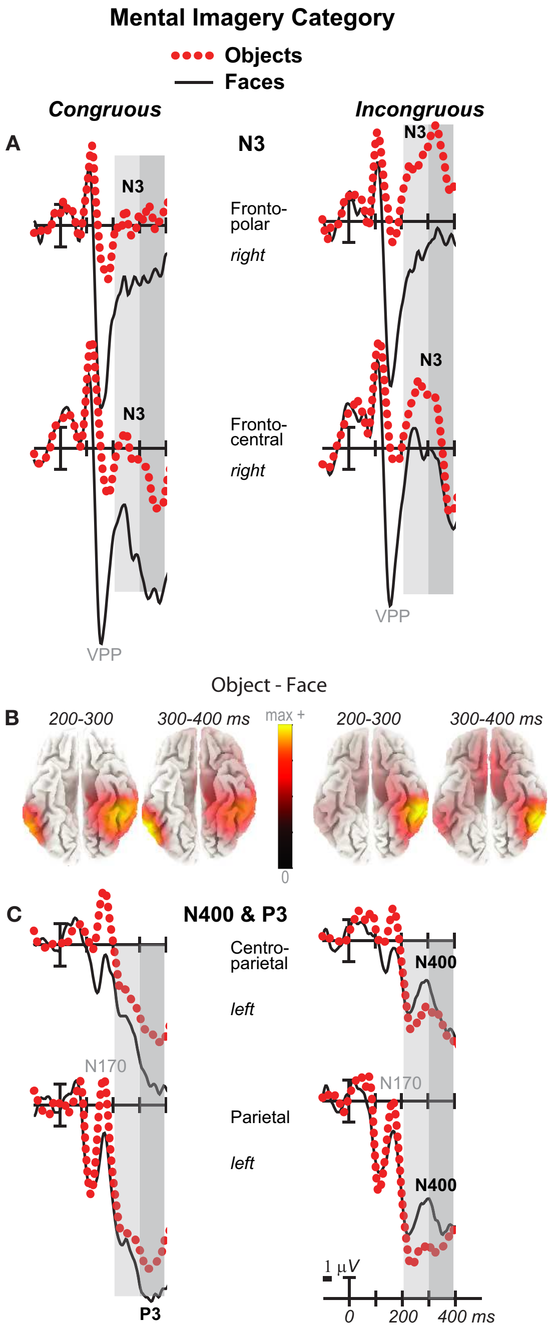


FIGURE 6 | Mental imagery category effects on the N3 and N400. Grand average ERPs between −100 and 400 ms at **(A)** right frontopolar (2) and right frontocentral (12) N3 sites and **(C)** left centroparietal N400 (19) and parietal P3 (25) sites. Filtered low-pass 30 Hz. VPP, vertex positive potential. **(B)** sLORETA sources of category difference waves at times of the N3, N400, and P3. Images plot the magnitude of the estimated current density based on the standardized electrical activity in each of 6,239 voxels of 5 mm^3 size.

P3 PEAK LATENCY AND AMPLITUDE

Visual inspection revealed prominent parietal P3-like peaks between 300 and 700 ms, resembling immediate perceptual repetition priming that makes the P3 earlier and larger (Bentin and McCarthy, 1994). Likewise, here, the P3 appeared to peak earlier for congruous than incongruous stimuli, resulting in a P3 that is more positive for congruous than incongruous initially and then later shows the opposite (**Figures 2**, **3**, and **5A**). Results of ANOVAs on local positive peak latency data at midline occipitoparietal site 30 confirmed that the P3 peaked earlier for congruous (426 ms) than incongruous (496 ms) stimuli [congruity, $F(1, 17) = 11.41$, $p = 0.004$], regardless of category. Peak latency captured the P3 pattern better than mean amplitude due to overlapping N400 and LPC effects. P3 mean amplitude results at site 30 showed only that the P3 was more positive for congruous than incongruous faces from 300 to 700 ms and marginally the opposite (more positive for incongruous than congruous) for objects from 400 to 500 ms; there were significant effects of congruity (300–400 ms, $F = 20.61$, $p < 0.001$) and congruity by type (300–700 ms: $Fs > 13$, $ps < 0.003$) due to congruity being significant for faces (300–500 ms, $Fs > 7.45$, $ps < 0.015$) and marginal for objects (400–500 ms, $F = 3.63$, $p = 0.074$).

MENTAL IMAGERY SOURCES

Because faces and objects recruit distinct occipitotemporal areas (Hasson et al., 2003), the cortical sources of mental imagery should differ between these categories. Estimated cortical sources of each of the four difference waves were consistent with known prefrontal and posterior face (object) processing areas. MNI coordinates are reported for the maximum activated region ($x\ y\ z$) and up to four anatomically distinct sources, the Brodmann's areas (BA) for all, and the BA for up to four secondary sources that are contiguous with the maximum; this captured all clear sources.

Congruity (incongruous–congruous)

For faces, results were consistent with ERP and fMRI evidence for prefrontal and temporal lobe generators during recognition and priming of faces (Henson et al., 2003). Specifically, congruity for faces from 200 to 400 ms during the N3 and N400 (and overlapping P3) reflects sources in medial prefrontal cortex, VLPFC, and superior temporal cortex: **Figure 4B** shows sources observed, from 200 to 300 ms, in medial frontal gyrus [BA9 at 5 45 25 and, from 300 to 400 ms in VLPFC (BA47 at 15 35 −30; BA11)] and superior temporal gyrus (STG; BA38). After 400 ms, P3 and LPC congruity reflect sources in prefrontal and middle temporal cortex: **Figure 5B** shows sources observed in middle temporal gyrus (400–500 ms BA21 at 70 −35 −5; 500–700 ms BA21 at 70 −35 −10) and, after 500 ms, also in medial and superior frontal gyrus (not shown; 500–600 ms, BA9/10 at 0 55 25; 600–700 ms, BA9/10 at 5 60 30).

For objects, N3 congruity effects occurred from 200 to 400 ms and so this time was of primary interest. Results were consistent with N3 and fMRI evidence for ventral object processing stream generators during categorization and priming (Henson et al., 2004; Schendan and Stern, 2008; Schendan and Maher, 2009; Schendan and Lucia, 2010). Specifically, during the N3 (**Figure 4B**), sources were observed, from 200 to 300 ms, in middle temporal gyrus (BA21 at 70 −35 −10) extending to inferior (BA37) and superior temporal (BA22) and fusiform (BA37) and middle occipital gyri (BA19) and, from 300 to 400 ms, at the junction

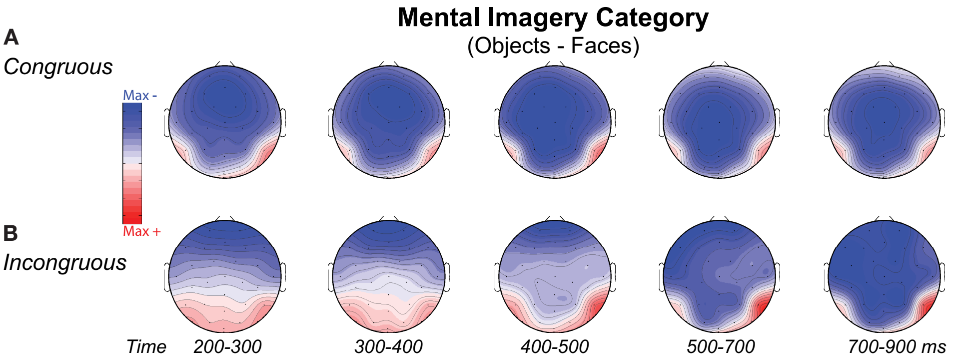


FIGURE 7 | Volt maps of mental imagery category effects. Maps plot voltage interpolated across electrode locations in each analysis time window for difference waves that define the category effect on **(A)** Congruous and **(B)** Incongruous trials, as highlighted in **Figure 6** for the N3, N400, and P3. Maps produced using *EEGLab* software.

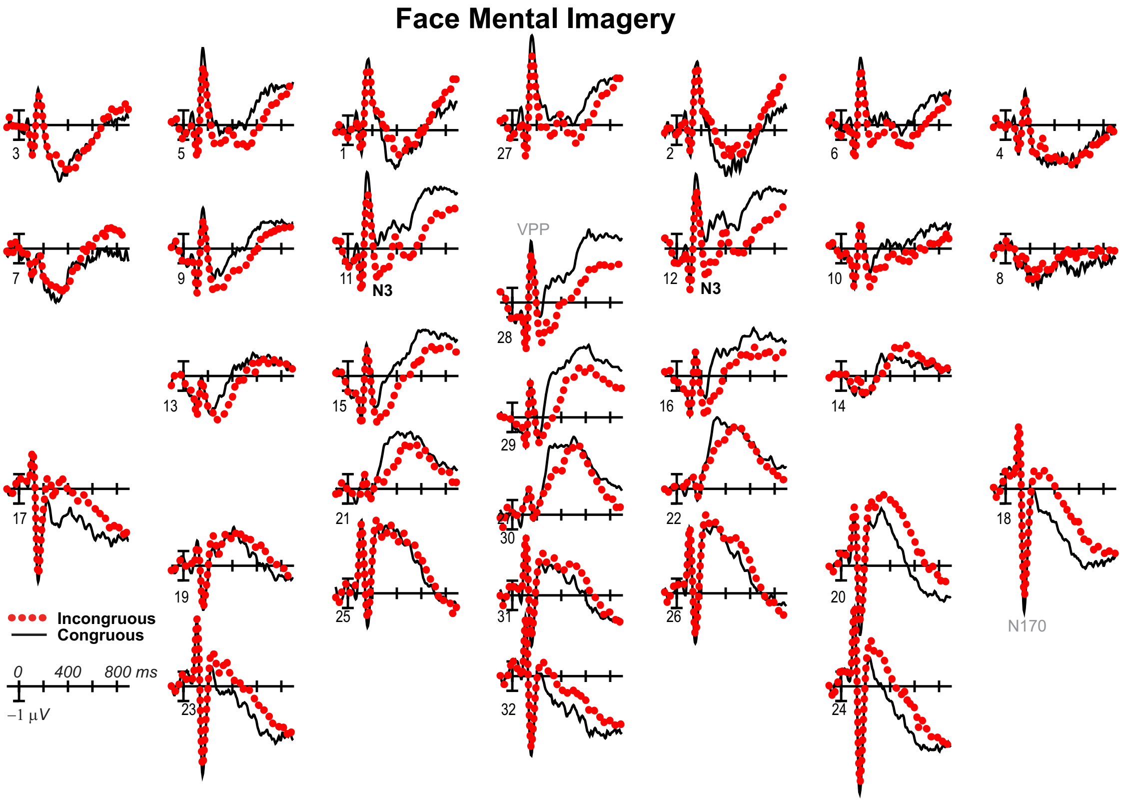


FIGURE 8 | Mental imagery congruity effects for faces with common average reference. Same as **Figure 2A**, except ERPs were re-referenced to the common average and plotted positive up for comparison with some other work.

of posterior fusiform and inferior occipital gyri (BA18 at 25 −90 −25 and 35 −90 −20) extending posteriorly to lingual (BA17) and anteriorly to fusiform (BA20 at −45 −25 −30; BA 37) and parahippocampal gyri (BA36). Afterward, during the later P3 peak to incongruous objects and the LPC (**Figure 5B**), various ventral stream sources continued, and prefrontal ones occurred initially from 400 to 500 ms: Sources were observed from (i) 400–500 ms, in VLPFC (BA47 at 20 30 −30), (ii) 400–700 ms, in fusiform (400–500 ms, BA37 at 55 −55 −25; 500–700 ms, BA20 at 55 −40 −30; BA36 at all times; BA19 at 400–500 ms), (iii) 400–900 ms, in inferior (BA20 and 37; a maximum also from 800–900 ms at BA20 at −60 −55 −20) and middle temporal gyri (BA20 until 700 ms), (iv) 600–700 ms, in parahippocampal gyrus (BA36).

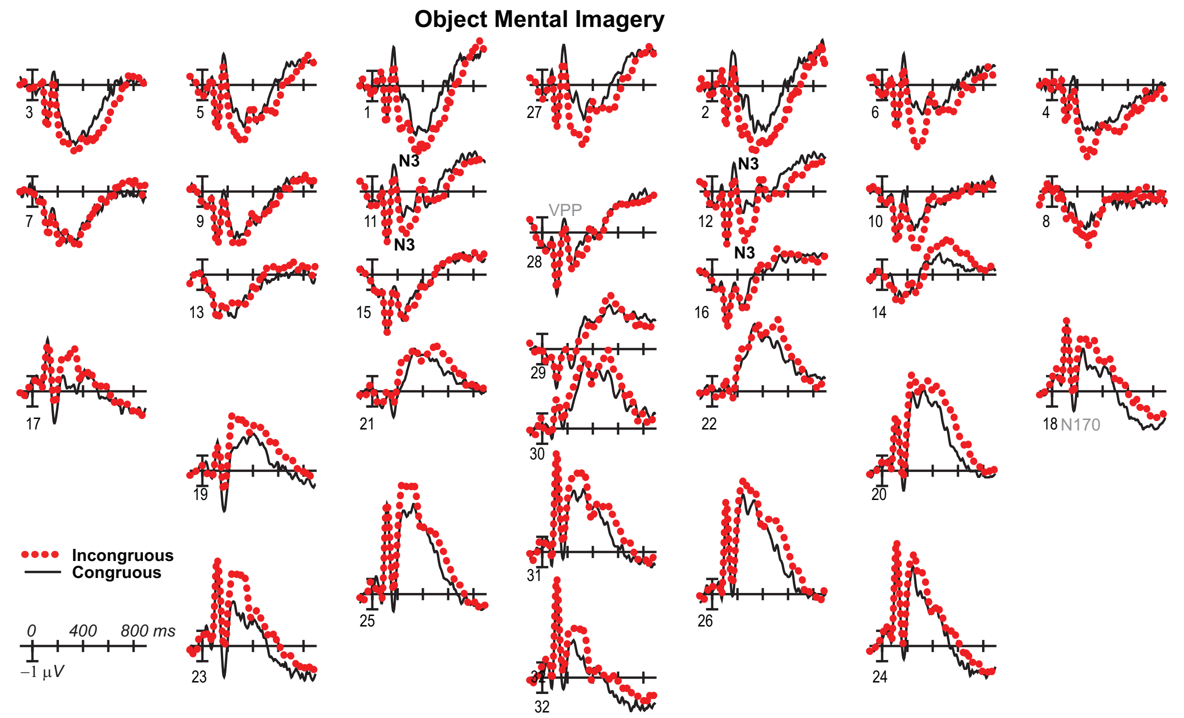


FIGURE 9 | Mental imagery congruity effects for objects with common average reference. Same as **Figure 3A**, except ERPs were re-referenced to the common average and plotted positive up for comparison with some other work.

Table 1 | ***F*-values for significant effects in lateral (lat) and midline (mid) omnibus ANOVAs with congruity (C) and category type (T) factors at each time period.**

ERP	N3		N3, N400, P3				LPC			
Time (ms)	**200–300**		**300–400**		**400–500**		**500–700**		**700–900**	
Source	**Lat**	**Mid**	**Lat**	**Mid**	**Lat**	**Mid**	**Lat**	**Mid**	**Lat**	**Mid**
Type	14.52**	5.78*	13.03**	8.84**	13.71**	13.26**	20.26**	27.61**	32.67**	32.8**
T × E	23.57**	20.59**	11.91**	11.6**	9.23**	7.78*	7.62**	6.19*	7.5**	8.24*
T × H	–	43.05**	–	20.17**	28.32**	13.15**	49.96**	5.49*	9.53**	8.04**
T × E × H	2.63*	–	–	–	–	6.04*	2.53*	12.56**	2.57*	10.1**
Congruity	12.99**	11.98**	28.72**	26.52**	–	–	–	–	–	5.94*
C × E	18.29**	24.73**	13.52**	14.7**	3.62*	–	9.38**	17.72**	–	–
C × H	–	20.38**	–	6.83**	–	–	–	8.84**	–	–
C × E × H	–	–	–	10.4**	–	–	–	–	–	–
T × C	–	11.62**	8.05*	16.65**	23.51**	23.72**	8.18*	14.38**	–	–
T × C × E	10.16**	–	9.66**	–	6.05**	–	6.89**	–	–	–
T × C × E × H	–	15.93**	–	19.21**	–	7.72**	–	10.93**	–	–

***$p < 0.01$, *$p \leq 0.05$, –$p > 0.05$. E, electrode; H, hemisphere for lateral or anterior-posterior site for midline.*

Category type (object–face)

As expected for domain-specificity (Downing et al., 2006), objects and faces differed primarily in object and face-sensitive areas of the ventral visual pathway (**Figure 6B** shows only 200–400 ms as later sources remained similar). Specifically, incongruous stimuli showed sources (i) continuously until 900 ms in inferior temporal (maximum 200–300 ms: BA20 at 60 −35 −20; maximum 400–500 ms: BA37 at 60 −55 −10) and (ii) middle temporal gyri (maximum 300–400 for BA20 at 60 −45 −20), (iii) in fusiform gyrus at most times (BA36, 37: 200–500 ms), and (iv) in middle occipital gyrus from 300 to 500 ms (BA37; BA19). Likewise, congruous stimuli also showed category differences in these regions:

Table 2 | F-values for significant effects of congruity (C) and category type (T) at lateral electrode pairs and time periods in focal ANOVAs (upper) and corresponding planned simple effects tests of congruity for each category type (lower).

Start time (ms)	200	300	400	500	700
FRONTOPOLAR (PAIR 1–2) FOCAL ANOVA					
T	45.13**	15.22**	13.29**	5.80*	5.76*
C	11.80**	16.96**	–	–	–
T × C	5.20*	–	–	–	–
Congruity effect for each category					
Objects	17.38**	12.28**	–	–	–
Faces	–	5.34*	–	–	–
FRONTOCENTRAL (PAIR 11–12) FOCAL ANOVA					
T	17.37**	13.84**	10.56**	18.08**	29.59**
C	23.87**	26.35**	–	9.80**	7.41*
T × C	5.31*	13.60**	21.49**	12.26**	–
Congruity effect for each category					
Objects	15.47**	7.94*	–	–	–
Faces	19.96**	31.84**	9.13**	15.25**	9.17**
OCCIPITOTEMPORAL (PAIR 17–18) FOCAL ANOVA					
T	9.10**	9.37**	10.79**	–	–
T × H	–	–	7.36*	10.13**	4.85*
C	–	–	4.97*	–	–
Congruity effect for each category					
Objects	5.22*	–	7.05*	7.76*	–
CENTROPARIETAL (PAIR 19–20) FOCAL ANOVA					
T	–	–	–	8.19*	7.33*
T × H	–	–	23.27**	33.56**	30.06**
C	–	10.42**	5.34*	6.32*	–
T × C	12.27**	5.91*	9.51**	–	–
T × C × H	–	4.63*	–	–	–
Congruity effect for each category					
Objects	9.36**	–	23.50**	13.82**	–
Faces	–	12.30**	–	–	–

***$p < 0.01$. *$p \leq 0.05$. –$p > 0.05$. H, hemisphere. Results for 100 ms time periods from 200 to 500 ms; results for 200 ms time periods after 500 ms.*

Sources were observed (i) in middle temporal gyrus (200–300 ms maximum for BA37 at 50 −40 −15; BA20), (ii) from 200 to 300 ms in fusiform gyrus (BA36/37), (iii) at most times in inferior temporal gyrus (BA37 at −60 −65 −10; BA20: 300–500 ms), (iv) from 300 to 400 ms in middle occipital gyrus (BA19/37). In addition, congruous stimuli showed sources of category differences in (v) STG from 400 to 500 ms (BA22 at 70 −25 5; BA41/42), consistent with superior temporal face-specific processes (Puce and Perrett, 2003).

PERCEPTION CONTROL ERPs

Results have already been reported for face ERPs before 500 ms and comparisons of the early VPP/N170 between perception and mental imagery of faces, demonstrating typical perceptual adaptation reduction of the VPP/N170 for repeated faces, as well as adaptation of later ERPs until 500 ms (Ganis and Schendan, 2008). Here, we focus on results for faces after 500 ms, results for objects, and comparisons between faces and objects. For faces (**Figure 10**), adaptation continued until 900 ms [i.e., more negative anteriorly and more positive at occipitotemporal sites for congruous (adapted) than incongruous], as shown by significant effects of congruity [700–899 ms $F(1, 17) = 8.47$, $p = 0.01$], congruity by electrode [500–699 ms $F(12, 204) = 8.91$, $p = 0.001$; 700–899 ms $F = 15.88$, $p < 0.001$], congruity by hemisphere [500–699 ms $F(1, 17) = 15.63$, $p = 0.001$; 700–899 ms $F = 11.72$, $p = 0.003$], and congruity by electrode by hemisphere [700–899 ms $F(12, 204) = 2.37$, $p = 0.048$]. Objects largely showed the opposite (**Figure 11**), resembling instead ERPs during rapid masked repetition priming that is associated with faster response times when the prime is more similar, relative to different from, the target (i.e., a positive priming benefit; Forster and Davis, 1984; Eddy et al., 2006). Such masked priming is associated with modulations of the VPP (labeled N190 in such work), frontal N3, and centroparietal N400, which are more negative for incongruous than congruous objects; an occipitotemporal N170 (labeled P190) and occipitotemporal N3 counterpart show the opposite (Eddy et al., 2006; Eddy and Holcomb, 2010). However, the present results would suggest that such positive priming benefits any category within the domain of non-face objects (i.e., between different non-face, basic, object categories) relative to the cross domain case (i.e., between faces and non-face objects), whereas all ERP masked object priming studies to date compared identical object pictures to unrelated non-face objects (Eddy et al., 2006, 2007; Eddy and Holcomb, 2009, 2010, 2011). To capture this similarity between the present object perception results and the prior masked object priming findings, the times and sites of masked priming effects were analyzed (Eddy et al., 2006; Eddy and Holcomb, 2010). Results confirmed the similarity of the present results to those in studies of masked priming: From 100 to 250 ms at frontocentral pair 11-12 for the VPP (i.e., N190), congruity was significant [$F(1, 17) = 29.34$, $p < 0.001$] and marginal at pair 23-24 for the occipitotemporal N170 (i.e., P190; $F = 3.44$, $p = 0.081$), significant from 250 to 350 ms for the N3 at frontocentral pair 11-12 ($F = 21.84$, $p < 0.001$), and 350–500 ms for the N400 at centroparietal pair 19-20 ($F = 23.49$, $p < 0.001$), but the centroparietal effect appeared to reflect only the overlapping LPC, suggesting minimal N400 modulation if any.

For completeness, the same focal spatiotemporal analyses compared perceptual congruity between types, as done for imagery, and a 140–180 ms time window was added to assess the VPP/N170, as had been done previously for faces (Ganis and Schendan, 2008). Results confirmed that perceptual congruity effects differed between categories during the VPP/N170, frontal N3, centroparietal N400, and posterior LPC. Specifically, at all times after 200 ms [all *df*s (1, 17)], frontopolar ERPs showed effects of type (type at 500–899 ms $F\text{s} > 5.72$, $p\text{s} < 0.03$; type by hemisphere at 200–899 ms $F\text{s} > 11$, $p\text{s} < 0.005$) and congruity (200–899 ms: congruity $F\text{s} > 43$, $p\text{s} < 0.001$; congruity by hemisphere $F\text{s} > 10$, $p\text{s} < 0.006$), and their interaction (type by congruity by hemisphere at 200–899 ms $F\text{s} > 4.99$, $p\text{s} < 0.04$). Frontocentral ERPs showed effects of type during the VPP and after 500 ms (140–180 and 500–899 ms $F\text{s} > 7.96$, $p\text{s} < 0.02$), effects of congruity at all times (congruity at 140–180 and 200–899 ms $F\text{s} > 6.2$, $p\text{s} < 0.03$; congruity by hemisphere at 200–299 ms $F\text{s} > 4.91$, $p\text{s} < 0.042$), and their interaction during the VPP and N3 (type by congruity by

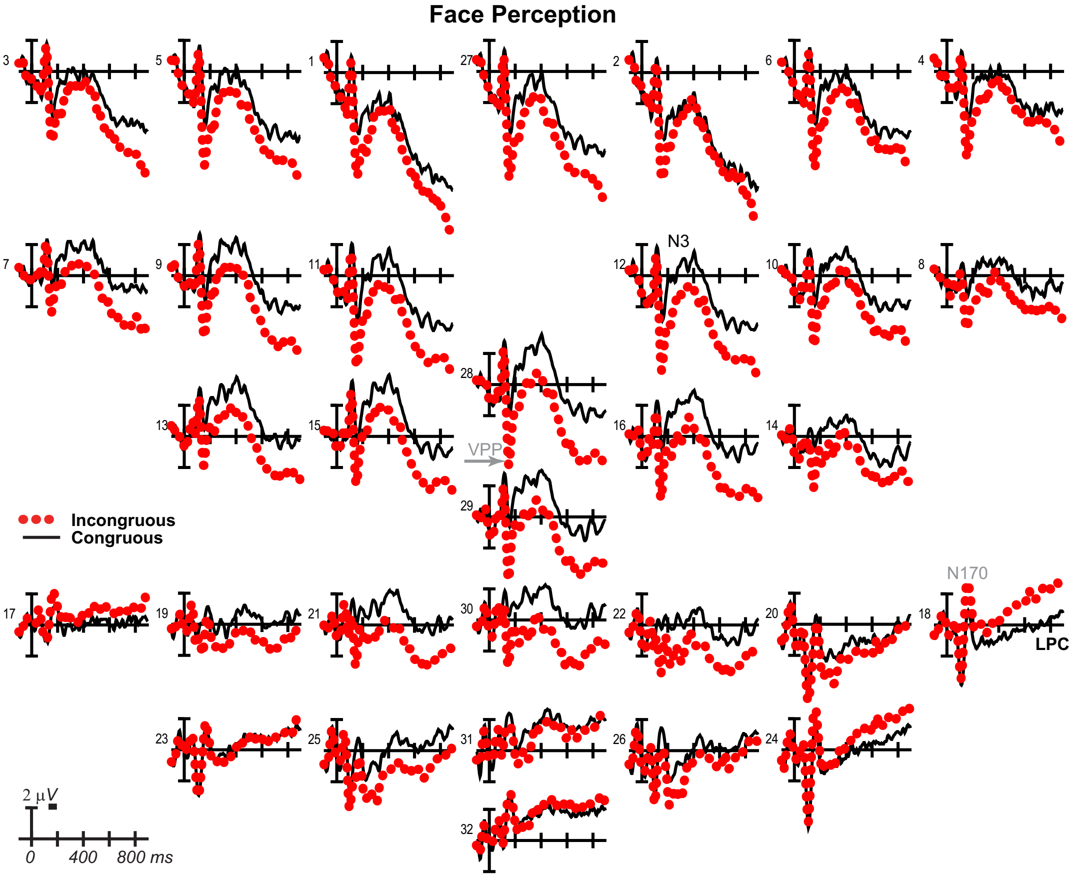


FIGURE 10 | Perception congruity effects for faces. Grand average ERPs between −100 and 900 ms at all 32 channels shown filtered low-pass 30 Hz. LPC, late positive complex. For congruous, pictures of two different people's faces were presented sequentially, and for incongruous, a picture of a non-face object preceded a picture of a face.

hemisphere at 140–180 and 200–299 ms $Fs > 6.31$, $ps < 0.01$). Occipitotemporal ERPs showed effects of type on the N170 and LPC (type at 140–180 and 400–499 ms $Fs > 4.65$, $ps < 0.05$; type by hemisphere 200–299 ms $Fs > 12$, $ps < 0.003$), and only the N170 showed effects of congruity (140–180 ms, $F = 15.81$, $p = 0.001$) and the interaction (type by congruity at 140–180 ms, $F = 16.80$, $p = 0.001$). Centroparietal sites showed effects of type (type by hemisphere at 200–499 ms, $Fs > 6.02$, $ps < 0.03$), congruity (400–499 ms $Fs > 6.77$, $ps < 0.002$), and interactions of type by congruity (200–499 ms $Fs > 7.21$, $ps < 0.02$), but reflected N400 adaptation for faces and instead overlapping LPC priming for objects. After 300 ms, the LPC at posterior pair 25-26 showed effects of type (300–499 ms $Fs > 8.31$, $ps < 0.02$) and the interactions (type by congruity at 500–899 ms, $Fs > 4.94$, $ps < 0.05$; type by congruity by hemisphere at 300–399, $Fs > 6.56$, $ps < 0.03$).

PERFORMANCE

As stimulus timing is critical to interpreting the results, time to report mental image generation (mental imagery experiment) or identification (perception experiment) is reported in more detail than before (Ganis and Schendan, 2008). Results of an ANOVA with experiment (imagery, perception), congruity, and category factors demonstrated that timing differences cannot explain category and congruity effects. Mental imagery was slower than perception [$F(1, 17) = 267.30$, $p < 0.001$], as subjects took a long time, 3,889 ms, from the onset of the word cue to report generation of the mental image, and only 1,070 ms from the onset of the first picture during perception to report identification of the face (categorization of the object); note, with the 200 ms ISI included, these times corresponded to an average, stimulus onset asynchrony (SOA) of 4,089 ms for imagery ($SD = 792$ ms; range 2,548–5,252 ms) and 1,270 ms for perception ($SD = 346$ ms; range 754–2,113 ms). Critically, these SOA times were indistinguishable between category and congruity conditions ($Fs < 1$, $ps > 0.8$), and, no interaction was significant ($Fs < 2.1$, $ps > 0.17$). Specifically, for mental imagery, SOAs were similar between categories and between congruous (faces 4,071 ms; objects 4,107 ms) and incongruous conditions (faces 4,081 ms; objects 4,096 ms). Likewise, for perception, SOAs were similar between categories and between congruous (faces 1,311 ms; objects 1,224 ms) and incongruous conditions (faces 1,247 ms; objects 1,298 ms).

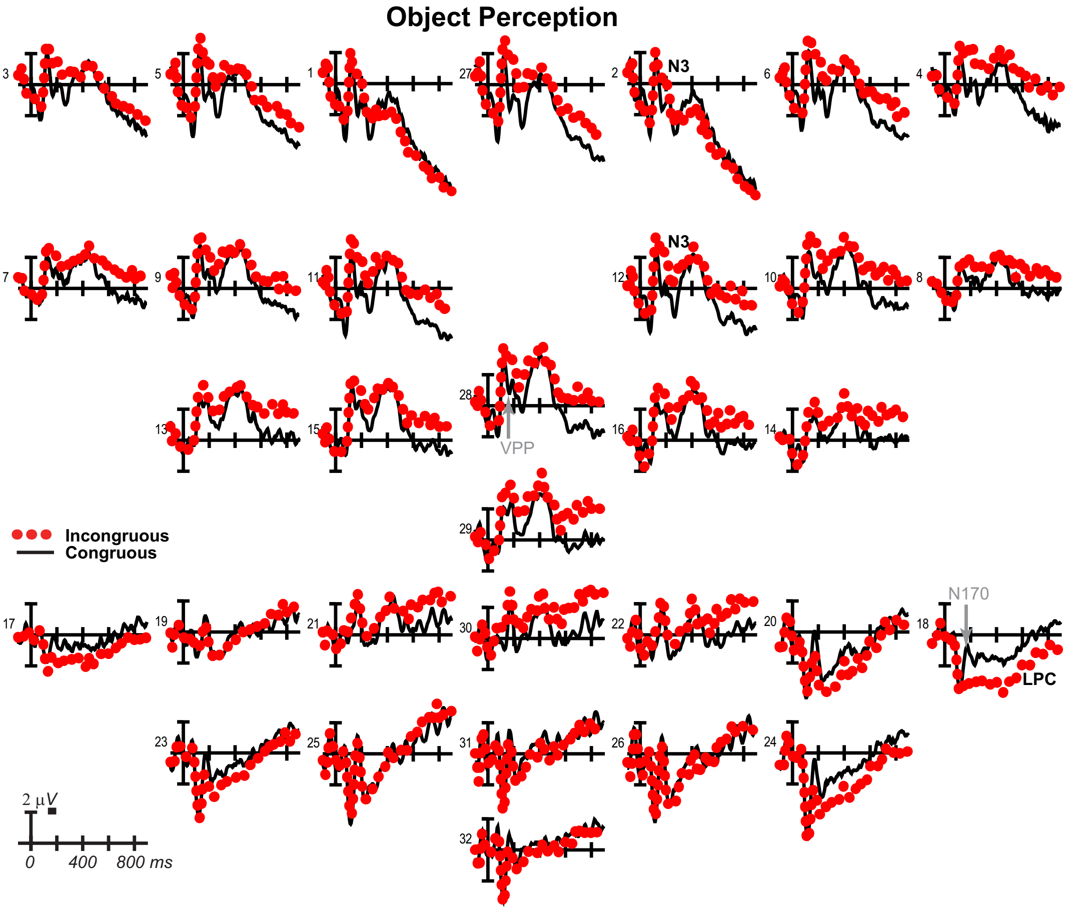


FIGURE 11 | Perception congruity effects for objects. Grand average ERPs between −100 and 900 ms at all 32 channels shown filtered low-pass 30 Hz. LPC, late positive complex. For congruous, pictures of two different objects were presented sequentially, and for incongruous, a picture of a face preceded a picture of an object.

DISCUSSION

The findings reveal the cortical dynamics of ongoing top-down processes of mental imagery of visual shape during later knowledge, meaning, and decision processing of a perceived picture. Overall, ongoing mental imagery facilitates categorical perception during the early VPP/N170, as reported previously (Ganis and Schendan, 2008), and higher cognitive processes during later ERPs, as reported here. Specifically, the frontal N3(00) complex, which indexes object and face cognition, knowledge, and category decisions from 200 to 400 ms and the centroparietal linguistic N400 index of semantic memory from 300 to 500 ms are more negative for incongruous than congruous stimuli. Notably, N3 and N400 effects dissociate from each other not only temporally (i.e., earlier for N3) and spatially (i.e., more frontal for N3) but also in how congruity effects differ between categories. N3 effects are frontopolar for objects and frontocentral for faces and associated with different brain sources between categories. While the N400 congruity effect is small but clear for faces, it is smaller, and, indeed, minimal and not clearly evident for objects. In addition, the parietal P3 peaks earlier for congruous (∼400 ms) than incongruous (∼500 ms) stimuli. After 400 ms, the posterior LPC shows the opposite congruity effect from the N3 and N400, being instead more positive for incongruous than congruous stimuli until 700 ms. This pattern of effects resembles a combination of ERP effects of semantic congruity (on N3, N400, and LPC) and immediate repetition priming (on N3, N400, and P3). In contrast, perception shows a different pattern of congruity effects, as predicted due to the bottom-up contributions to perception but not imagery. Further, the pattern differs between categories: Faces show adaptation; objects show rapid priming effects in the opposite direction. Altogether, these findings indicate that top-down processes of mental imagery can induce a powerful imagistic mental representation of visual shape that mimics top-down processes recruited also for picture perception and facilitates knowledge, meaning, and categorization processes.

MENTAL IMAGERY

Next we consider the evidence that these mental imagery effects are related to semantic congruity and immediate repetition priming processes and the implications of this for how mental simulation can ground cognition in shape processing. First, it is important to be reminded of key, unique aspects of the present methods (Ganis and Schendan, 2008). (a) People had extensive training generating mental images of each picture of a real person or object, a standard procedure used in validated mental imagery tasks. (b) The name (i.e., written word) for the person or object cued subjects to visualize mentally each associated trained picture. (c) A picture probed ongoing mental imagery of faces and objects, and ERPs were recorded to this picture, which either was the picture subjects were cued to visualize mentally or another picture from the opposite category. Further, two categories (faces and objects) that are supported by different visual processing areas (Hasson et al., 2003; Downing et al., 2006; Tsao et al., 2006) were used to manipulate congruity, and the opposite category was assigned to the incongruous condition. Consequently, congruity effects reflect large differences in shape processing. This is because congruent face imagery (identical face) maximally affects face processes, while incongruent imagery (object) minimally affects face processes, and analogously for object imagery. (d) The delay between the name cue and target picture was relatively long (SOA 4,089 ms, on average). (e) No task was performed on the target picture to minimize decision and response related ERPs that can complicate interpretation of the waveform, thereby defining knowledge and semantic memory processes as clearly as possible. While a limitation of this study is that the target picture was not followed by a task that assessed the mental imagery, evidence that subjects did as instructed is that they took much longer (2,819 ms) to visualize the named picture in the mental imagery experiment than to identify the perceived face (object) picture in the perception control experiment.

Mental imagery produces semantic congruity effects and facilitates cross modal priming

The most important finding here is that mental imagery produces ERP effects that resemble N3, N400, and LPC effects observed in studies of short latency, cross modal priming, semantic congruity, and semantic priming phenomena. No prior behavioral or ERP study on these phenomena or mental imagery would have predicted this finding, as mentioned in the introduction. One reason is that the timing for mental imagery here is well beyond that for automatic spreading activation associated with semantic priming, which is thought to underlie semantic congruity effects and to reveal semantic memory processes (Kutas and Federmeier, 2011). Consider that, in the typical semantic priming task, two related words are presented sequentially with a brief delay (usually under 1 s): A target word (e.g., “doctor”) is preceded by a prime word that is different and either semantically related (“nurse”) or unrelated (“truck”) to the target. Response times are faster to targets preceded by primes that are semantically related (congruous) than unrelated (incongruous). Findings with SOAs between prime and target of under 500 ms reflect automatic spreading activation in the semantic network, whereas SOAs between 500 and 1000 ms or so reflect controlled semantic analysis (Rossell et al., 2003; O’Hare et al., 2008). Thus for the timing in our study (for both imagery and perception), the results can reflect only controlled semantic processing. In contrast, most semantic priming work studied automatic spreading activation using short SOAs under 500 ms and so evidence of behavioral and ERP priming with longer delays is scarce and more so for the much longer delays here.

A second reason these findings are novel is that the procedure here of having a word (the name) precede a picture is used in cross modal priming studies. Critically, for word-picture priming, behavioral evidence has been mixed and, if such cross modal priming is found, it occurs mainly with much shorter delays than used here. Such studies typically use an SOA of about 1 s (Bajo, 1988) and often much less (Carr et al., 1982) but also often mask the prime. At such short SOAs, word-picture priming can be comparable (Bajo, 1988) or much less than within modality (e.g., picture to picture; Carr et al., 1982). At slightly longer SOAs of less than about 2 s, priming can be absent (Biggs and Marmurek, 1990). Like behavioral effects, ERP effects of cross modal priming (on the N3, N400, and P3/LPC) have been found most consistently in studies using SOAs briefer than 500 ms or unmasked primes composed of multiple words in sentences or noun phrases (Ganis et al., 1996; Federmeier and Kutas, 2001, 2002; Stanfield and Zwaan, 2001; Zwaan et al., 2002; Hirschfeld et al., 2012). Cross modal priming from a single word to a picture is more variable but has been found at SOAs ranging from 120 to 700 ms on the N400 (and perhaps P3/LPC) between 350 and 550 ms when the prime is unmasked (Auchterlonie et al., 2002; Johnson and Olshausen, 2003, 2005; Dobel et al., 2010; Kiefer et al., 2011) and masked (Blackford et al., 2012). In contrast, N3 cross modal priming has been found only in studies that (a) mask the word prime, use the shortest SOAs (120 ms or less), and overt naming, or (b) use long SOAs of about 1–2 s and name verification (Johnson and Olshausen, 2003, 2005). Notably, for priming from a word to a picture, visually impoverishing the objects (by occlusion or fragmentation) yields a more frontopolar distribution of congruity effects. This scalp distribution is consistent with the frontopolar N3 in studies of object cognition and priming with non-canonical views, fragmented real objects, and pseudo objects (Holcomb and McPherson, 1994; Schendan et al., 1998; McPherson and Holcomb, 1999; Schendan and Kutas, 2002, 2003, 2007; Schendan and Lucia, 2009; Schendan and Maher, 2009), which recruit top-down processes more (Michelon et al., 2003; Ganis et al., 2007), and as found here for mental imagery of objects. In sum, behavioral and ERP (N3, N400, P3/LPC) effects of cross modal priming can occur at shorter SOAs, with multiple words as the prime, and during naming tasks, but, crucially, none of these procedures apply to the mental imagery task used here.

In contrast, for long SOAs well beyond about 1 s and more like mental imagery here, behavioral cross modal priming from a single word to a picture has not been found (Morton, 1979; Warren and Morton, 1982) or is much smaller than that within modality (Carr et al., 1982). The ineffectiveness of word primes for picture targets at long delays, however, can be overcome, by (a) varying prime modality only between- (i.e., not within-) subjects, (b) blocking prime modality (Brown et al., 1991), or (c), critically here, instructing subjects to use mental imagery. A word that is used to cue mental imagery during a study session primes later performance with the picture at a long delay (minutes) on an implicit memory test with objects (McDermott and Roediger, 1994) or famous faces (Cabeza et al., 1997) and can do so as well or better than a perceived picture (Michelon and Koenig, 2002; Michelon and Zacks, 2003). However, it is important to note that these priming studies do not use the picture target at test to reveal ongoing mental imagery sustained within working memory, as herein, but rather its long-term consequences for a memory test much later (i.e., beyond the time limits for working memory). Notably, prior ERP studies, in which a single word could prime a picture target at a long delay, have not used blocking or mental imagery procedures, and, accordingly, no behavioral or ERP priming effects were found. Although episodic recognition does show effects (Kazmerski and Friedman, 1997; Spironelli et al., 2011), these do not apply here because we assess ongoing mental imagery, not the consequences for later episodic recollection. Further, even if N400 effects are found at such long lags, they likely reflect morphological (linguistic) representations, not semantic or phonological representations which do not seem to survive lags beyond SOAs of 300 ms (Feldman, 2000; Koester and Schiller, 2008). This would suggest that N400 effects for mental imagery here reflect linguistic, not semantic, memory representations (Kousta et al., 2011), but future work needs to assess this.

Mental imagery primes perception like immediate picture repetition does

Altogether, these direct neurophysiological findings are consistent with behavioral evidence that mental imagery facilitates object categorization via priming mechanisms (Peterson and Graham, 1974). Most striking is the finding that the N3, N400, and P3 congruity effects mimic ERP immediate repetition priming, providing direct neurophysiological evidence that mental imagery can affect neural processing like actual perception of a picture can. When the exact same image repeats immediately with no intervening stimuli, ERPs after 300 ms become earlier and larger for objects (Nielsen-Bohlman and Knight, 1994; Zhang et al., 1995), faces, and words (Bentin and McCarthy, 1994; Schendan et al., 1997). The P3 is more positive and peaks earlier (∼400 ms) and the following N400 is more positive for repeated (akin to congruent) than unrepeated (akin to incongruent) faces and objects, and a later P3 or LPC, peaking around 500 ms, is instead more positive for unrepeated than repeated pictures. These ERP effects of perceptual immediate repetition priming have been observed at relatively short SOAs of 1200–3500 ms, which is much longer than the 500 ms SOA necessary to observe automatic spreading activation in semantic priming and slightly longer than the about 1 s SOA for cross modal priming (without special conditions like mental imagery). For objects, most studies could not or did not assess frontal ERPs, but one study also shows the frontal N3 is more positive for repeated than unrepeated items (SOA 2400 ms; Henson et al., 2004). This study also reported VPP/N170 repetition adaptation for objects (i.e., smaller for repeated), but this adaptation direction is the opposite of the later repetition priming effects in that study and of the mental imagery effects here, and no other immediate repetition study found effects before 200 ms (Nielsen-Bohlman and Knight, 1994; Zhang et al., 1995). Altogether, these findings indicate that mental imagery mimics the pattern of immediate repetition priming of perceived pictures on the N3, N400, P3, and LPC. However, mental imagery also enhances the VPP/N170, unlike immediate repetition priming, which typically has little or no effect on early ERPs.

The similarity between mental imagery effects and ERP immediate repetition priming is consistent with the role of working memory in both. Immediate repetition priming is due to working memory for the first image that is sustained across the brief delay until the second image appears (Bentin and McCarthy, 1994). This working memory facilitates categorization of the repeated percept with minimal or no reactivation of semantic memory, minimizing the N400. Likewise, the priming effect on the frontal N3 could indicate that visual knowledge and cognitive decision processes for objects and faces are also largely bypassed. Immediate repetition effects on the P3 reflect modality-specific (i.e., visual) working memory that speeds the category decision (Bentin and McCarthy, 1994; Nielsen-Bohlman and Knight, 1994; Zhang et al., 1995). Consistent with maintaining modal visual information in working memory, P3 facilitation is not associated with semantic priming, as reviewed above. Further, immediate repetition effects on the P3 are likely also due to having subjects perform a task on the pictures, which was often episodic recognition, as such task requirements maximize P3 and other late posterior positivities (Dien et al., 2004). Mental imagery had no task requirements for the target picture, but the practice session required subjects to assess how well their mental image matched the picture; thus subjects likely continued to do so incidentally during the mental imagery test, resulting in P3 facilitation despite no overt task. Thus, mental imagery can simulate the top-down cortical dynamics that are produced by an actual perceived picture, and the episodic memories encoded during training and practice contain visual details sufficient to enable mental imagery representations to operate like an actual perceived picture (as in immediate repetition priming). This provides strong and direct neurophysiological support for the pictorial theory of mental imagery and implicates these strong pictorial representations in episodic memory of personally experienced, autobiographical information that depends upon the mediotemporal lobe (Ganis and Schendan, 2011). This finding also constitutes evidence for the visual detail achievable by the episodic memory system. Such evidence will be crucial for developing theories of mental simulation for episodic memory (Schacter et al., 2008).

Reflexive top-down processes for mental imagery support automatic mental simulation

By using faces and objects, which have partially non-overlapping visual processing pathways, these mental imagery findings define the largest possible set of top-down mechanisms that support mental simulation of face (object) shape, including non-conscious automatic simulation. Mental simulation has been proposed to operate via a pattern completion process that re-enacts modal processing that had occurred during learning when later retrieving the memory (Barsalou, 2009). We proposed that, at the level of brain mechanisms, the top-down feedback mechanisms that support automatic simulation are a subset of those that support mental imagery (Ganis and Schendan, 2011). Specifically, automatic simulations unfold via reflexive top-down signals from higher to lower level areas along modal information processing pathways, such as the ventral stream: Perceiving a stimulus triggers these processes reflexively (Ganis and Kosslyn, 2007). Through such distributed multi-regional activity, seeing an object or reading its name (e.g., "dog"), for example, re-enacts associated modal features that were stored during earlier learning experiences (e.g., its shape, color, motion, actions with it), thereby constructing cognition, memory, and meaning. This is consistent with the MUSI account that proposes top-down feedback processes after 200 ms have the greatest role in visual cognition, constructing knowledge, meaning, memory, and decisions (Schendan and Kutas, 2007; Schendan and Maher, 2009; Schendan and Lucia, 2010). These same processes are triggered by strategic top-down signals from the prefrontal cortex during mental imagery (Ganis and Kosslyn, 2007) and so mental imagery time courses like those here can define when and how mental simulation grounds cognition. Previously, semantic priming has revealed the most about automatic mental simulation and its brain basis, especially with words (e.g., Marinkovic et al., 2003; Rossell et al., 2003; Kutas and Federmeier, 2011). This is because automatic spreading activation across a semantic memory network, which explains such priming, is thought to operate via the same automatic and reflexive, top-down processes that have been implicated in automatic simulation (e.g., Collins and Loftus, 1975; Franklin et al., 2007; Kutas and Federmeier, 2011). The similarity between the results here using a validated mental imagery task and ERP findings related to semantic congruity and immediate repetition priming supports this conclusion.

PERCEIVED PICTURE IDENTIFICATION ADAPTS FACES BUT PRIMES OBJECTS

Overall, perception control results confirm that common top-down processes underlie similarities between imagery and perception, while bottom-up processes for perception (but not imagery) underlie their differences (Ganis and Schendan, 2008). The time precision of ERPs offers advantages over fMRI and behavior for characterizing such similarities and differences between perception and imagery. Specifically, perception results dissociate between categories, consistent with the domain-specificity of object and face processing (Downing et al., 2006): Perceptual repetition adapts processing of perceived faces from categorical perception onward, as predicted, but instead unexpectedly primes processing of objects during categorical perception, visual knowledge processing, and strategic semantic analysis. Critically, identification time for the first picture is similar for objects and faces and so cannot explain differences in congruity effects. In fact, the timing was chosen to replicate classic face adaptation effects on the VPP/N170(Jeffreys, 1996) obtained with an 1,100 ms SOA, 800 ms duration, and 300 ms ISI (i.e., like the 1,070 ms identification RT and 200 ms ISI here), as was achieved (Ganis and Schendan, 2008). To understand the perception control findings, it is necessary to highlight that both prime and target were always different pictures, even in the congruous condition. Thus, perception results show how perceiving a picture of a face (object) is affected by having previously identified a perceived picture of a different face (object), compared to having previously identified the opposite category picture [i.e., of an object (face)].

For faces, seeing two different people sequentially adapts the ERPs, thereby producing the opposite congruity effect from that for mental imagery. For congruous relative to incongruous perception, the VPP is less positive (and N170 less negative), as reported previously (Ganis and Schendan, 2008), the N3 and N400 are more negative, and the LPC is more positive. Altogether this finding and the mental imagery finding indicate that early congruity effects on the VPP/N170 extend to later ERPs. The direction of these effects suggests their interpretation. Consider that all these adaptation effects go in the opposite direction compared to the facilitation effects found for mental imagery: Adaptation reduces the VPP/N170 but enhances later ERPs. Such enhancements of later ERPs are thought to reflect greater (not less) recruitment of the underlying processes (Schendan and Kutas, 2003). Hence, early adaptation of categorical perception during the VPP/N170 impedes later cognitive processing (i.e., due to failure of an earlier critical process), thereby recruiting additional top-down processing resources to accomplish these later cognitive functions (Kosslyn et al., 1994; Ganis et al., 2007).

The direction of the object perception findings tells a different story, which further supports a facilitation interpretation of the mental imagery findings because the direction of the effect is the same as for imagery. Specifically, under rapid, immediate serial presentation, the perception of an object picture primes (facilitates) a subset of the ERPs to a target object picture that mental imagery also primes. For congruous relative to incongruous conditions, the VPP is more positive (and N170 more negative), the N3 less negative, and LPC less positive, whereas the N400 shows minimal or no priming. Unlike imagery, though, perceptual repetition shows no P3 modulation, but this is presumably due to no overt or implied task on the perceived target picture. Surprisingly, therefore, priming of whatever processes are shared among a set of real objects from different basic categories can facilitate processing of each other (i.e., congruous perception), in contrast to the cross domain case of perceiving a face and then a non-face object (i.e., incongruous perception). The resemblance between these perception (and the mental imagery) results and those for certain kinds of rapid perceptual and semantic priming supports a facilitation interpretation. After all, the N3, N400, and LPC effects of object perception congruity resemble a subset of effects for mental imagery and immediate repetition and semantic priming. Moreover, the waveforms resemble those associated with priming under the most rapid, serial presentation of pictures (SOA < 500 ms) when the prime is either not masked and semantically related pictures repeat (Holcomb and McPherson, 1994; McPherson and Holcomb, 1999; Kiefer et al., 2011) or masked and identical pictures repeat (Eddy and Holcomb, 2010). In masked repetition priming, the VPP (labeled N190 in these studies), N3, and N400 are more negative [and occipitotemporal N170 (labeled P190) and N3 counterpart are more positive] for unrelated than repeated (identical) pictures of objects (Eddy et al., 2006; Eddy and Holcomb, 2010). The waveform similarity between object perception here and masked priming must be due to the very short 200 ms ISI used here for perception, causing the ERPs to the prime and target to overlap temporally, as they do in the masked priming work, which uses short ISIs of 100 ms or less. The short ISI is probably also responsible for some of the effects resembling effects for rapid repetition and masked semantic priming (with ISIs of 200 ms or less) more than for longer lag, immediate repetition priming (with SOAs of 1200–3500 ms). In turn, the longer SOA here may explain why the later perceptual congruity effects also resemble some longer lag, immediate repetition priming effects on the N3 and LPC at SOAs of about 1 to 1.5 s, which show the same pattern of congruity effects on the

N3, N400, and a posterior P3/LPC (Barrett and Rugg, 1990; Holcomb and McPherson, 1994; McPherson and Holcomb, 1999; Bach et al., 2009). Thus, for objects (but not faces) perception-driven priming and mental imagery congruity effects differ quantitatively and qualitatively but nonetheless all follow a direction that indicates priming facilitation. The similarity between the perception and imagery ERP congruity effects with objects further bolsters the idea that mental imagery mimics perception: Mental imagery of objects, primes (facilitates) object processing like repeating a perceived picture.

CAVEATS

Future work will need to investigate why, despite identical timing, perception of faces and objects produce opposite congruity (repetition) effects on the ERPs to the probe picture. N170 rapid adaptation evidence (Nemrodov and Itier, 2012) suggests that faces are stronger adaptors than some non-face categories. Face (and car) primes reduce (i.e., adapt) the N170 more than chair and house primes for all categories of test objects (i.e., faces, cars, chairs, houses). Prime category and ISIs are the key factors determining how much, if at all, the prime adapts the test stimulus; after all, the ISIs of 232–268 ms and test stimulus duration of 200 ms resemble the present timing, but the faces were of unknown people, whereas here they were famous, and the prime was much briefer (i.e., 200 ms) than here, suggesting neither knowledge nor prime duration can explain the findings. The reason that the category of the prime matters is unclear but generally consistent with transfer appropriate processing and encoding specificity accounts of memory (Tulving and Thomson, 1973; Morris et al., 1977). Other proposals include interference from ongoing late processing of the adaptor due to the short ISI and neural fatigue that is not category specific but selectively tuned to adaptor properties (Nemrodov and Itier, 2012). Regardless, faces can adapt more than other object categories. The present results suggest an additional twist: Non-face objects can prime better than faces. One may speculate that, here (and in related work), face perception shows a substantial adaptation pattern because faces (congruent) adapt better than other objects (incongruent), while objects prime better, and object perception shows a substantial priming pattern because objects (congruent) prime better than faces, while faces (incongruent) adapt better; both adaptation and priming influences could affect any result, making congruity effects larger than either influence alone. The relative strengths of adaptation and priming explain the pattern observed, which will be important for future research to tease apart.

This caveat also speaks to an alternative hypothesis to explain why both perception and mental imagery prime, but perception also adapts a stimulus: Perception adapts not only face (object) processing but also adapts visual sensory processing, and the latter results in reduced sensitivity in lower-level areas, including occipital cortex (Wilson and Humanski, 1993; Anderson and Wilson, 2005; Loffler et al., 2005). This is another way of stating that both perception and mental imagery can recruit top-down processes, while perception is also driven by bottom-up processes (Ganis and Schendan, 2008) but adds the idea that perception also adapts lower-level occipital areas. After all, both perceptual priming and adaptation effects on perception are well-documented (Grill-Spector et al., 2006), and top-down processes from the frontal lobe, which support mental imagery, have been implicated in perceptual repetition priming (Schacter et al., 2007). Consider that different faces share more lower-level features, such as spatial frequency spectra, and so have less interstimulus variance than different objects (Costen et al., 1996; Thierry et al., 2007; Ganis et al., 2012). Hence, different faces can mutually adapt not only face processing but also lower-level feature processing more than different objects. For example, the N170 face-specificity effect overlaps spatiotemporally with an N1(00) component (i.e., N170 face-specificity modulates the N1), which reflects widespread bottom-up and reflexive feedback processing along the visual pathways from lower to higher level areas (e.g., Vogel and Luck, 2000; Bullier, 2001). Adaptation of both the face-specific and the low-level visual processes results in a large decrease, as both N170 face and N1 visual components are affected. In contrast, mental imagery of face-specific processes increases N170 amplitude, but this is a smaller change (than the decrease for adaptation) because only the N170 face component is affected. This explanation needs to be considered to resolve whether perception and mental imagery share properties (Freyd and Finke, 1984) or not (Craver-Lemley and Reeves, 1992).

Finally, given the novelty of this mental imagery probe paradigm (Ganis and Schendan, 2008), several issues remain to be resolved. For example, future work will need to evaluate how individual differences in mental imagery and other abilities (Kozhevnikov et al., 2005) affect mental imagery processes, as well as how congruity effects differ as a function of the vividness of the perceived stimulus and subjective vividness of mental imagery (Herholz et al., 2012). Also, mental imagery will need to be compared directly with cross modal priming from a word to a picture and potential processing differences (e.g., depth of semantic network activation) will need to be addressed.

CONCLUSION

The ERP results described in this report define the neurophysiological characteristics and time course of top-down processes for mental imagery of the visual shape of faces and objects that can ground cognition in these modal processes. These findings provide striking direct neural evidence that top-down feedback processes of mental imagery sustain an imagistic representation that mimics perception well enough to prime subsequent perception and cognition like an actual picture. By manipulating congruity by switching between face and object categories, which involve different modal processes along the ventral visual stream, the ERPs reveal the largest set of top-down processes for mental imagery of these shape categories. The subset of these mental imagery processes that correspond to the reflexive top-down inputs from higher to lower level areas along the posterior ventral face (object) processing pathway also constitute the automatic mental simulation processes that can ground cognition of faces (objects). The ERP congruity effects here therefore provide direct neurophysiological markers for these visual shape mental simulations that can be used to determine precisely when, how, and how much these cortical mental simulation mechanisms ground cognition. Together, the robust frontal N3 and minimal centroparietal N400 mental imagery congruity findings confirm the visual imagistic (non-linguistic) nature of shape mental imagery, and further implicate the N3 as an index of visual knowledge (Schendan and Maher, 2009) and the N400 as an index of linguistic knowledge. This is consistent with grounded cognition distinctions between non-linguistic ("experiential" modal sensorimotor and mental state information) and linguistic systems (word-related associations) for semantic memory (Kousta et al., 2011; Paivio and Sadoski, 2011). Thus, future work on the cortical dynamics of the contribution of mental simulation of visual shape to semantic memory should focus on the frontal N3. Further, finding both mid-latency (N3, N400) and LPC congruity effects suggests that, to ground cognition in modal processing, two types of mental simulation can operate at two distinct times. Automatic simulations of visual shape (due to reflexive top-down processes) and linguistic processes operate between 200 and 500 ms during the N3 and N400, respectively, and effortful simulation (due to strategic top-down processes) operates between 400 and 700 ms during the LPC. Altogether, these markers, and others defined using the methods developed here, can be used to characterize and probe these mental simulation processes in future research on grounded cognition theory, especially for discovering the neural mechanisms of how mental simulation works.

In addition, for both imagery and perception, functional dissociations and spatial distribution differences between faces and objects provide further evidence for domain-specificity and the modality-specific processing posited for grounded cognition. In contrast to mental imagery, perception of a face (instead of an object) adapts categorical perception, which consequently impairs later processing of a different face, whereas perception of an object (instead of a face) primes categorical perception, activation of visual knowledge, and later categorization of a different object. The differences between mental imagery and perception are consistent with the strategic top-down processes required to construct and maintain mental imagery versus the bottom-up input required for perception, whereas the similarities are consistent with the common, automatic, top-down, modal processing supporting both, and further support the imagistic nature of mental imagery.

AUTHOR NOTE

Preparation of this manuscript was supported by the University of Plymouth and funded by an International Research and Collaboration grant to Haline E. Schendan. The authors are grateful for the assistance of Roderick Elias, B. S., Dr. Stephen M. Maher, and Lisa C. Lucia, M.Sc. for recording the EEG data.

REFERENCES

Allison, T., Puce, A., Spencer, D. D., and Mccarthy, G. (1999). Electrophysiological studies of human face perception I: potentials generated in occipitotemporal cortex by face and non-face stimuli. *Cereb. Cortex* 9, 415–430.

Anderson, N. D., and Wilson, H. R. (2005). The nature of synthetic face adaptation. *Vision Res.* 45, 1815–1828.

Anderson, S. E., Chiu, E., Huette, S., and Spivey, M. J. (2011). On the temporal dynamics of language-mediated vision and vision-mediated language. *Acta Psychol. (Amst.)* 137, 181–189.

Auchterlonie, S., Phillips, N. A., and Chertkow, H. (2002). Behavioral and electrical brain measures of semantic priming in patients with Alzheimer's disease: implications for access failure versus deterioration hypotheses. *Brain Cogn.* 48, 264–267.

Bach, P., Gunter, T. C., Knoblich, G., Prinz, W., and Friederici, A. D. (2009). N400-like negativities in action perception reflect the activation of two components of an action representation. *Soc. Neurosci.* 4, 212–232.

Bajo, M. T. (1988). Semantic facilitation with pictures and words. *J. Exp. Psychol. Learn. Mem. Cogn.* 14, 579–589.

Barrett, S. E., and Rugg, M. D. (1989). Event-related potentials and the semantic matching of faces. *Neuropsychologia* 27, 913–922.

Barrett, S. E., and Rugg, M. D. (1990). Event-related potentials and the semantic matching of pictures. *Brain Cogn.* 14, 201–212.

Barsalou, L. W. (2008). Grounded cognition. *Annu. Rev. Psychol.* 59, 617–645.

Barsalou, L. W. (2009). Simulation, situated conceptualization, and prediction. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 364, 1281–1289.

Bentin, S., and McCarthy, G. (1994). The effects of immediate stimulus repetition on reaction time and event-related potentials in tasks of different complexity. *J. Exp. Psychol. Learn. Mem. Cogn.* 20, 130–149.

Biggs, T. C., and Marmurek, H. H. C. (1990). Picture and word naming – is facilitation due to processing overlap. *Am. J. Psychol.* 103, 81–100.

Blackford, T., Holcomb, P. J., Grainger, J., and Kuperberg, G. R. (2012). A funny thing happened on the way to articulation: N400 attenuation despite behavioral interference in picture naming. *Cognition* 123, 84–99.

Brown, A. S., Neblett, D. R., Jones, T. C., and Mitchell, D. B. (1991). Transfer of processing in repetition priming: some inappropriate findings. *J. Exp. Psychol. Learn. Mem. Cogn.* 17, 514–525.

Bub, D. N., and Masson, M. E. (2010). On the nature of hand-action representations evoked during written sentence comprehension. *Cognition* 116, 394–408.

Bullier, J. (2001). Integrated model of visual processing. *Brain Res. Brain Res. Rev.* 36, 96–107.

Burton, A. M., Jenkins, R., and Schweinberger, S. R. (2011). Mental representations of familiar faces. *Br. J. Psychol.* 102, 943–958.

Cabeza, R., Burton, A. M., Kelly, S. W., and Akamatsu, S. (1997). Investigating the relation between imagery and perception: evidence from face priming. *Q. J. Exp. Psychol. A* 50, 274–289.

Carr, T. H., Mccauley, C., Sperber, R. D., and Parmelee, C. M. (1982). Words, pictures, and priming: on semantic activation, conscious identification, and the automaticity of information processing. *J. Exp. Psychol. Hum. Percept. Perform.* 8, 757–777.

Chatterjee, A. (2010). Disembodying cognition. *Lang. Cogn.* 2, 79–116.

Clarke, A., Taylor, K. I., and Tyler, L. K. (2011). The evolution of meaning: spatio-temporal dynamics of visual object recognition. *J. Cogn. Neurosci.* 23, 1887–1899.

Collins, A. M., and Loftus, E. F. (1975). A spreading activation theory of semantic processing. *Psychol. Rev.* 82, 407–428.

Costen, N. P., Parker, D. M., and Craw, I. (1996). Effects of high-pass and low-pass spatial filtering on face identification. *Percept. Psychophys.* 58, 602–612.

Craver-Lemley, C., and Reeves, A. (1992). How visual imagery interferes with vision. *Psychol. Rev.* 99, 633–649.

David, O., Harrison, L., and Friston, K. J. (2005). Modeling event-related responses in the brain. *Neuroimage* 25, 756–770.

David, O., Kiebel, S. J., Harrison, L. M., Mattout, J., Kilner, J. M., and Friston, K. J. (2006). Dynamic causal modeling of evoked responses in EEG and MEG. *Neuroimage* 30, 1255–1272.

Dien, J., Spencer, K. M., and Donchin, E. (2004). Parsing the late positive complex: mental chronometry and the ERP components that inhabit the neighborhood of the P300. *Psychophysiology* 41, 665–678.

Dobel, C., Junghofer, M., Breitenstein, C., Klauke, B., Knecht, S., Pantev, C., and Zwitserlood, P. (2010). New names for known things: on the association of novel word forms with existing semantic information. *J. Cogn. Neurosci.* 22, 1251–1261.

Doniger, G. M., Foxe, J. J., Murray, M. M., Higgins, B. A., Snodgrass, J. G., Schroeder, C. E., and Javitt, D. C. (2000). Activation timecourse of ventral visual stream object-recognition areas: high density electrical mapping of perceptual closure processes. *J. Cogn. Neurosci.* 12, 615–621.

Doniger, G. M., Foxe, J. J., Schroeder, C. E., Murray, M. M., Higgins, B. A., and Javitt, D. C. (2001). Visual perceptual learning in human object recognition areas: a repetition priming study using high-density electrical mapping. *Neuroimage* 13, 305–313.

Downing, P. E., Chan, A. W., Peelen, M. V., Dodds, C. M., and Kanwisher, N. (2006). Domain specificity in visual cortex. *Cereb. Cortex* 16, 1453–1461.

Eddy, M. D., and Holcomb, P. J. (2009). Electrophysiological evidence for size invariance in masked picture repetition priming. *Brain Cogn.* 71, 397–409.

Eddy, M. D., and Holcomb, P. J. (2010). The temporal dynamics of masked repetition picture priming effects: manipulations of stimulus-onset asynchrony (SOA) and prime duration. *Brain Res.* 1340, 24–39.

Eddy, M. D., and Holcomb, P. J. (2011). Invariance to rotation in depth measured by masked repetition priming is dependent on prime duration. *Brain Res.* 1424, 38–52.

Eddy, M. D., Schmid, A., and Holcomb, P. J. (2006). Masked repetition priming and event-related brain potentials: a new approach for tracking the time-course of object perception. *Psychophysiology* 43, 564–568.

Eddy, M. D., Schnyer, D., Schmid, A., and Holcomb, P. J. (2007). Spatial dynamics of masked picture repetition effects. *Neuroimage* 34, 1723–1732.

Farah, M. J., Peronnet, F., Gonon, M. A., and Giard, M. H. (1988). Electrophysiological evidence for a shared representational medium for visual images and visual percepts. *J. Exp. Psychol. Gen.* 117, 248–257.

Federmeier, K. D., and Kutas, M. (2001). Meaning and modality: influences of context, semantic memory organization, and perceptual predictability on picture processing. *J. Exp. Psychol. Learn. Mem. Cogn.* 27, 202–224.

Federmeier, K. D., and Kutas, M. (2002). Picture the difference: electrophysiological investigations of picture processing in the two cerebral hemispheres. *Neuropsychologia* 40, 730–747.

Feldman, L. B. (2000). Are morphological effects distinguishable from the effects of shared meaning and shared form? *J. Exp. Psychol. Learn. Mem. Cogn.* 26, 1431–1444.

Fischer, M. H., and Zwaan, R. A. (2008). Embodied language: a review of the role of the motor system in language comprehension. *Q. J. Exp. Psychol. (Colchester)* 61, 825–850.

Fodor, J. A. (1983). *The Modularity of Mind*. Cambridge, MA: MIT Press.

Forster, K. I., and Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. *J. Exp. Psychol. Learn. Mem. Cogn.* 10, 680–698.

Franklin, M. S., Dien, J., Neely, J. H., Huber, E., and Waterson, L. D. (2007). Semantic priming modulates the N400, N300, and N400RP. *Clin. Neurophysiol.* 118, 1053–1068.

Freyd, J. J., and Finke, R. A. (1984). Facilitation of length discrimination using real and imaged context frames. *Am. J. Psychol.* 97, 323–341.

Ganis, G., and Kosslyn, S. M. (2007). "Multiple mechanisms of top-down processing in vision," in *Representation and Brain*, ed. S. Funahashi (Tokyo: Springer Verlag), 21–45.

Ganis, G., and Kutas, M. (2003). An electrophysiological study of scene effects on object identification. *Brain Res. Cogn. Brain Res.* 16, 123–144.

Ganis, G., Kutas, M., and Sereno, M. I. (1996). The search for common sense: an electrophysiological study of the comprehension of words and pictures in reading. *J. Cogn. Neurosci.* 8, 89–106.

Ganis, G., and Schendan, H. E. (2008). Visual mental imagery and perception produce opposite adaptation effects on early brain potentials. *Neuroimage* 42, 1714–1727.

Ganis, G., and Schendan, H. E. (2011). Mental imagery. *Wiley Interdiscip. Rev. Cogn. Sci.* 2, 239–252.

Ganis, G., Schendan, H. E., and Kosslyn, S. M. (2007). Neuroimaging evidence for object model verification theory: role of prefrontal control in visual object categorization. *Neuroimage* 34, 384–398.

Ganis, G., Smith, D., and Schendan, H. E. (2012). The N170, not the P1, indexes the earliest time for categorical perception of faces, regardless of interstimulus variance. *Neuroimage* 62, 1563–1574.

Ganis, G., Thompson, W. L., Mast, F. W., and Kosslyn, S. M. (2003). Visual imagery in cerebral visual dysfunction. *Neurol. Clin.* 21, 631–646.

Grill-Spector, K., Henson, R., and Martin, A. (2006). Repetition and the brain: neural models of stimulus-specific effects. *Trends Cogn. Sci. (Regul. Ed.)* 10, 14–23.

Hasson, U., Harel, M., Levy, I., and Malach, R. (2003). Large-scale mirror-symmetry organization of human occipito-temporal object areas. *Neuron* 37, 1027–1041.

Heil, M. (2002). The functional significance of ERP effects during mental rotation. *Psychophysiology* 39, 535–545.

Henson, R. N., Goshen-Gottstein, Y., Ganel, T., Otten, L. J., Quayle, A., and Rugg, M. D. (2003). Electrophysiological and haemodynamic correlates of face perception, recognition and priming. *Cereb. Cortex* 13, 793–805.

Henson, R. N., Rylands, A., Ross, E., Vuilleumeir, P., and Rugg, M. D. (2004). The effect of repetition lag on electrophysiological and haemodynamic correlates of visual object priming. *Neuroimage* 21, 1674–1689.

Herholz, S. C., Halpern, A. R., and Zatorre, R. J. (2012). Neuronal correlates of perception, imagery, and memory for familiar tunes. *J. Cogn. Neurosci.* 24, 1382–1397.

Hirschfeld, G., Feldker, K., and Zwitserlood, P. (2012). Listening to "flying ducks": individual differences in sentence-picture verification investigated with ERPs. *Psychophysiology* 49, 312–321.

Hirschfeld, G., and Zwitserlood, P. (2011). How vision is shaped by language comprehension – top-down feedback based on low-spatial frequencies. *Brain Res.* 1377, 78–83.

Holcomb, P. J. (2003). *"ERPSystem," 2.0a ed.* Medford: Tufts University.

Holcomb, P. J., and McPherson, W. B. (1994). Event-related brain potentials reflect semantic priming in an object decision task. *Brain Cogn.* 24, 259–276.

Jeffreys, D. A. (1996). Evoked potential studies of face and object processing. *Vis. Cogn.* 3, 1–38.

Johnson, J. S., and Olshausen, B. A. (2003). Time course of neural signatures of object recognition. *J. Vis.* 3, 499–512.

Johnson, J. S., and Olshausen, B. A. (2005). The earliest EEG signatures of object recognition in a cued-target task are postsensory. *J. Vis.* 5, 299–312.

Johnson-Laird, P. N. (1983). *Mental Models*. Cambridge, MA: Harvard University Press.

Kazmerski, V. A., and Friedman, D. (1997). Old/new differences in direct and indirect memory tests using pictures and words in within- and cross-form conditions: event-related potential and behavioral measures. *Brain Res. Cogn. Brain Res.* 5, 255–272.

Kemmerer, D., Castillo, J. G., Talavage, T., Patterson, S., and Wiley, C. (2008). Neuroanatomical distribution of five semantic components of verbs: evidence from fMRI. *Brain Lang.* 107, 16–43.

Kiefer, M., Sim, E. J., Helbig, H., and Graf, M. (2011). Tracking the time course of action priming on object recognition: evidence for fast and slow influences of action on perception. *J. Cogn. Neurosci.* 23, 1864–1874.

Koester, D., and Schiller, N. O. (2008). Morphological priming in overt language production: electrophysiological evidence from Dutch. *Neuroimage* 42, 1622–1630.

Kosslyn, S. M. (1994). *Image and Brain: The Resolution of the Imagery Debate.* Cambridge, MA: MIT Press.

Kosslyn, S. M., Alpert, N. M., Thompson, W. L., Chabris, C. F., Rauch, S. L., and Anderson, A. K. (1994). Identifying objects seen from different viewpoints. A PET investigation. *Brain* 117, 1055–1071.

Kosslyn, S. M., Ganis, G., and Thompson, W. L. (2001). Neural foundations of imagery. *Nat. Rev. Neurosci.* 2, 635–642.

Kosslyn, S. M., Thompson, W. L., and Ganis, G. (2006). *The Case for Mental Imagery*. New York, NY: Oxford University Press.

Kousta, S. T., Vigliocco, G., Vinson, D. P., Andrews, M., and Del Campo, E. (2011). The representation of abstract words: why emotion matters. *J. Exp. Psychol. Gen.* 140, 14–34.

Kozhevnikov, M., Kosslyn, S., and Shephard, J. (2005). Spatial versus object visualizers: a new characterization of visual cognitive style. *Mem. Cognit.* 33, 710–726.

Kutas, M., and Federmeier, K. D. (2011). Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annu. Rev. Psychol.* 62, 621–647.

Lau, E. F., Phillips, C., and Poeppel, D. (2008). A cortical network for semantics: (de)constructing the N400. *Nat. Rev. Neurosci.* 9, 920–933.

Loffler, G., Yourganov, G., Wilkinson, F., and Wilson, H. R. (2005). fMRI evidence for the neural representation of faces. *Nat. Neurosci.* 8, 1386–1390.

Luck, S. J. (2005). *An Introduction to the Event-Related Potential Technique.* Cambridge, MA: MIT Press.

Mahon, B. Z., and Caramazza, A. (2009). Concepts and categories: a cognitive neuropsychological perspective. *Annu. Rev. Psychol.* 60, 27–51.

Marinkovic, K., Dhond, R. P., Dale, A. M., Glessner, M., Carr, V., and Halgren, E. (2003). Spatiotemporal dynamics of modality-specific and supramodal word processing. *Neuron* 38, 487–497.

McDermott, K. B., and Roediger, H. L. III. (1994). Effects of imagery on perceptual implicit memory tests. *J. Exp. Psychol. Learn. Mem. Cogn.* 20, 1379–1390.

McPherson, W. B., and Holcomb, P. J. (1999). An electrophysiological investigation of semantic priming with pictures of real objects. *Psychophysiology* 36, 53–65.

Meteyard, L., Cuadrado, S. R., Bahrami, B., and Vigliocco, G. (2011). Coming of age: a review of embodiment and the neuroscience of semantics. *Cortex* 48, 788–804.

Michelon, P., and Koenig, O. (2002). On the relationship between imagery and visual perception: evidence from priming studies. *Eur. J. Cogn. Psychol.* 14, 161–184.

Michelon, P., Snyder, A. Z., Buckner, R. L., Mcavoy, M., and Zacks, J. M. (2003). Neural correlates of incongruous visual information. An event-related fMRI study. *Neuroimage* 19, 1612–1626.

Michelon, P., and Zacks, J. M. (2003). What is primed in priming from imagery? *Psychol. Res.* 67, 71–79.

Morris, C. D., Bransford, J. D., and Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *J. Verbal Learn. Verbal Behav.* 16, 519–533.

Morton, J. (1979). "Facilitation in word recognition: experiments causing change in the Logogen model," in *Proceedings of the Conference on the Processing of Visible Language*, eds P. A. Kolers, M. E. Wrolstad, and H. Bouma (New York: Plenum), 259–268.

Moulton, S. T., and Kosslyn, S. M. (2009). Imagining predictions: mental imagery as mental emulation. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 364, 1273–1280.

Nemrodov, D., and Itier, R. J. (2012). Is the rapid adaptation paradigm too rapid? Implications for face and object processing. *Neuroimage* 61, 812–822.

Nessler, D., Mecklinger, A., and Penney, T. B. (2005). Perceptual fluency, semantic familiarity and recognition-related familiarity: an electrophysiological exploration. *Brain Res. Cogn. Brain Res.* 22, 265–288.

Neumann, M. F., and Schweinberger, S. R. (2008). N250r and N400 ERP correlates of immediate famous face repetition are independent of perceptual load. *Brain Res.* 1239, 181–190.

Nielsen-Bohlman, L., and Knight, R. T. (1994). Electrophysiological dissociation of rapid memory mechanisms in humans. *Neuroreport* 5, 1517–1521.

O'Hare, A. J., Dien, J., Waterson, L. D., and Savage, C. R. (2008). Activation of the posterior cingulate by semantic priming: a co-registered ERP/fMRI study. *Brain Res.* 1189, 97–114.

Paivio, A., and Sadoski, M. (2011). Lexicons, contexts, events, and images: commentary on Elman (2009) from the perspective of dual coding theory. *Cogn. Sci.* 35, 198–209.

Pascual-Marqui, R. D. (2002). Standardized low resolution electromagnetic tomography (sLORETA): technical details. *Methods Find. Exp. Clin. Pharmacol.* 24, 5–12.

Peterson, M. J., and Graham, S. E. (1974). Visual detection and visual imagery. *J. Exp. Psychol.* 103, 509–514.

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. (2006). Temporal characterization of the neural correlates of perceptual decision making in the human brain. *Cereb. Cortex* 16, 509–518.

Philiastides, M. G., and Sajda, P. (2007). EEG-informed fMRI reveals spatiotemporal characteristics of perceptual decision making. *J. Neurosci.* 27, 13082–13091.

Puce, A., Allison, T., and Mccarthy, G. (1999). Electrophysiological studies of human face perception III: effects of top-down processing on face-specific potentials. *Cereb. Cortex* 9, 445–458.

Puce, A., and Perrett, D. (2003). Electrophysiology and brain imaging of biological motion. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 358, 435–445.

Pulvermuller, F. (1999). Words in the brain's language. *Behav. Brain Sci.* 22, 253–279; discussion 280–336.

Pylyshyn, Z. (2003). Return of the mental image: are there really pictures in the brain? *Trends Cogn. Sci. (Regul. Ed.)* 7, 113–118.

Rossell, S. L., Price, C. J., and Nobre, A. C. (2003). The anatomy and time course of semantic priming investigated by fMRI and ERPs. *Neuropsychologia* 41, 550–564.

Rugg, M. D., and Curran, T. (2007). Event-related potentials and recognition memory. *Trends Cogn. Sci. (Regul. Ed.)* 11, 251–257.

Rumiati, R. I., Papeo, L., and Corradi-Dell'acqua, C. (2010). Higher-level motor processes. *Ann. N. Y. Acad. Sci.* 1191, 219–241.

Schacter, D. L., Addis, D. R., and Buckner, R. L. (2008). Episodic simulation of future events: concepts, data, and applications. *Ann. N. Y. Acad. Sci.* 1124, 39–60.

Schacter, D. L., Wig, G. S., and Stevens, W. D. (2007). Reductions in cortical activity during priming. *Curr. Opin. Neurobiol.* 17, 171–176.

Schendan, H. E., Ganis, G., and Kutas, M. (1998). Neurophysiological evidence for visual perceptual categorization of words and faces within 150 ms. *Psychophysiology* 35, 240–251.

Schendan, H. E., Kanwisher, N. G., and Kutas, M. (1997). Early brain potentials link repetition blindness, priming and novelty detection. *Neuroreport* 8, 1943–1948.

Schendan, H. E., and Kutas, M. (2002). Neurophysiological evidence for two processing times for visual object identification. *Neuropsychologia* 40, 931–945.

Schendan, H. E., and Kutas, M. (2003). Time course of processes and representations supporting visual object identification and memory. *J. Cogn. Neurosci.* 15, 111–135.

Schendan, H. E., and Kutas, M. (2007). Neurophysiological evidence for the time course of activation of global shape, part, and local contour representations during visual object categorization and memory. *J. Cogn. Neurosci.* 19, 734–749.

Schendan, H. E., and Lucia, L. C. (2009). Visual object cognition precedes but also temporally overlaps mental rotation. *Brain Res.* 1294, 91–105.

Schendan, H. E., and Lucia, L. C. (2010). Object-sensitive activity reflects earlier perceptual and later cognitive processing of visual objects between 95 and 500 ms. *Brain Res.* 1329, 124–141.

Schendan, H. E., and Maher, S. M. (2009). Object knowledge during entry-level categorization is activated and modified by implicit memory after 200 ms. *Neuroimage* 44, 1423–1438.

Schendan, H. E., and Stern, C. E. (2008). Where vision meets memory: prefrontal-posterior networks for visual object constancy during categorization and recognition. *Cereb. Cortex* 18, 1695–1711.

Scott, L. S., Tanaka, J. W., Sheinberg, D. L., and Curran, T. (2006). A reevaluation of the electrophysiological correlates of expert object processing. *J. Cogn. Neurosci.* 18, 1453–1465.

Sehatpour, P., Molholm, S., Javitt, D. C., and Foxe, J. J. (2006). Spatiotemporal dynamics of human object recognition processing: an integrated high-density electrical mapping and functional imaging study of "closure" processes. *Neuroimage* 29, 605–618.

Sehatpour, P., Molholm, S., Schwartz, T. H., Mahoney, J. R., Mehta, A. D., Javitt, D. C., Stanton, P. K., and Foxe, J. J. (2008). A human intracranial study of long-range oscillatory coherence across a frontal-occipital-hippocampal brain network during visual object processing. *Proc. Natl. Acad. Sci. U.S.A.* 105, 4399–4404.

Sitnikova, T., Goff, D., and Kuperberg, G. R. (2009). Neurocognitive abnormalities during comprehension of real-world goal-directed behaviors in schizophrenia. *J. Abnorm. Psychol.* 118, 256–277.

Sitnikova, T., Perrone, C., Goff, D., and Kuperberg, G. R. (2010). Neurocognitive mechanisms of conceptual processing in healthy adults and patients with schizophrenia. *Int. J. Psychophysiol.* 75, 86–99.

Spironelli, C., Galfano, G., Umilta, C., and Angrilli, A. (2011). Word position affects stimulus recognition: evidence for early ERP short-term plastic modulation. *Int. J. Psychophysiol.* 82, 217–224.

Stanfield, R. A., and Zwaan, R. A. (2001). The effect of implied orientation derived from verbal context on picture recognition. *Psychol. Sci.* 12, 153–156.

Thierry, G., Martin, C. D., Downing, P., and Pegna, A. J. (2007). Controlling for interstimulus perceptual variance abolishes N170 face selectivity. *Nat. Neurosci.* 10, 505–511.

Tsao, D. Y., Freiwald, W. A., Tootell, R. B., and Livingstone, M. S. (2006). A cortical region consisting entirely of face-selective cells. *Science* 311, 670–674.

Tucker, M., and Ellis, R. (1998). On the relations between seen objects and components of potential actions. *J. Exp. Psychol. Hum. Percept. Perform.* 24, 830–846.

Tulving, E., and Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychol. Rev.* 80, 359–380.

Vigliocco, G., Warren, J., Siri, S., Arciuli, J., Scott, S., and Wise, R. (2006). The role of semantics and grammatical class in the neural representation of words. *Cereb. Cortex* 16, 1790–1796.

Vogel, E. K., and Luck, S. J. (2000). The visual N1 component as an index of a discrimination process. *Psychophysiology* 37, 190–203.

Voss, J. L., Schendan, H. E., and Paller, K. A. (2010). Finding meaning in novel geometric shapes influences electrophysiological correlates of repetition and dissociates perceptual and conceptual priming. *Neuroimage* 49, 2879–2889.

Warren, C., and Morton, J. (1982). The effects of priming on picture recognition. *Br. J. Psychol.* 73, 117–129.

Wassenburg, S. I., and Zwaan, R. A. (2010). Readers routinely represent implied object rotation: the role of visual experience. *Q. J. Exp. Psychol. (Colchester)* 63, 1665–1670.

Wilson, H. R., and Humanski, R. A. (1993). Spatial frequency adaptation and contrast gain control. *Vision Res.* 33, 1133–1149.

Wilson, M. (2002). Six views of embodied cognition. *Psychon. Bull. Rev.* 9, 625–636.

Zhang, X. L., Begleiter, H., Porjesz, B., Wang, W. Y., and Litke, A. (1995). Event related potentials during object recognition tasks. *Brain Res. Bull.* 38, 531–538.

Zwaan, R. A., Stanfield, R. A., and Yaxley, R. H. (2002). Language comprehenders mentally represent the shapes of objects. *Psychol. Sci.* 13, 168–171.

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: 07 May 2012; accepted: 17 August 2012; published online: 14 September 2012.

Citation: Schendan HE and Ganis G (2012) Electrophysiological potentials reveal cortical mechanisms for mental imagery, mental simulation, and grounded (embodied) cognition. Front. Psychology ***3****:329. doi: 10.3389/fpsyg.2012.00329*
This article was submitted to Frontiers in Perception Science, a specialty of Frontiers in Psychology.

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 18 January 2013
doi: 10.3389/fpsyg.2012.00561

A cross-modal perspective on the relationships between imagery and working memory

Lora T. Likova*

The Smith-Kettlewell Eye Research Institute, San Francisco, CA, USA

Edited by:
Joel Pearson, The University of New South Wales, Australia

Reviewed by:
Holly Bridge, University of Oxford, UK
Axel Kohler, University of Münster, Germany

Correspondence:
Lora T. Likova, The Smith-Kettlewell Eye Research Institute, 2318 Fillmore Street, San Francisco, CA, USA.
e-mail: lora@ski.org

Mapping the distinctions and interrelationships between imagery and working memory (WM) remains challenging. Although each of these major cognitive constructs is defined and treated in various ways across studies, most accept that both imagery and WM involve a form of internal representation available to our awareness. In WM, there is a further emphasis on goal-oriented, active maintenance, and use of this conscious representation to guide voluntary action. Multicomponent WM models incorporate representational buffers, such as the visuo-spatial sketchpad, plus central executive functions. If there is a visuo-spatial "sketchpad" for WM, does imagery involve the same representational buffer? Alternatively, does WM employ an imagery-specific representational mechanism to occupy our awareness? Or do both constructs utilize a more generic "projection screen" of an amodal nature? To address these issues, in a cross-modal fMRI study, I introduce a novel Drawing-Based Memory Paradigm, and conceptualize drawing as a complex behavior that is readily adaptable from the visual to non-visual modalities (such as the tactile modality), which opens intriguing possibilities for investigating cross-modal learning and plasticity. Blindfolded participants were trained through our Cognitive-Kinesthetic Method (Likova, 2010a, 2012) to draw complex objects guided purely by the memory of felt tactile images. If this WM task had been mediated by transfer of the felt spatial configuration to the visual imagery mechanism, the response-profile in visual cortex would be predicted to have the "top-down" signature of propagation of the imagery signal *downward* through the visual hierarchy. Remarkably, the pattern of cross-modal occipital activation generated by the non-visual memory drawing was essentially the inverse of this typical imagery signature. The sole visual hierarchy activation was isolated to the primary visual area (V1), and accompanied by deactivation of the entire extrastriate cortex, thus 'cutting-off' any signal propagation from/to V1 through the visual hierarchy. The implications of these findings for the debate on the interrelationships between the core cognitive constructs of WM and imagery and the nature of internal representations are evaluated.

Keywords: visual imagery, visuo-spatial sketchpad, working memory, primary visual cortex V1, drawing, fMRI

INTRODUCTION

Mapping the distinctions and interrelationships between imagery and working memory (WM) remains challenging. Although each of these major cognitive constructs is defined and treated in various ways across studies, most accept that both imagery and WM involve a type of internal representation available to our awareness; in WM, however, there is a further emphasis on goal-oriented, active maintenance and use of this conscious representation to guide voluntary action.

WM refers to the temporary storage and manipulation of information, and is invoked as the mechanism for information processing during the performance of a wide range of everyday tasks (e.g., Baddeley and Hitch, 1974; Baddeley, 1986, 1992, 2000, 2003; Logie et al., 1989; Logie and Marchetti, 1991; Logie, 1995; Baddeley and Andrade, 2000). Initially, the proposed structure included the central executive component and two active storage buffers – the *visuo-spatial sketchpad* and the *articulatory/phonological loop*. While the visuo-spatial sketchpad is considered to be responsible for the temporarily storage and manipulation of visuo-spatial material, the phonological loop is posited to provide a similar function for verbal material. An enhanced version of the multicomponent WM model added an *episodic buffer* (Baddeley, 2003).

Interestingly, the involvement of the early visual cortex, and area V1 in particular, has been a critical issue in discussions both on the neural substrate of the putative visuo-spatial sketchpad and on the nature of imagery. The V1-substrate propositions are based mainly on the fact that, although most areas in occipital cortex are topographically organized, area V1 has the unique status of being the *largest topographic map* in the brain, with the *highest spatial resolution*, in addition to parallel processing of the information from the whole map surface (in contrast to the sequential processing in some other modalities).

All of these features are critically important for a successful "sketchpad" implementation, which is why previous theoretical as well as neurophysiological studies in non-human primates (e.g., Mumford, 1991, 1996; Lee et al., 1998; Super et al., 2001a,b; Lee and Mumford, 2003; Super, 2003) had suggested V1 as the source of the high-resolution visuo-spatial "sketchpad" function: "instead of being the first stage in a feedforward pipeline, V1 is better described as the unique high-resolution buffer in the visual system" (Lee and Mumford, 2003).

However, the same characteristics of this region are also key requirements for the existence of a pictorial-code form of visual imagery. Thus, the issue of V1 involvement has been central in the long-standing debate about the nature of visual imagery, which relates to the question of whether the imagery "code" is pictorial or propositional (Kosslyn et al., 2001; Kosslyn and Thompson, 2003).

Neuroimaging is a valuable tool that can help to resolve these issues, but while there has been much neuroimaging work on imagery, this is not the case with the putative WM sketchpad. The neural substrate for visual imagery has been found to largely *overlap* with that for visual perception (e.g., Ishai and Sagi, 1995; Kosslyn et al., 1999; Kreiman et al., 2000; O'Craven and Kanwisher, 2000; Kosslyn et al., 2001; Kosslyn and Thompson, 2003; Mechelli et al., 2004), with the activation pattern implying that the signal propagates from higher cortical regions in a *top-down* manner through the visual hierarchy toward V1. The resultant top-down gradient of activation provides a notable signature of the visual imagery activation pattern; consistent with this employment of the visual pathway for imagery, there is no significant negative signal (i.e., no flow interruption) for imagery in occipital cortex (Ganis et al., 2004).

While a substantial activation in the higher areas has been consistently found across the imagery studies, this has not been the case with V1. Although some level of V1 activation during imagery has been reported in several studies (e.g., Kosslyn et al., 1993; Le Bihan et al., 1993; Sabbah et al., 1995; Kosslyn et al., 1996; Chen et al., 1998; Shin et al., 1999; Thompson et al., 2001; Ishai et al., 2002; Lambert et al., 2002; Ganis et al., 2004), a larger number of studies did not find any V1 activation at all (e.g., Goldenberg et al., 1991; Charlot et al., 1992; Mellet et al., 1995, 1998a,b; D'Esposito et al., 1997; Ishai et al., 2000; Knauff et al., 2000; Trojano et al., 2000; Wheeler et al., 2000; Formisano et al., 2002; Sack et al., 2002; Mazard et al., 2004; Kaas et al., 2010). Importantly, even when V1 was activated during imagery, the signal there was *significantly weaker* than in the extrastriate visual areas. Thus, the level of V1 activation is of great importance for the imagery debate, as imagery activation in V1 implies the usage of a pictorial code. Kaas et al. (2010), however, whose primary goal was "to eliminate the effects of (short- or long-term) memory in the investigation of the effects of mental imagery," conclude that their results "suggest that the activation in early visual areas observed in previous imagery studies might be related to short- or long-term memory retrieval of specific sensory experiences."

To analyze the disparate results on the V1 involvement in the imagery literature, Kosslyn and Thompson (2003) defined sets of variables associated with each of three theories, which were then fit to the observed results using logistic regression analysis to discover how well each theory predicted when early visual cortex was activated. The three theories were Perceptual Anticipatory Theory (pictorial imagery coding), Propositional Theory (non-pictorial propositional imagery coding), and Methodological Factors (determining factors, such as low neuroimaging resolution or differential resting activation in postulated imagery loci, that may need to be controlled in order to resolve imagery-specific activation). Their analysis identified three variables that optimally predicted the differences in the probability of activation across imagery studies. Notably, two of the variables were *task-dependent* requirements (the requirement to note *high-resolution details* in the stimuli and the requirement to visualize *shapes* rather than abstract spatial relations), while the third was purely technical (sufficiently high sensitivity of the technique).

Thus, the operation of the (pictorial) visual imagery and of the visuo-spatial sketchpad concepts share (i) similar task requirements, and (ii) similar need of V1 usage (although note that, in contrast to the putative sketchpad, the top-down theories of imagery are not restricted to V1 but require activation of the whole visual hierarchy). A number of behavioral studies have addressed possible interactions between WM and visual imagery (e.g., Bruyer and Scailquin, 1998; Baddeley, 2000), with the most recent by Keogh and Pearson (2011) suggesting an imagery-dependent dichotomy in cognitive strategies for visual WM.

QUESTIONS

Despite the array of studies on the issue, the above review indicates that the neural substrates as well as the interactions between these two cognitive constructs are still far from being definitively resolved. In particular, the visuo-spatial sketchpad remains an almost entirely theoretical construct. If there is a visuo-spatial "sketchpad" for WM, does imagery use the same representational buffer? Alternatively, does WM employ an imagery-specific representational mechanism to occupy our awareness? Or do both constructs utilize a more generic "projection screen" of an amodal nature?

A NOVEL APPROACH

The drawing-based memory paradigm

Likova (2010a, 2012) recently conceptualized the drawing task as the basis for a novel memory paradigm to address these questions. Drawing, and in particular memory-guided drawing, challenges both the encoding of detailed spatial representations and their explicit retrieval from memory for "projection" back onto a mental high-resolution "screen" to guide the movements of the drawing hand with the requisite precision. A cortical region such as V1 would be an ideal neural implementation of the required "screen"; thus, the putative *V1 visuo-spatial sketchpad* is a plausible theoretical construct that provides for memory retrieval for just the kinds of spatial representations involved in the drawing task, allowing for the active maintenance of information about stimuli no longer in view. (It is not by chance that this ubiquitous tool of the real drawing process – the use of a disposable sketchpad – was the metaphor employed for the memory module in question.)

Importantly, the drawing-from-tactile-memory task effectively transcends simple "recognition memory." An easy demonstration makes the point: Close your eyes and try to imagine the objects on your desk, the face of your close friend, or even your own face; in particular, try to "see" the detailed shapes as though you are going to draw them. It is amazing how misleading the feeling is that we "know" any of those very well. Despite the fact that we would effortlessly recognize them (*recognition memory*), when we try to retrieve the details, they fade or somehow escape our grasp, although we do not feel these information gaps when recognizing immediately the object as a whole. Many early representational details seem to be lost along the passage through the visual pathway before they have been integrated into the internal reconstruction of the face/object that they represent. In other words, recognition memory seems to operate at the higher-level of object category processing and does not need to retain the more "local" level of detail. In contrast, a highly detailed kind of spatial memory is engaged to meet the needs of the efferent drawing task; this "*memory-for-drawing*" preserves and recalls details sufficient to enable the complex spatiomotor act of producing an accurate drawing.

Conceptual framework

Both the novel Drawing-Based Memory Paradigm and the Cognitive-Kinesthetic Training Method, are based on a framework of principles (Likova, 2012), including:

1. *Space transcends any specific sensory modality*. As emphasized by the phenomenon of drawing by the blind (e.g., Kennedy, 1993, 2000; Kennedy and Igor, 2003; Kennedy and Juricevic, 2006; Ponchillia, 2008; Likova, 2010a, 2012), space, and spatial structure are not represented solely by the visual modality. The visual system is best suited to process spatial information, but it is not the only one. Thus, when deprived of visual input, the brain is capable of employing the "free" visual resources in the most relevant way. (As there is an ambiguity in the use of the term "spatial," particularly in the WM and imagery literature, note that when used in this paper, "spatial" refers to the perception of any spatial structure – 2D or 3D, static or dynamic – *independently* of the sensory modality exploring it. For example, a face can be recognized by exploring its spatial structure with the hands, or a geometric function can be "seen" by audio-graphics, etc.) My view is that *drawing* deals with spatial structures in this general sense, and consequently it has the advantage that can readily be "translated" from a visual into a tactile form.
2. *Closing the perception-cognition-action loop is a powerful amplifier for learning*, so the *task* selection is critical. *Drawing* is a complex task precisely orchestrating multiple brain mechanisms, and consequently, it provides for an integrative, perception-cognition-action paradigm.
3. *Training in highly engaging unfamiliar task*s that provide *fun and inspiring outcomes* is a fruitful paradigm for driving brain reorganization and assessing its earliest stages. *Drawing* inherently embodies all of these components, particularly when studied under the circumstances of visual deprivation (which are unusual for a "visual" art); the characteristic for drawing sense of completion, creativity, and fulfillment amplify the experience-based plasticity.
4. *Tasks demanding detailed re-expression of memory-representations force the development of precise and robust memory. Drawing*-from-memory demands such explicit *re-expression* through the motor loop, and consequently it demands "high-resolution" internal representations to be communicated back through the drawing act.
5. *Studies on memory would highly benefit from tasks providing "direct" memory-control. Drawing*-from-memory incorporates such direct control by providing direct memory "readout," as it ensures an explicit expression of the remembered information by externalization of the mental representation that guides the drawing hand.

These considerations led me to the choice of *non-visual drawing*, which incorporates all of the above principles, as a paradigm for both *training* and *studying* cross-modal memory. The role of WM and imagery in mediating the training effects were evaluated by functional Magnetic Resonance Imaging (fMRI).

The cognitive-kinesthetic drawing method

To employ this novel memory paradigm for studying learning-based plasticity, a method is needed to train non-sighted people to draw not simply without visual feedback, but guided solely by *non-visual* memory. Recently I have developed a novel technique, the Cognitive-Kinesthetic Drawing Method, which proved to be very effective in the successful training of people under total visual deprivation. Congenitally blind, late-onset blind and blindfolded were successfully trained in only a week of 1.5 h/day sessions to draw complex face and object structures (as opposed to simple geometric or grid structures), guided solely by tactile-memory (Likova, 2010a,b, 2012).

In contrast to standard imaging studies of tactile activation, the fMRI evaluation that was run before and after the training was specifically designed to probe the *memory* involvement by recording the brain activity while drawing-from-memory in the absence of any visual or tactile input from the learned raised-line drawing templates (see Materials and Methods). This novel Drawing-Based Memory Paradigm has the unique advantage of providing an *explicit* memory "readout" of the specific memory representation that guides it. Importantly, the Cognitive-Kinesthetic training allows subjects to learn to draw from memory the *specific memorized* objects and faces that they had explored, not just some long-standing "clichés," thus showing that the particular memory-representations generated during the *tactile exploration* phase were guiding their drawing.

To provide for comparative pre/post-training analyses of brain activation, an innovative platform was developed including the first MRI-compatible multisensory drawing tablet, with a stylus incorporating a fiber-optic motion-capture system to record the drawing movements in the scanner for off-line analyses.

These unique capabilities allow for testing between the following hypotheses:

Hypothesis I. If the memory-based drawing task is mediated by a transfer of the *tactually* felt spatial configuration to the *visual imagery* mechanism, the predicted response-profile in the visual cortex would have the top-down "imagery signature" of propagation of the imagery signal *downward* through the visual hierarchy, with activation significantly *decreasing* from the higher extrastriate areas toward V1.

Hypothesis II. If V1 was being employed as a WM sketchpad *independently* from the visual imagery process, it would be activated by a *separate pathway* external to the visual hierarchy, together with activation of WM-related sites beyond the occipital lobe. Moreover, if V1 activation is found in *non-sighted* drawing, it will also confirm a re-conceptualization of the sketchpad buffer from being visuo-spatial to being independent from sensory modality, or *amodal* as previously proposed *on the basis of a study in the congenitally blind* (Likova, 2012).

MATERIALS AND METHODS

EXPERIMENTAL DESIGN

A battery of raised-line models of faces and objects was developed as the drawing targets in a three-task block fMRI paradigm, with interleaved baseline conditions (**Figure 1A**). The three tasks were as follows: *Explore/Memorize* (*E/M*) – perceptual exploration and memorization of the model to be drawn; *MemoryDraw* (*MD*) – a memory-guided non-visual drawing task; and *Scribble* (*S*) – a motor-control and negative memory-control task. Each task duration was 20 s, with a 20 s baseline condition [*NullInterval* (*NI*)] intervening between the tasks. Importantly, as opposed to the usual null periods, the subjects not only rested motionless but were instructed and practiced to clear any memory or image structure from awareness ("mind-blank"). The start of each task or null interval was prompted by an auditory cue. The whole three-task sequence with interleaved null intervals (*NI*, *E/M*, *NI*, *MD*, *NI*, *S*, *NI*) was repeated 12 times in each 1-hour fMRI session using a new image for each repeat.

In *Explore/Memorize*, using the left hand only, the subjects had to tactually explore a raised-line drawing model on the left slot of the drawing tablet, and to develop a full memory representation of the image in preparation for the *MD* task. Then the model image was removed, and the subjects rested motionless for 20 s with no image in mind (*NI*). In the following *MD* phase the fiber-optic stylus was used to draw the image (from tactile-memory) on the right slot of the tablet using only the right hand; the trajectory of the stylus tip was recorded with high precision. *Scribble* was a control for both movement-specific activation (due to hand or eye movements) and absence of a memory involvement; the subjects had to move the stylus with the right hand in a random trajectory over the right slot of the tablet matching the extent and rate of the drawing movements, but under instructions to avoid planning or imagining any particular trajectory or cognitive content.

Rational of the drawing requirements

(i) The requirement that the models were always explored with the left hand but drawn by the right hand was an advanced aspect of the experimental design ensuring that in the *MD* task the cortex controlling the right (drawing) hand did not have any direct "haptic knowledge" of the image. This design enforces the development of a *detailed memory* representation in order to transfer the information later to the opposite (drawing) hand. (ii) Furthermore, the left hand was not allowed to follow the contour being drawn by the right hand, ensuring that the subjects learned to draw from memory *without* relying on specific tactile feedback of the raised-line configuration. Moreover, the reliance on detailed enough memory was further guaranteed by the fact that the virtual stylus left no tactually perceivable trace, eliminating any possibility for tactile tracing during drawing in the scanner. Together, these design features enforce the encoding of a robust *spatial memory* representation needed to guide the drawing trajectory. The quality of the reproductions was assessed by a masked rating procedure, based on recognition and similarity to the templates.

TACTILE STIMULUS PRESENTATION AND HAND MOVEMENT CONTROL

Multisensory MRI-compatible drawing system

To run drawing studies in the scanner is not a conventional protocol and faced a lot of challenging technological problems. We developed a multisensory drawing-system that: (1) is MRI-compatible, (2) is ergonomically adaptable to the small space available inside the scanner bore, (3) allows multiple tactile images to be presented sequentially in the scanner without the need

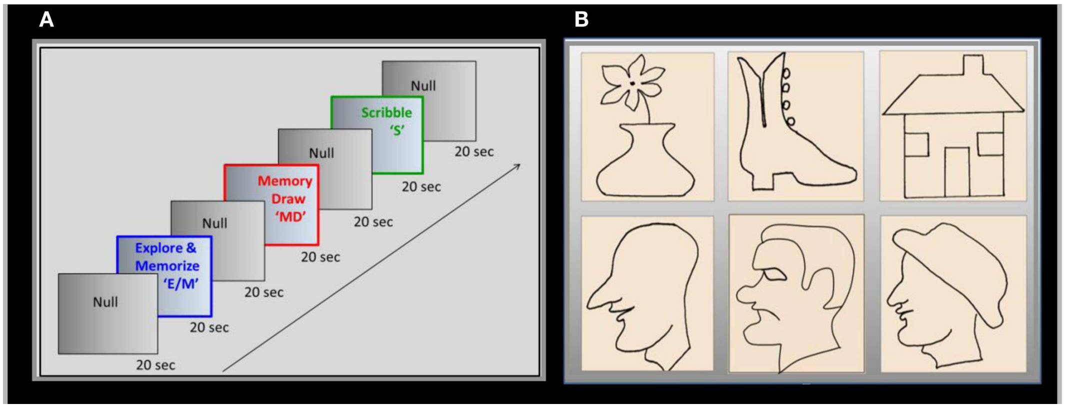


FIGURE 1 | Experimental design. (A) Drawing was investigated in a three-phase paradigm consisting of a memory-guided drawing task, abbreviated as "MemoryDraw" (*MD*), plus two control tasks: a motor-control and negative memory-control task "Scribble" (*S*), and a task of perceptual exploration and memorization of the model to be drawn "Explore/Memorize" (*E/M*). Each task's duration was 20 s, with 20 s null intervals interposed between the tasks, the whole 140 s trial sequence being repeated 12 times in each scanning session using a new image for each repeat. **(B)** Raised-line drawings of realistic faces and objects were presented as templates to be explored by the subject using her left hand. The quality of the reproductions was assessed by a masked rating procedure, based on recognition and similarity to the templates (examples of reproduction are shown in **Figure 3**).

of any operator assistance, (4) captures and records the drawing trajectory with high precision, and (5) provides a real-time visual feedback when appropriate (Likova, 2010a). It incorporates a dual-slot drawing tablet that is height/distance adjustable for the subject's arm length and a specially adapted version of a fiber-optic device for motion capture of the drawing movements with high-resolution (**Figure 2**). This unique drawing-system supports the fMRI investigation of both *tactilely* and *visually* guided drawing. It also allows us to record relevant behavioral and feedback events and to correlate them to the brain activation for full off-line analysis.

Auditory cue presentation

The auditory stimuli were presented through Resonance Technologies Serene Sound earphones (Resonance Technologies, Salem, MA, USA). To reduce scanner noise, this equipment employs external ear protectors with perforated earplugs that conduct the auditory cues directly into the auditory passage while blocking much of the scanner noise.

MRI DATA COLLECTION, ANALYSIS, AND VISUALIZATION

Subjects

The study was conducted on a group of six subjects with normal vision who were blindfolded during the experiments. The subjects ranged in age from 25 to 59 and were four females and two males. All subjects gave informed consent for the experimental protocol approved by the local research ethics committee, Institutional Review Board.

FMRI acquisition

MR data were collected on a Siemens Trio 3T scanner equipped with eight-channel EXCITE capability, a visual stimulus presentation system, response buttons. A high-resolution anatomical ($T1$-weighted) volume scan of the entire brain was obtained for each subject (voxel size = 0.8 mm × 0.8 mm × 0.8 mm). The fMRI blood-oxygenation-level-dependent (BOLD) responses were collected with EPI acquisition from the whole head coil. There were 34 axial slices at 2 s TR, with TE of 28 ms and flip angle of 80˚, providing 3.0 mm × 3.0 mm × 3.5 mm voxels throughout the brain. The functional activations were processed for slice-time correction and motion correction. An anatomical segmentation algorithm (mrGray) was applied to the T1 scan, ensuring localization of the signal within the cortical gray matter close to the activated neurons and greatly reducing the blood drain artifacts of BOLD signals displaced from the neural activation sites, which afflict studies in which cortical segmentation is not used.

FMRI time-course analyses

The analysis software was Stanford Vision and Imaging Science and Technology (VISTA) software. The data were analyzed to estimate the effective neural activation amplitudes (for each task across the 12 repeats of the three-task sequence in the 1-h scan) by the following procedure. A General Linear Model (GLM) consisting of a $(3 + 1)$-parameter boxcar neural activation model convolved with an estimated hemodynamic response function (HRF) was fitted to the BOLD responses (i.e., a 1-parameter boxcar for each of the three tasks, plus a 1-parameter 8-boxcar sequence to model the auditory cue presentations). An additive fourth-order polynomial was applied to capture low-frequency drift in the BOLD signal. (The HRF parameters were determined once per session by optimizing this model to a subset of gray matter voxels identified as most responsive to the task/null alternation frequency in this experiment.) Thus, the parameters of the activation model consisted of the boxcar activation amplitudes for the three-task periods, relative to the remainder of the 140 s scan duration.

Voxel-wise parametric maps

For each task statistical parametric maps were generated, based on the estimated activation amplitudes from the above GLM in each voxel. As is standard for GLM, the boxcar neural activation model for each 20 s task period was contrasted with the entire remainder of the 140 s scan duration. All three task-models were optimized

FIGURE 2 | A subject on the scanner bed operating our novel multimodal MRI-compatible drawing system. The plexiglass gantry supports a drawing tablet while a fiber-optic drawing stylus captures and records the drawing movements with high precision. The motion-capture information synchronized with the fMRI allows the effect of behavioral events to be analyzed to high precision.

jointly to the detrended BOLD waveform. Also, for the pre-post comparison, voxel-wise maps of the change in activation following the training were generated, scaled in terms of *z-score* of the pre-post difference signals. These maps could be viewed in the 3D volume or projected onto 3D views of the inflated cortex or flat-maps of cortical regions of particular interest.

ROI activation analysis

The effective neural activation amplitudes (bar-graphs) for each condition in each region of interest (ROI) were estimated by the same GLM procedure but now applied to the *average* signal across all voxels within the ROI. This procedure also provided high-quality time courses for evaluation of the response dynamics and its comparison across tasks and stages of training.

The confidence intervals in the bar-graphs were defined by the amplitude variability of the 12 repeats of the three-task sequence in each 1-h scan. The *dashed lines* and the *error bars* represent confidence intervals for *two different forms* of statistical comparison of the activation levels (i.e., of the beta weights for the event types in the GLM): (1) The *dashed lines* represent the 99% "zero" *confidence interval* ($p < 0.01$, uncorrected) within which the activation amplitudes are not significantly different from zero (i.e., relative to the noise variance for no stimulus-related activation defined as the residual variance after the GLM model fit of the *FMRI time-course analyses* section described above); thus this statistical criterion is designed to indicate the significance of each individual activation (at $p < 0.05$, corrected for multiple applications within each figure); (2) The *error bars* are "difference" *confidence intervals* designed to illustrate the *t-test* for the significance of differences *between* activation levels in each figure (i.e., the differences are not significant unless they exceed the confidence intervals for both compared activations). In the text, all ROI-comparisons are specified as significant by the *t-test* using a statistical criterion threshold of $p < 0.05$ corrected for multiple comparisons.

Topographic mapping

The boundaries of the retinotopic projection areas V1, V2d, V2v, V3d, and V3v were established as described in Sereno et al. (1995); Tootell et al. (1996); Engel et al. (1997). Retinotopic projection areas V3A, V3B, hV4, and V7 were specified in accordance with Tyler et al. (2005). The retinotopic mapping was done by using standard retinotopic stimuli – expanding rings and traveling wedges. An innovative 14-step procedure (Likova, 2010a) allowed us to warp the brains to the same MNI brain coordinates, within which other localizers could also be specified on the basis of prior studies.

RESULTS

The subjects were all able to improve their drawing skills so as to complete each drawing with its particular characteristics within the 20 s allotted for the drawing tasks (which took many times longer before training). Examples of drawing recorded in the post-training scanning session are shown in **Figure 3**, illustrating the level of detail required to complete each drawing.

BOLD RESPONSES TO DRAWING GUIDED BY TACTILE-MEMORY

The initial analysis gives an overview of the averaged cortical activation for the *MD* task in the brains of the group of blindfolded subjects. **Figure 4** shows the post-training map, projected on inflated representations of the lateral and medial views of the two hemispheres. As expected, there is strong activation in the pre-motor, motor, and somatosensory cortex in the dorsal regions of the brain (the pre-central and post-central sulci, PreCS and PostCS) predominantly for the left hemisphere controlling the right hand that was performing the drawing task (**Figures 4A,B**). The anterior and posterior dorsal regions also showed enhanced activation bilaterally, implying enhanced kinesthetic processing for the drawing movements. The supplementary motor area (SMA, on the medial surface), which plays a role in the planning of complex coordinated movements is also activated bilaterally, although again the left hemisphere responds more strongly. The dorsolateral-prefrontal cortex, known to be of key importance in WM, decision-making, executive control, etc. is activated in both hemispheres. Temporal lobe activation can be also seen in the LOtv region suggested to be involved in tactile object processing. The involvement of many of these regions would be very much predicted on the basis of prior studies. There is also an extensive network of deactivation

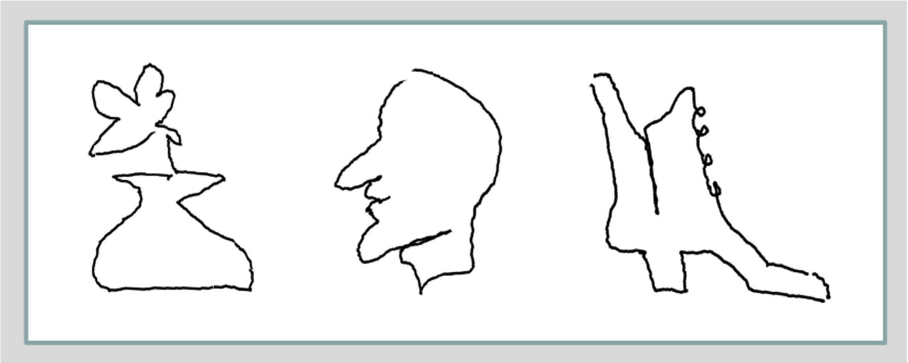


FIGURE 3 | Examples of blindfolded drawings of the vase with a flower, the face profile, and the boot (the corresponding templates shown in Figure 1B are easy to recognize: the first and the second in the top row, and the first in the bottom row). Remarkably, the post-training drawings, recorded in the scanner by the motion-capture system show a lot of specific detail, which makes them readily recognizable as specific examples of their category, although they were drawn without visual or tactile input (i.e., with eye-hand coordination eliminated), but were guided solely by tactile-memory.

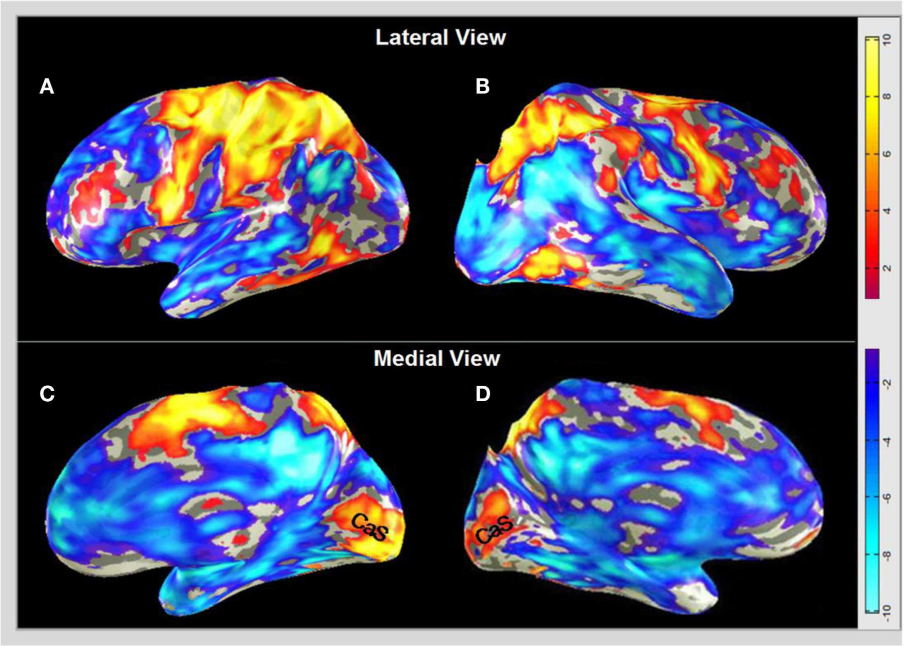


FIGURE 4 | BOLD activation and deactivation in *non-visual memory drawing* in the blindfolded. Post-training group responses from the *MD* task are derived according to the GLM described in Section "Materials and Methods," and projected on inflated representations of the lateral left **(A)** and right **(B)**, and medial left **(C)** and right **(D)** hemispheres in MNI brain coordinates. Dark-gray, sulci; light gray, gyri. **(A,B)** A non-occipital network of temporal, parietal, and frontal regions is activated (yellow-orange coloration), together with strong deactivation (blue-cyan coloration) in a network that corresponds broadly to the default-mode network (except for the occipital lobe portion). **(C,D)** Both medial views show *massive activation* along the *calcarine sulcus* (CaS) corresponding to V1 surrounded by deactivation, which extends throughout the lateral regions of the visual cortex. Activation is shown down to $-1 < z < 1$; the scale bar indicates the color-coding for the respective *z-score* levels. Note that, interestingly, the medial CaS activation spreads to the same eccentricity in both hemispheres.

that, *beyond* the occipital lobe regions, largely overlaps with the default-mode network (e.g., Raichle and Snyder, 2007).

What is surprising within the traditional view of brain architecture, however, is the massive activation in the occipital region along the calcarine sulcus (CaS) for this non-visually and not even sensory (neither visually nor tactually) guided task (in the sense that, as described in Methods, the task has been accomplished with no visual input at any stage, and the drawing phase has been accomplished with guidance only by tactile-memory, i.e., with no concurrent sensory input of any form about the image to be drawn). This region corresponds to the location of area V1 and is the key focus of the present analysis. The strong V1 activation (orange-yellow coloration) can be seen on the medial surface in both hemispheres (**Figures 4C,D**).

Activation is shown down to $-1 < z < 1$ because the lower amplitudes appear to form a consistent fringe around the cluster regions of high significance. Indeed, such a visualization approach is not often taken; however, in cases where not only the positive but the negative clusters are also of interest, this visualization approach allows the following to be shown: (i) whether the surrounding regions similarly show activation but just below the threshold level, or (ii) whether there is an extended cluster of deactivation. Higher-threshold presentations mask such differences; thus, it seems worthwhile to provide maximum information in the figures. For restricting to the individual voxel (rather than cluster) significances, the color-coding provides explicit information as to the activation pattern for any preferred threshold level.

To better evaluate the *activation/deactivation* pattern in the context of the functional "geography" of the visual cortex, **Figure 5A** shows a flat-map representation of the occipital cortex of a representative blindfolded subject who had all of the visual-hierarchy areas mapped, as well as the motion complex hMT+ and LOC. The flat-maps are centered on the occipital pole and oriented as if viewed from the back of the head, with activation/deactivation designated as in **Figure 4**. The retinotopic boundaries were determined in a separate scan using a 20° circular field. Note that the peripheral boundaries of these regions thus correspond to about 10° of eccentricity, which corresponds approximately to the maximum extent of the BOLD activation (yellow-orange coloration). Area V1 is outlined by a red contour. It is remarkable to see that the *tactile-memory* drawing not only generates activation specific to V1 (although no sensory visual or even tactile information about the drawing templates was available), while the entire extrastriate hierarchy that surrounds V1 is massively *deactivated.* The extended non-visual response in what traditionally is considered as primary "*visual*" cortex was accompanied by activation in a region of cortex at the occipitotemporal border known to be involved in tactile object recognition, LOtv (e.g., Amedi et al., 2001, 2002; Reed et al., 2004); activation is also seen dorsally in the caudal intraparietal sulcus (cIPS).

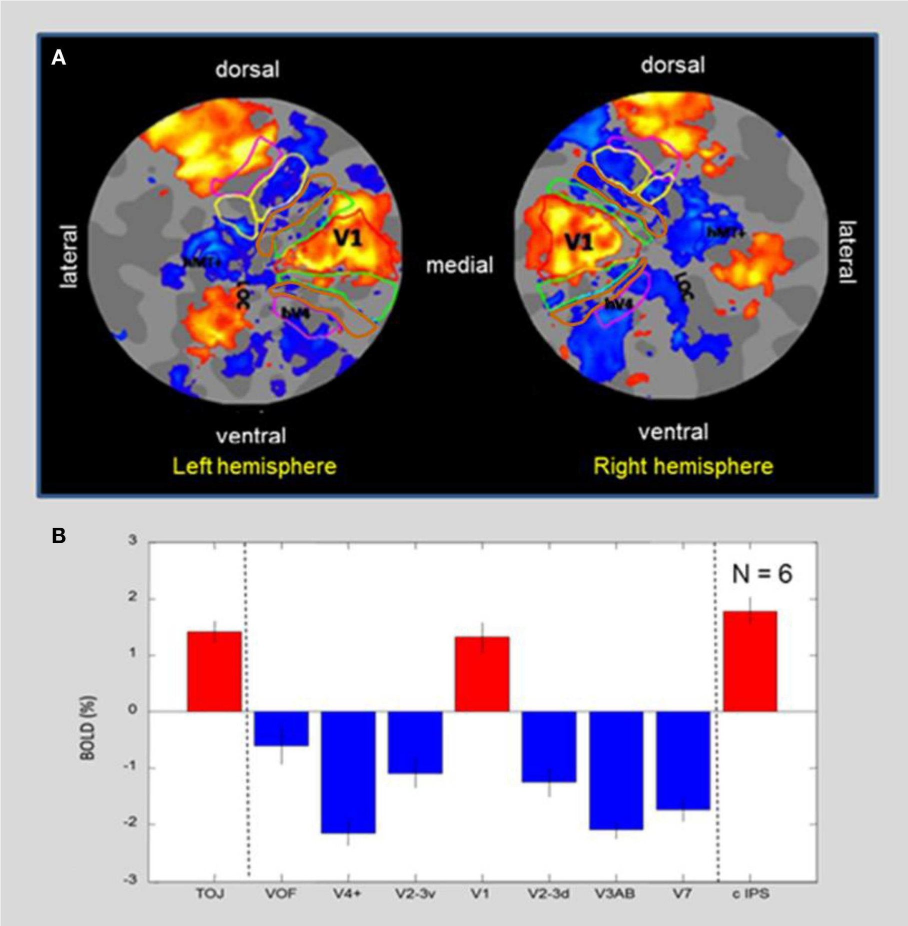


FIGURE 5 | (A) *MD* flat-maps centered on the occipital pole. ROIs for the retinotopic hierarchy are indicated by colored outlines, with hMT+ and LOC based on functional localizers. The post-training *MD* map shows a "triad" of three activation regions (orange-yellow coloration). Note in particular that the (non-stimulated visually) primary visual cortex, V1, forms an unusual isolated "island" of *activation* surrounded by a "sea" of *suppression* in the adjacent retinotopic areas. The other two activated regions seen on the flat map are the caudal intraparietal sulcus (cIPS) dorsally and an additional locus at the occipitotemporal border (LOtv). **(B)** Average response amplitude with standard errors for blindfolded *memory-guided drawing* in a group of six subjects, showing positive signal in the triad of areas – primary visual area V1, cIPS, and LOtv; these three "islands" of positive activation are separated by strong deactivation in both the ventral and the dorsal extrastriate areas. Error bars represent 1 standard error of the means.

Average amplitudes across the ROIs shown in **Figure 5A** are quantified for a group of six normally sighted subjects performing the drawing task under blindfolded conditions in **Figure 5B**. As in the example subject, the only three regions in this part of the cortex showing significant activation are V1, and the two regions beyond the visual retinotopic hierarchy, LOtv, and cIPS. All the other regions show significant reductions in the BOLD below baseline.

A difference *MD* map, which represents voxel-wise comparison of the *post-training* BOLD activation relative to the *pre-training* level is presented in the CaS region in **Figure 6**. It shows a pronounced bilateral increase of activation after the Cognitive-Kinesthetic training.

CROSS-TASK COMPARISON OF V1 ACTIVATION

Figure 7A below shows the average time-courses of BOLD activity (black lines) in V1 for the sequence of the three-task intervals (white bars); the four dark-gray bars indicate the 20 s null intervals separating *E/M*, *MD*, and *S* tasks. The bar-graphs in **Figure 7B** show the estimated activation in each hemisphere for each task. Importantly, the *MD* task (red color), which requires retrieval of *detailed tactile-memory* representation is the one that produces the strongest activation in V1; note that eliminating the memory component (in the "non-memory" drawing task *S*, green color) correspondingly eliminates any response in this purely motor form of the task. Although the *E/M* task (blue color), which represents the memory encoding phase, also seems to employ V1 to some degree, significantly less activation is observed than for the *MD* task. Correspondingly, **Figures 7C,D** show the average time-courses and bar-graphs for the deactivated extrastriate regions. The *MD* response for these regions is the most prominent and highly significantly negative. *E/M* and *S*, on the other hand, have marginally significant responses, but are not significantly different from each other. All cross-task ROI-comparisons in the text

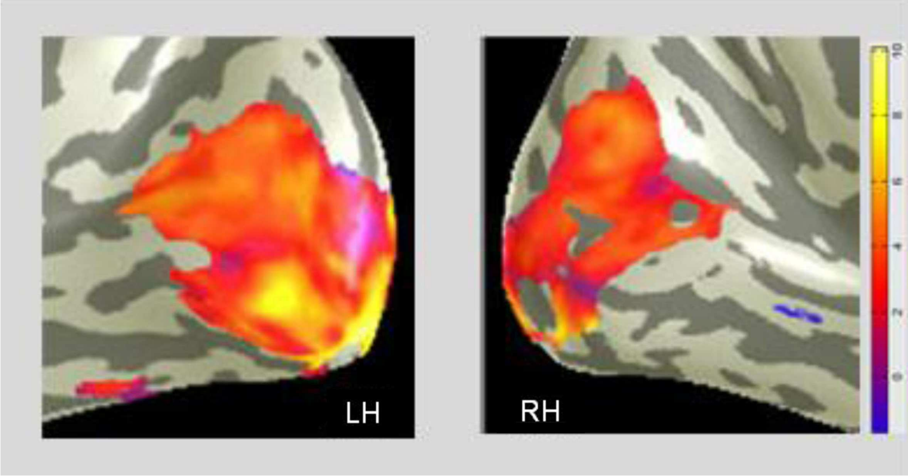


FIGURE 6 | The *MD* task shows the predominant training effect in the primary visual cortex. A voxel-wise comparison, projected on inflated representations of the posterior left (LH) and right (RH) hemispheres, with orange-yellow coloration showing the average *pre/post increase* in BOLD activation for *MD* in the CaS (V1) region.

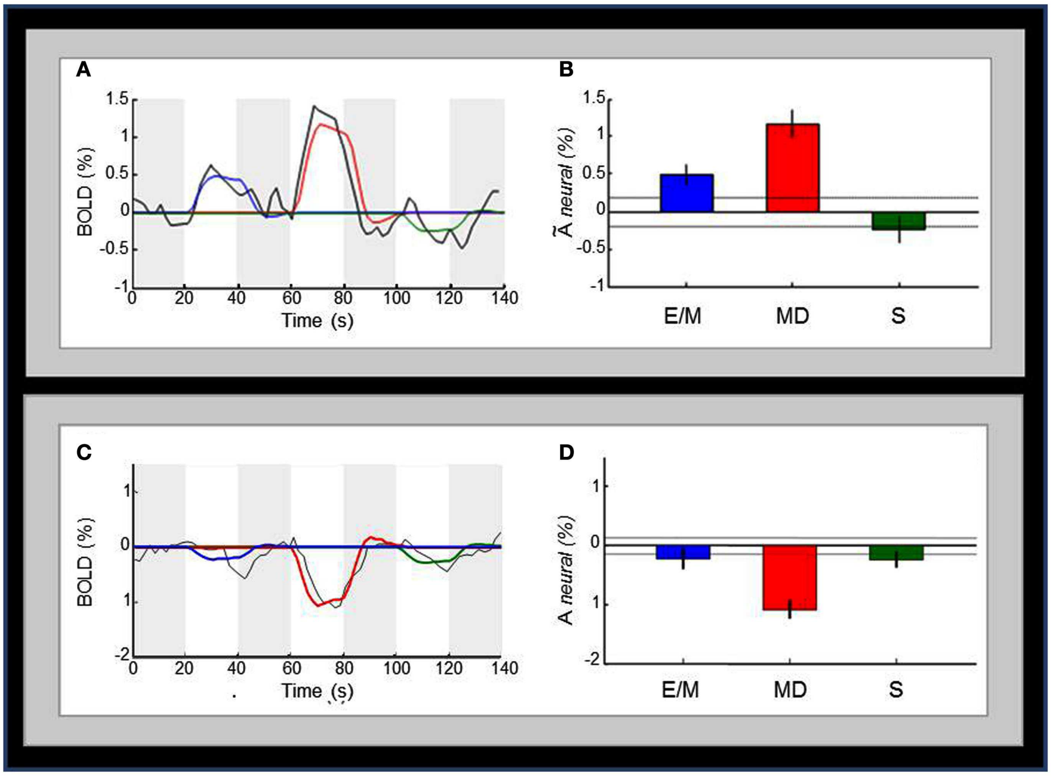


FIGURE 7 | Comparison of the activation pattern across the three tasks. (A) Average time-courses of BOLD activity (black curve) in V1 for the sequence of the *E/M*, *MD*, and *S* task intervals (white bars); the colored curves are the best fits of the model predictions for the three tasks to the time course; the four dark-gray bars indicate the 20 s null intervals separating the three tasks. **(B)** Bar-graphs for the estimated V1 activation for each task; the activation levels refer to the beta weights for the event types in the GLM. Dotted lines and the error bars represent confidence intervals for *two different forms* of statistical comparison of the activation levels. *Dotted lines* represent the 99% "zero" *confidence interval*, within which the activations are not significantly different from zero. *Error bars* are 99% "difference" *confidence intervals* designed to illustrate the *t-test* to assess the significance of the differences *between pairs of* activation levels in each figure, i.e., amplitude differences are not significant unless they exceed the confidence intervals for both compared activations. **(C)** Average time-courses of BOLD activity in the deactivated regions surrounding V1. **(D)** Bar-graphs of estimated activation for each task for the deactivated regions presented in **(C)**. Conventions in **(C,D)** are as in **(A,B)**, respectively.

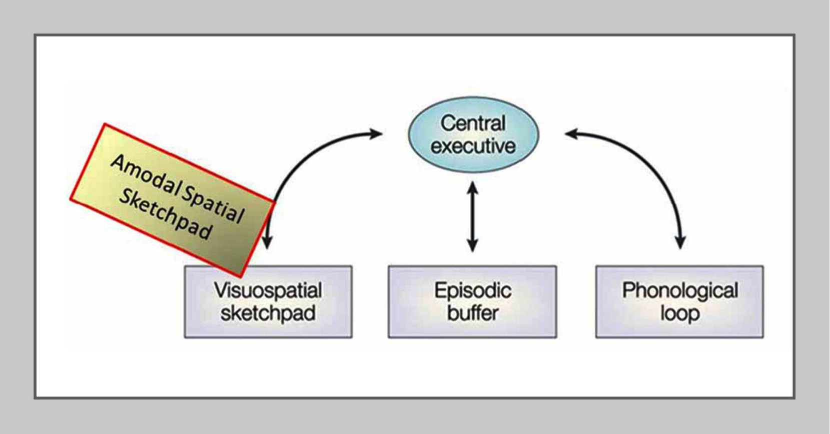


FIGURE 8 | The proposed re-conceptualization of the *visuo-spatial* sketchpad as an *amodal-spatial* sketchpad. Modified schematic of the main modules of Baddeley's classic model of working memory including the visuo-spatial sketchpad (after Baddeley, 2003), where the added "Amodal-Spatial Sketchpad" block depicts our re-conceptualization of the *visuo-spatial* sketchpad as being accessible to any sensory modality (from Likova, 2012).

are specified as significant by the *t-test* at a statistical criterion threshold of $p < 0.05$, corrected for multiple comparisons.

Although the time-course plots show that there was no strong activation during the "mind-blank" null periods between the task epochs, to formally assess the degree of memory involvement in V1 following the *E/M* phase, we ran regressors for all null periods, and compared the one after *E/M* with the average for the remainder of the null periods. As expected based on the "mind-blank" design of these periods, there were no significant differences in the V1 ROI ($p \gg 0.1$), implying that there was no evidence for memory engagement of V1 during this null period. The corresponding analysis was also run on the deactivated extrastriate ROI with the same result.

DISCUSSION

Employing a novel memory paradigm based on drawing in normally sighted subjects while blindfolded revealed that the primary "visual" area V1 can be strongly activated in a *non-visual WM* task: the task of blindfolded drawing guided solely by memory acquired during haptic exploration of complex spatial structures, such as raised-line objects and faces. Furthermore, the pattern of response showed a number of unique characteristics:

(i) The occipital activation was largely restricted to area V1.
(ii) The strong V1 activation was remarkably well structured, ceasing rapidly at a specific eccentricity.
(iii) Surprisingly, the V1 activation was surrounded by massive *deactivation* of the entire extrastriate visual hierarchy.

In contrast to the parallel nature of typical visual processing, haptic exploration operates in a sequential manner. As the blindfolded are subjects who have an intact visual system, and presumably a well developed visual imagery mechanism, one possibility it is that once the spatiotemporal integration of the sequentially explored template-objects was completed, the memory retrieval was implemented by the visual imagery mechanism (*Hypothesis I*). That is, in principle, it is possible that the memory of the complex spatial structures could in some way be transferred to higher-order visual imagery "processors," which in turn may have mediated the corresponding conscious experience via a top-down propagation through the visual hierarchy to a high-resolution memory representation in V1.

However, this hypothesis seems not to be supported by the data because, although we find strong activation in the iconic visual area V1, it is implausible that the signal is "delivered" through the visual hierarchy, as this hierarchy is not only not activated but is massively *deactivated*. This implies that V1 is "cut-off" from the higher-level cortical regions that could generate and propagate the imagery signals. It is important to emphasize that this *pattern of V1-activation/extrastriate-deactivation* is quite distinct, almost the *inverse* of the known hierarchical pattern for visual imagery (which is a top-down process strongest in the higher extrastriate areas and *decreasing* toward the lower areas, often not reaching V1 at all; see Introduction). The implications of the deactivation cutting-off V1 from receiving signal through the extrastriate visual pathway go further, beyond simple judgment of "similarity" or "difference" of activation patterns, to imply functional incompatibility with the main principle of visual imagery as a process propagating through the visual pathway downward to V1.

Consequently, the unique pattern of response in the blindfolded is not compatible with an explicit role for visual imagery in this form of WM. Instead, the strong but isolated V1 activation seems to be more consistent with *Hypothesis II* that V1 is operating as a WM component, such as the spatial memory-buffer/sketchpad of the composite WM model.

The *training* paradigm of the current study also provides a *causal* manipulation that links the memory enhancement to the *increase* in V1 activation as a result of the (non-visual) Cognitive-Kinesthetic training, consistent with the memory-buffer interpretation of the role of V1. Moreover an important twist for this interpretation is the lack of any visual stimulation under the blindfolded conditions, implying that activation of the V1 buffer should be independent of the input modality. In other words, the present results imply that the nature of the V1-buffer is not "visuo-spatial" but "*amodal*-spatial".

Importantly, the results in the blindfolded reinforce the previous implication from a parallel study of memory drawing in a congenitally blind subject (Likova, 2012) that V1 was operating as an *amodal* memory-buffer because the subject had had a complete absence of visual experience and visual memory throughout life and had performed the task entirely based on the memory from the tactile input. The strong post-training activation of V1 in that study thus could not meaningfully be attributed to visual imagery, but is more consistent with the hypothesis that V1 uses an amodal-spatial representation in its operation as the putative memory-buffer. Our re-conceptualization of the *visuo-spatial* sketchpad as being *amodal-spatial* is depicted by the yellowish block in **Figure 8**.

It is important to note, however, that such an interpretation does not mean that V1 is employed for *storage* of the memory trace. Moreover, in contrast to the usual format of a *baseline* condition, we instructed and practiced the subjects to *eliminate any rehearsal* of either the just-explored templates or of any other memory images for the full 20 s duration of each *null* interval. This duration is also too long to account for the known retention time of any iconic image of a memory trace, that is of the order of a second or less (Sperling, 1963; Di Lollo, 1980). Since the drawings were not experienced as spatial images during the *null* interval, they were evidently held in some other, non-conscious storage location until it was needed for the subsequent drawing task.

Further support for the general idea of early sensory areas in human being involved in some form in WM comes from the seminal study of Harrison and Tong (2009), which "demonstrated that early visual areas can *retain* specific information about visual features held in WM." However, similarly to visual imagery (and *in contrast* to our data), the *whole sequence* of early visual areas – V1, V2, V3, V3A–V4 was activated in that study. Thus, as should be expected for a visual process, *visual* WM did *not* suppress but activated the extrastriate areas, i.e., it did *not* "cut-off" the V1 signal propagation through the visual pathways. Their result of "*widespread* activity throughout the early visual system" makes it clear that *visually* driven WM uses pathways different from those in our *non-visually* driven WM task, which had *deactivated* all of the extrastriate visual areas.

Thus, visual WM can not account for our data, while on the other hand, our concept of an amodal (modality-independent) spatial-sketchpad in V1 is consistent with the V1 involvement in both visual and non-visual WM, as its amodal nature implies availability through the visual as through the tactile modality. Therefore, the present results encourage an *expanded* view of the V1 functionality to a *cross-modal involvement in WM processing* (the access to which, however, would require different pathways for different modalities).

GENERALIZATION BEYOND THE SKETCHPAD: V1 AS A GENERIC "PROJECTION SCREEN" OF AN AMODAL NATURE?

Although the present results are consistent with our Hypothesis II, namely that V1 was operating as the neural substrate for the putative WM sketchpad in the MD task, and not as a visual imagery component, they have further significance beyond this specific dichotomy.

In a more general sense, the current results provide a strong demonstration of massive employment of V1 in a higher-order cognitive task that involves *no* visual (or even tactile) *sensory* stimulation. Thus, in principle, these results do not exclude a *broader hypothesis*, specifically that V1 may play the role of a more *generic* "*projection screen*" of an *amodal* nature, which could be utilized by each of the two main cognitive constructs discussed here, as well as by visual WM and by other forms of cognitive functions requiring such a high-resolution "projection screen." Depending on the specific *task needs*, it may be utilized in either a cross-modal or intramodal manner.

POSSIBLE MECHANISMS

The general field of visuotactile interactions, especially with respect to primary visual cortex in non-visual sensory stimulation, began from studies in the blind (e.g., Sadato et al., 1996; Cohen et al., 1997; Zangaladze et al., 1999; Pascual-Leone and Hamilton, 2001; Block, 2003). Although further analyses are needed to investigate what underlying mechanisms may mediate the cross-modal activation observed in V1 in the current tactile-memory task, there is a range of general theoretical possibilities, such as unmasking of pre-existing inter-region connections, changes in synaptic weights, modulation of long-range intercortical influences, up-regulation of non-local transmitter sources, or a variety of subcortical mechanisms (e.g., Florence and Kaas, 1995; Jones, 2000; Pascual-Leone and Hamilton, 2001; Raineteau and Schwab, 2001; Block, 2003; Merabet et al., 2008; Van Brussel et al., 2011).

Indeed, the extent of V1 connectivity is currently undergoing an extensive re-evaluation. Recent electrophysiological and anatomical studies in non-human primates reveal a picture of multiple reciprocal connections at lower hierarchical levels, including the primary areas. In addition to the well-known direct feedback projections to V1 originating from the visual hierarchy (Perkel et al., 1986; Ungerleider and Desimone, 1986a,b; Shipp and Zeki, 1989; Rockland, 1994; Budd, 1998; Barone et al., 2000; Suzuki et al., 2000), there are direct feedforward projections to V1 originating from a variety of subcortical structures, including the pulvinar, LGNd, claustrum, nucleus paracentral, raphe system, locus coeruleus, and the nucleus basalis of Meynert (Ogren and Hendrickson, 1976; Rezak and Benevento, 1979; Graham, 1982; Blasdel and Lund, 1983; Doty, 1983; Perkel et al., 1986; Lachica and Casagrande, 1992; Hendry and Yoshioka, 1994; Adams et al., 2000; Schmolesky, 2000). In principle, any of these subcortical structures could be involved in the processes of memory storage and retrieval for the performance of the high-resolution drawing task.

Additionally, Clavagnier et al. (2004), examined feedback projections to area V1 using retrograde tracer injections. Notably, in addition to well-known areas and a number of long-distance feedback connections originating from auditory (A1) and multisensory (STP) cortices, they also found connections from a perirhinal area. The perirhinal-to-V1 connections appear of particular interest (Likova, 2012) in the context of our finding of a memory-related role for V1, as the perirhinal area is adjacent to the hippocampus and has a well-established role in memory storage and retrieval.

These connections therefore could represent another potential pathway for the involvement of V1 in WM and the active processing of stored spatial information. Nevertheless, none of the multiple connections above directly predicts the way that V1 is *cut-off* from the surrounding visual hierarchy by the deactivation. Definitive studies on these issues remain to be conducted.

CONCLUSION

This novel experimental approach, showing how a WM task accesses the highest resolution topographic map in the brain (V1) even under non-visual conditions, provides a "real-life" yet tractable paradigm for addressing the role of such high-order cognitive processes in a cross-modal manner. V1 was activated in an isolated fashion in the drawing-from-memory task, supporting our Hypothesis II that V1 operates as the active spatial "sketchpad" underlying the accurate drawing performance under non-visual conditions. The blindfolded drawing results were also consistent with the previous conclusion from a congenitally blind study (Likova, 2012) that the spatial WM sketchpad may operate in an *amodal* (rather than exclusively visual) fashion. The converse hypothesis that the memory retrieval would activate the visual hierarchy in a top-down fashion (as would be expected if this task were mediated through the imagery network), was not supported by the data. In a more general sense, the results are a strong demonstration of a massive *cross-modal activation in V1* in a high-level cognitive function. In combination with our previous work with the Drawing-Based Memory Paradigm and training effects of the Cognitive-Kinesthetic protocol, these studies further propel the emerging re-conceptualization of brain architecture as highly interactive and capable of plastic reorganization even after short-term sensory deprivation.

ACKNOWLEDGMENTS

This research is supported by NSF/SLC Grant to Lora T. Likova. The author thanks Spero Nicholas for his help in data preprocessing and analysis tools, and Christopher W. Tyler for fruitful discussions.

REFERENCES

Adams, M. M., Hof, P. R., Gattass, R., Webster, M. J., and Ungerleider, L. G. (2000). Visual cortical projections and chemoarchitecture of macaque monkey pulvinar. *J. Comp. Neurol.* 419, 377–393.

Amedi, A., Jacobson, G., Hendler, T., Malach, R., and Zohary, E. (2002). Convergence of visual and tactile shape processing in the human lateral occipital complex. *Cereb. Cortex* 12, 1202–1212.

Amedi, A., Malach, R., Hendler, T., Peled, S., Zohary, E. (2001). Visuo-haptic object-related activation in the ventral visual pathway. *Nat. Neurosci.* 4, 324–330.

Baddeley, A. D. (1986). *Working Memory*. New York, NY: Oxford University Press.

Baddeley, A. D. (1992). Working memory. *Science* 255, 556–559.

Baddeley, A. D. (2000). The episodic buffer: a new component of working memory? *Trends Cogn. Sci. (Regul. Ed.)* 4, 417–423.

Baddeley, A. D. (2003). Working memory: looking back and looking forward. *Nat. Rev. Neurosci.* 4, 829–839.

Baddeley, A. D., and Andrade, J. (2000). Working memory and the vividness of imagery. *J. Exp. Psychol. Gen.* 129, 126–145.

Baddeley, A. D., and Hitch, G. J. (1974). "Working memory," in *The Psychology of Learning and Motivation*, Vol. 8, ed. G. Bower (San Diego, CA: AcademicPress), 47–90.

Barone, P., Batardiere, A., Knoblauch, K., and Kennedy, H. (2000). Laminar distribution of neurons in extrastriate areas projecting to visual areas V1 and V4 correlates with the hierarchical rank and indicates the operation of a distance rule. *J. Neurosci.* 20, 3263–3281.

Blasdel, G. G., and Lund, J. S. (1983). Termination of afferent axons in macaque striate cortex. *J. Neurosci.* 3, 1389–1413.

Block, N. (2003). Tactile sensation via spatial perception. *Trends Cogn. Sci. (Regul. Ed.)* 7, 285–286.

Bruyer, R., and Scailquin, J. C. (1998). The visuospatial sketchpad for mental images: Testing the multicomponent model of working memory. *Acta Psychol. (Amst.)* 98, 17–36.

Budd, M. L. (1998). Extrastriate feedback to primary visual cortex in primates: a quantitative analysis of connectivity. *Proc. Biol. Sci.* 265, 1037–1044.

Charlot, V., Tzourio, N., Zilbovicius, M., Mazoyer, B., and Denis, M. (1992). Different mental imagery abilities result in different regional cerebral blood flow activation patterns during cognitive tests. *Neuropsychologia* 30, 565–580.

Chen, W., Kato, T., Zhu, X.-H., Ogawa, S., Tank, D. W., and Ugurbil, K. (1998). Human primary visual cortex and lateral geniculate nucleus activation during visual imagery. *Neuroreport* 9, 3669–3674.

Clavagnier, S., Falchier, A., and Kennedy, H. (2004). Long-distance feedback projections to area V1: implications for multisensory integration, spatial awareness, and visual consciousness. *Cogn. Affect. Behav. Neurosci.* 4, 117–126.

Cohen, L. G., Celnik, P., Pascual-Leone, A., Corwell, B., Falz, L., Dambrosia, J., et al. (1997). Functional relevance of cross-modal plasticity in blind humans. *Nature* 389, 180–183.

D'Esposito, M., Detre, J. A., Aguirre, G. K., Stallcup, M., Alsop, D. C., Tippet, L. J., et al. (1997). A functional MRI study of mental image generation. *Neuropsychologia* 35, 725–730.

Di Lollo, V. (1980). Temporal integration in visual memory. *J. Exp. Psychol. Gen.* 109, 75–97.

Doty, R. W. (1983). Nongeniculate afferents to striate cortex in macaques. *J. Comp. Neurol.* 218, 159–173.

Engel, S. A., Glover, G. H., and Wandell, B. A. (1997). Retinotopic organization in human visual cortex and the spatial precision of functional MRI. *Cereb. Cortex* 7, 181–192.

Florence, S. L., and Kaas, J. H. (1995). Large-scale reorganization at multiple levels of the somatosensory pathway follows therapeutic amputation of the hand in monkeys. *J. Neurosci.* 15, 8083–8095.

Formisano, E., Linden, D. J., Di Salle, F., Trojano, L., Esposito, F., Sack, A. T., et al. (2002). Tracking the mind's image in the brain: I. Timeresolved fMRI during visuospatial mental imagery. *Neuron* 35, 185–194.

Ganis, G., Thompson, W. L., and Kosslyn, S. M. (2004). Brain areas underlying visual mental imagery and visual perception: An fMRI study. *Cogn. Brain Res.* 20, 226–241.

Goldenberg, G., Podreka, I., Steiner, M., Franzen, P., and Deecke, L. (1991). Contributions of occipital and temporal brain regions to visual and acoustic imagery-A SPECT study. *Neuropsychologia* 29, 695–702.

Graham, J. (1982). Some topographical connections of the striate cortex with subcortical structures in Macaca fascicularis. *Exp. Brain Res.* 47, 1–14.

Harrison, S. A., and Tong, F. (2009). Decoding reveals the contents of visual working memory in early visual areas. *Nature* 458, 632–635.

Hendry, S. H., and Yoshioka, T. (1994). A neurochemically distinct third channel in the macaque dorsal lateral geniculate nucleus. *Science* 264, 575–577.

Ishai, A., Haxby, J. V., and Ungerleider, L. G. (2002). Visual imagery of famous faces: Effects of memory and attention revealed by fMRI. *Neuroimage* 17, 1729–1741.

Ishai, A., and Sagi, D. (1995). Common mechanisms of visual imagery and perception. *Science* 268, 1771–1974.

Ishai, A., Ungerleider, L., and Haxby, J. V. (2000). Distributed neural systems for the generation of visual images. *Neuron* 28, 979–990.

Jones, E. G. (2000). Plasticity and neuroplasticity. *J. Hist. Neurosci.* 9, 37–39.

Kaas, A., Weigelt, S., Roebroeck, A., Kohler, A., and Muckli, L. (2010). Imagery of a moving object: the role of occipital cortex and human MT/V5+. *Neuroimage* 49, 794–804.

Kennedy, J. (1993). *Drawing and the Blind*. New Haven, CT: Yale University Press.

Kennedy, J. M. (2000). "Recognizing outline pictures via touch: Alignment theory," in *Touch, Representation and Blindness*, ed. M. A. Heller (Oxford: Oxford University Press), 67–98.

Kennedy, J. M., and Igor, J. (2003). Haptics and projection: drawings by Tracy, a blind adult. *Perception* 32, 1059–1071.

Kennedy, J. M., and Juricevic, I. (2006). Foreshortening, convergence and drawings from a blind adult. *Perception* 35, 847–851.

Keogh, R., and Pearson, J. (2011). Mental imagery and visual working memory. *PLoS ONE* 6:e29221. doi:10.1371/journal.pone.0029221

Knauff, M., Kassubek, J., Mulack, T., and Greenlee, M. W. (2000). Cortical activation evoked by visual mental imagery as measured by fMRI. *Neuroreport* 11, 3957–3962.

Kosslyn, S. M., Alpert, N. M., Thompson, W. L., Maljkovic, V., Weise, S. B., Chabris, C. F., et al. (1993). Visual mental-imagery activates topographically organized visual cortex – pet investigations. *J. Cogn. Neurosci.* 5, 263–287.

Kosslyn, S. M., Ganis, G., and Thompson, W. L. (2001). Neural foundations of imagery. *Nat. Rev. Neurosci.* 2, 635–642.

Kosslyn, S. M., Pascual-Leone, A., Felician, O., Camposano, S., Keenan, J. P., Thompson, W. L., et al. (1999). The role of area 17 in visual imagery: convergent evidence from PET and rTMS. *Science* 284, 167–170.

Kosslyn, S. M., Shin, L. M., Thompson, W. L., McNally, R. J., Rauch, S. L., Pitman, R. K., et al. (1996). Neural effects of visualizing and perceiving aversive stimuli: A PET investigation. *Neuroreport* 7, 1569–1576.

Kosslyn, S. M., and Thompson, W. L. (2003). When is early visual cortex activated during visual mental imagery? *Psychol. Bull.* 129, 723–746.

Kreiman, G., Koch, C., and Fried, I. (2000). Imagery neurons in the human brain. *Nature* 408, 357–361.

Lachica, E. A., and Casagrande, V. A. (1992). Direct W-like geniculate projections to the cytochrome-oxidase (CO) blobs in primate visual cortex: axon morphology. *J. Comp. Neurol.* 319, 141–158.

Lambert, S., Sampaio, E., Scheiber, C., and Mauss, Y. (2002). Neural substrates of animal mental imagery: Calcarine sulcus and dorsal pathway involvement – an fMRI study. *Brain Res.* 924, 176–183.

Le Bihan, D., Turner, R., Zeffiro, T. A., Cue´nod, C. A., Jezzard, P., and Bonnerot, V. (1993). Activation of human primary visual cortex during visual recall: a magnetic resonance imaging study. *Proc. Natl. Acad. Sci. U.S.A.* 90, 11802–11805.

Lee, T. S., and Mumford, D. (2003). Hierarchical Bayesian inference in the visual cortex. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* 20, 1434–1448.

Lee, T. S., Mumford, D., Romero, R., and Lamme, V. A. F. (1998). The role of the primary visual cortex in higher level vision. *Vision Res.* 38, 2429–2454.

Likova, L. T. (2010a). "Drawing in the blind and the sighted as a probe of cortical reorganization," in *Proceedings of the SPIE7527 Human Vision and Electronic Imaging*, eds B. E. Rogowitz and T. N. Pappas, 752708–752714.

Likova, L. T. (2010b). The primary visual cortex as a modality-independent 'screen' for working memory. *J. Vision* 10:776.

Likova, L. T. (2012). Drawing enhances cross-modal memory plasticity in the human brain: A case study in a totally blind adult. *Front. Hum. Neurosci.* 6:44. doi:10.3389/fnhum.2012.00044.

Logie, R. (1995). *Visio-spatial Working Memory*. Hove: Lawrence Erlbaum Associates Ltd.

Logie, R. H., Baddeley, A. D., Mane, A., Donchin, E., and Sheptak, R. (1989). Working memory and the analysis of a complex skill by secondary task methodology. *Acta Psychol.* 71, 53–87.

Logie, R. H., and Marchetti, C. (1991). "Visio-soatial working memory: visual, spatial or central executive?" in *Mental Images in Human Cognition*, eds R. H. Logie and M. Denis (Amsterdam: Elsevier), 105–115.

Mazard, A., Tzourio-Mazoyer, N., Crivello, F., Mazoyer, B., and Mellet, E. (2004). A PET meta-analysis of object and spatial mental imagery. *Eur. J. Cogn. Psychol.* 16, 673–695.

Mechelli, A., Price, C. J., Friston, K. J., and Ishai, A. (2004). Where bottom-up meets top-down: neuronal interactions during perception and imagery. Cereb. *Cortex* 14, 1256–1265.

Mellet, E., Petit, L., Mazoyer, B., Denis, M., and Tzourio, N. (1998a). Reopening the mental imagery debate: Lessons from functional anatomy. *Neuroimage* 8, 129–139.

Mellet, E., Tzourio, N., Denis, M., and Mazoyer, B. (1998b). Cortical anatomy of mental imagery of concrete nouns based on their dictionary definition. *Neuroreport* 9, 803–808.

Mellet, E., Tzourio, N., Denis, M., and Mazoyer, B. (1995). A positron emission tomography study of visual and mental spatial exploration. *J. Cogn. Neurosci.* 7, 433–445.

Merabet, L. B., Hamilton, R., Schlaug, G., Swisher, J. D., Kiriakopoulos, E. T., Pitskel, N. B., et al. (2008). Rapid and reversible recruitment of early visual cortex for touch. *PLoS ONE* 3:e3046. doi:10.1371/journal.pone.0003046

Mumford, D. (1991). On the computational architecture of the neocortex I: the role of the thalamo-cortical loop. *Biol. Cyber.* 65, 135–145.

Mumford, D. (1996). "Commentary on banishing the homunculus by H. Barlow," in *Perception as Bayesian Inference*, eds D. C. Knill and W. Richards (Cambridge: Cambridge University Press), 501–504.

O'Craven, K. M., and Kanwisher, N. (2000). Mental imagery of faces and places activates corresponding stimulus-specific brain regions. *J. Cogn. Neurosci.* 12, 1013–1023.

Ogren, M., and Hendrickson, A. (1976). Pathways between striate cortex and subcortical regions in Macaca mulatta and Saimiri sciureus: evidence for a reciprocal pulvinar connection. *Exp. Neurol.* 53, 780–800.

Pascual-Leone, A., and Hamilton, R. (2001). The metamodal organization of the brain. *Prog. Brain Res.* 134, 427–445.

Perkel, D. J., Bullier, J., and Kennedy, H. (1986). Topography of the afferent connectivity of area 17 of the macaque monkey: a double labeling study. *J. Comp. Neurol.* 253, 374–402.

Ponchillia, P. E. (2008). "Non-visual sports and arts: fertile substrates for the growth of knowledge about brain plasticity in people who are blind or have low vision," in *Blindness and Brain Plasticity in Navigation and Object Perception*, eds J. J. Rieser, D. H. Ashmead, F. F. Ebner, and A. L. Corn (London: Lawrence Erlbaum Associates), 283–313.

Raichle, M. E., and Snyder, A. Z. (2007). A default mode of brain function: A brief history of an evolving idea. *Neuroimage* 37, 1083–1090.

Raineteau, O., and Schwab, M. E. (2001). Plasticity of motor systems after incomplete spinal cord injury. *Nat. Rev. Neurosci.* 2, 263–273.

Reed, C. L., Shoham, S., and Halgren, E. (2004). neural substrates of tactile object recognition: An fMRI study. *Hum. Brain Mapp.* 21, 236–246.

Rezak, M., and Benevento, L. A. (1979). A comparison of the organization of the projections of the dorsal lateral geniculate nucleus, the inferior pulvinar and adjacent lateral pulvinar to primary visual cortex (area 17) in the macaque monkey. *Brain Res.* 167, 19–40.

Rockland, K. S. (1994). "The organization of feedback connections from areaV2 (18) to V1 (17)," in *Cerebral Cortex, Vol. 10: Primary Visual Cortex in Primates*, eds A. Peters and K. S. Rockland (NewYork, NY: Plenum), 261–299.

Sabbah, P., Simond, G., Levrier, O., Habib, M., Trabaud, V., Murayama, N., et al. (1995). Functional magnetic resonance imaging at 1.5 T during sensorimotor and cognitive task. *Eur. Neurol.* 35, 131–136.

Sack, A. T., Sperling, J. M., Prvulovic, D., Formisano, E., Goebel, R., Di Salle, F., et al. (2002). Tracking the mind's image in the brain: II. Transcranial magnetic stimulation reveals parietal asymmetry in visuospatial imagery. *Neuron* 35, 195–204.

Sadato, N., Pascual-Leone, A., Grafman, J., Ibañez, V., Deiber, M. P., Dold, G., et al. (1996). Activation of the primary visual cortex by Braille reading in blind subjects. *Nature* 380, 526–528.

Schmolesky, M. (2000). *The Primary Visual Cortex*. Available at: http://webvision.med.utah.edu/VisualCortex.html.

Sereno, M. I., Dale, A. M., Reppas, J. B., Kwong, K. K., Belliveau, J. W., Brady, T. J., et al. (1995). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science* 268, 889–893.

Shin, L. M., McNally, R. J., Kosslyn, S. M., Thompson, W. L., Rauch, S. L., Alpert, N. M., et al. (1999). Regional cerebral blood flow during script-driven imagery in childhood sexual abuse-related PTSD: a PET investigation. *Am. J. Psychiatry* 156, 575–584.

Shipp, S., and Zeki, S. M. (1989). The organization of connections between areas V5 and V1 in macaque monkey visual cortex. *Eur. J. Neurosci.* 1, 309–332.

Sperling, G. (1963). A model for visual memory tasks. *Hum. Factors* 5, 19–31.

Super, H. (2003). Working memory in the primary visual cortex. *Arch. Neurol.* 60, 809–812.

Super, H., Spekreijse, H., and Lamme, V. A. F. (2001a). A neural correlate of working memory in the monkey primary visual cortex. *Science* 293, 120–124.

Super, H., Spekreijse, H., and Lamme, V. A. F. (2001b). Two distinct models of sensory processing observed in monkey primary visual cortex (V1). *Nat. Neurosci.* 4, 304–310.

Suzuki, W., Saleem, K. S., and Tanaka, K. (2000). Divergent backward projections from the anterior part of the inferotemporal cortex (area TE) in the macaque. *J. Comp. Neurol.* 422, 206–228.

Thompson, W. L., Kosslyn, S. M., Sukel, K. E., and Alpert, N. M. (2001). Mental imagery of high- and low-resolution gratings activates area 17. *Neuroimage* 14, 454–464.

Tootell, R. B., Dale, A. M., Sereno, M. I., and Malach, R. (1996). New images from human visual cortex. *Trends Neurosci.* 95, 818–824.

Trojano, L., Grossi, D., Linden, D. E. J., Formisano, E., Hacker, H., Zanella, F. E., et al. (2000). Matching two imagined clocks: The functional anatomy of spatial analysis in the absence of visual stimulation. *Cereb. Cortex* 10, 473–481.

Tyler, C. W., Likova, L. T., Chen, C. C., Kontsevich, L. L., Schira, M. M., and Wade, A. R. (2005). Extended concepts of occipital retinotopy. *Curr. Med. Imag. Rev.* 1, 319–329.

Ungerleider, L. G., and Desimone, R. (1986a). Projections to the superior temporal sulcus from the central and peripheral field representations of Vl and V2. *J. Comp. Neurol.* 248, 147–163.

Ungerleider, L. G., and Desimone, R. (1986b). Cortical connections of visual area MT in the macaque. *J. Comp. Neurol.* 248, 190–222.

Van Brussel, L., Gerits, A., and Arckens, L. (2011). Evidence for cross-modal plasticity in adult mouse visual cortex following monocular enucleation. *Cereb. Cortex* 21, 2133–2146.

Wheeler, M. E., Petersen, S. E., and Buckner, R. L. (2000). Memory's echo: Vivid remembering reactivates sensory-specific cortex. *Proc. Natl. Acad. Sci. U.S.A.* 97, 11125–11129.

Zangaladze, A., Epstein, C. M., Grafton, S. T., and Sathian, K. (1999). Involvement of visual cortex in tactile discrimination of orientation. *Nature* 401, 587–590.

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 June 2012; accepted: 28 November 2012; published online: 18 January 2013.

*Citation: Likova LT (2013) A cross-modal perspective on the relationships between imagery and working memory. Front. Psychology **3**:561. doi: 10.3389/fpsyg.2012.00561*

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

frontiers in
HUMAN NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 20 February 2012
doi: 10.3389/fnhum.2012.00024

Verbal to visual code switching improves working memory in older adults: an fMRI study

Mariko Osaka[1], **Yuki Otsuka**[2] **and Naoyuki Osaka**[2]*

[1] Department of Psychology, Graduate School of Human Sciences, Osaka University, Osaka, Japan
[2] Department of Psychology, Graduate School of Letters, Kyoto University, Kyoto, Japan

Edited by:
Joel Pearson, The University of New South Wales, Australia

Reviewed by:
Claude Alain, Rotman Research Institute, Canada
Axel Kohler, University of Münster, Germany

***Correspondence:**
Naoyuki Osaka, Department of Psychology, Graduate School of Letters, Kyoto University, Yoshida-Honmachi, Sakyo-ku, Kyoto 606-8501, Japan.
e-mail: nosaka@bun.kyoto-u.ac.jp

The effect of verbal to visual code switching training on working memory performance was investigated in individuals aged 63 and older. During verbal working memory task performance, the training group ($n = 25$) was introduced to a verbal to visual code switching strategy while the control group ($n = 25$) was not exposed to such a strategy. Working memory recognition accuracy was enhanced only in the training group. To explore the neural substrates underlying these strategy effects, fMRI was used to measure brain activity in both groups during working memory task performance before and after an attention training period. In a comparison between pre- and post-training sessions, results showed increased activation in the anterior cingulate cortex (ACC). Relative to the control group, the post-training group exhibited increased activation in the left and right inferior parietal lobules (IPLs) and right superior parietal lobule (SPL). These findings suggest that use of a verbal to visual code switching strategy may assist older individuals in the maintenance of information in working memory.

Keywords: working memory, older adults, fMRI, strategy training, attention switching

INTRODUCTION

Working memory serves simultaneous information, storage, and processing functions (Baddeley, 1986). According to the multicomponent model proposed by Baddeley (Baddeley, 1986, 2003; Baddeley and Logie, 1999), working memory is composed of two storage buffers: the phonological loop for storing verbal information and the visuo-spatial sketchpad for the storage of visual information. Because working memory resources are limited, an executive system assigns and coordinates these limited resources for storing and manipulating information (Daneman and Carpenter, 1980; Baddeley and Logie, 1999; Engle et al., 1999).

Working memory impairments are commonly found in older adults. In particular, older individuals often have difficulty performing dual tasks that require both storage and processing (Salthouse, 1996; West, 1996; Craik and Salthouse, 2000). It has been proposed that vulnerability of the frontal lobes accounts for the age-related impairments in memory functioning found in older adults (Raz et al., 1997; Raz, 2000; Madden et al., 2004). In accordance with this hypothesis, decreased prefrontal cortex (PFC) activation has been found in older adults through neuroimaging studies (Bunge et al., 2000; Cabeza, 2001). Age-related reductions in the ability to suspend non-task-related activity and engage areas for carrying out memory tasks have also been reported (Grady et al., 2006).

Several studies have examined whether training improves working memory in older adults (Nyberg et al., 2003; Erickson et al., 2007). Erickson et al. (2007) found that dual-task training induced performance improvements which were correlated with an increase in hemispheric asymmetry and a reduction in age differences in ventral and dorsal PFC activation. The investigators argued that these results provide evidence of brain plasticity in older individuals. Other studies indicated that training in the *n*-back task, which requires executive control of working memory, is associated with improvements in the working memory performance of older adults (Dahlin et al., 2008a; Li et al., 2008). Dahlin et al. (2008b) showed that training in updating information, which requires executive function in working memory, improved significantly in young and older adults relative to controls. Based on these reports, it is likely that the dual-task deficits typically observed in older individuals are caused not only by a decreased ability to maintain information in working memory, but also by deficits in the executive processing system.

The neural basis of the working memory system, particularly executive functioning, is thought to be located in the PFC; recent neuroimaging studies have suggested that the executive system is located in the dorsolateral PFC (DLPFC, BA9/46) and anterior cingulate cortex (ACC) (D'Esposito et al., 1995, 1999; Smith et al., 1996; Cohen et al., 1997; Owen et al., 1998; Smith and Jonides, 1999; Bunge et al., 2000; Osaka et al., 2003, 2004; Linden, 2007). The brain regions underlying working memory appear to drive different executive functions. The DLPFC supports the active maintenance of attention on task goals, whereas the ACC serves as an attentional coordinator, which enables individuals to inhibit inappropriate stimuli while monitoring task performance (MacDonald et al., 2000). ACC activation appears to increase when cognitive tasks require greater attentional demands (Petersen et al., 1988; Posner and Petersen, 1990).

Osaka et al. (2003, 2004) used the reading span test (RST) to compare PFC activation in high-performing normal young adults with those of low performers, and found that ACC and DLPFC activation was greater in high versus low performers. The RST is a commonly used dual task developed to measure working memory resources available for reading sentences and storing a target word associated with each sentence (Daneman and Carpenter, 1980; Just and Carpenter, 1992). Individual differences in working memory capacity are likely derived from recruited activity shared in the DLPFC and ACC (Osaka et al., 2003). In addition to these results for young adults, our previous research in older adults showed less ACC activation relative to younger individuals (Otsuka et al., 2006, 2008). Based on these findings, it is plausible to assume that working memory impairments in older adults are due to ineffective attentional coordination, as supported by a network including the ACC and DLPFC.

In the present study, we investigated the effect of attentional coordination training on dual-task performance in older adults. Strategic training is considered to be an effective method for increasing working memory performance (McNamara and Scott, 2001). For example, strategic training using visual imagery produces increased correct recall performance when older individuals perform a paired association task (Yesavage, 1983; Carretti et al., 2007). In our previous study (Osaka and Nishizaki, 2000) we found that high-RST performers tended to use a mental imagery strategy more often than low-RST performers, who appear to restrict themselves to verbal rehearsal. Based upon Baddeley's multi-component model of working memory (Baddeley, 2003), we introduced for the first time a unique training task that required participants to switch attention from verbal to visual images, (i.e., switching from phonological information to visuo-spatial information), with the intention that this approach would strengthen the executive control function of working memory in older adults. To our knowledge, there have been no published reports on the use of visual image formation as a training strategy to improve working memory in older adults. We hypothesized that attention switching training through the use of image formation would improve elder participants' working memory performance.

To investigate the neural substrates underlying such training effects, we measured fMRI activation while participants performed a working memory task (the RST) before and after the training period. We hypothesized that this training of attention switching from verbal to visual images activated brain regions related to attentional control. Moreover, it is predicted that an activation increase would be found in both the left and right hemispheres according to the engagement of phonological and visuo-spatial information processing simultaneously.

MATERIALS AND METHODS

PARTICIPANTS

Fifty volunteers participated in the experiment. All participants reported that they had no history of neuropsychological or psychiatric disease and all were right-handed. Their educational levels ranged from 9 to 16 years (mean = 12.76, SD = 1.9).

Participants performed the RST in a preliminary behavioral screening using a paper version. In the RST, participants were required to read aloud sets of one to five sentences presented one at a time and to remember a target word for each sentence. Each sentence condition included five trials, and all participants completed five total trials in the experiment. Span score was calculated as the highest level at which a participant could recall all words from the list in three of the five trials. Overall performance on the RST was assessed in terms of the total number of words recalled. RST scores of the participants ranged from 1.5 to 4.0.

Both the training and control groups were matched for age, educational level, and RST estimates in the screening test. Pretest, training, and post-test sessions all occurred on the same day. Twenty-five participants (17 males and eight females, age range = 63–73 years, mean = 68.08, SD = 2.7, educational level range = 9–16, mean = 12.0, SD = 2.3) participated in the strategy training (training group), whereas the other 25 (15 males and 10 females, age range = 63–75 years, mean = 68.50, SD = 3.4, educational level range = 9–16, mean = 11.8, SD = 2.0) did not receive training (control group). The initial RST scores of both groups fell within the same range, from 1.5 to 4.0 (mean = 2.22 in the training group, mean = 2.19 in the control group). Informed consent was obtained from each participant, in accordance with the protocol approved by the Advanced Telecommunications Research Institute International Brain Imaging Center Review Board. All participants were paid a modest honorarium for their participation in the study.

VISUAL IMAGERY TRAINING PROCEDURE

Specific training was provided in the form of a modified RST that required participants to read three sentences and remember three target words. During the training period, participants practiced imagining a target word by drawing a figure that represented the word. Drawing the figures required participants to switch their focus of attention from verbal to visual images. For example, as the sample protocol in **Figure 1** shows, participants were asked to draw the figure of the jet plane using a black pencil when the target word indicated "travel."

In the first part of the training, as **Figure 1** shows, participants read aloud one sentence and then drew a figure representing the target word on a piece of paper within 60 s. Although producing these drawings was initially challenging, the task became easier with practice. In the second part of the training, participants were

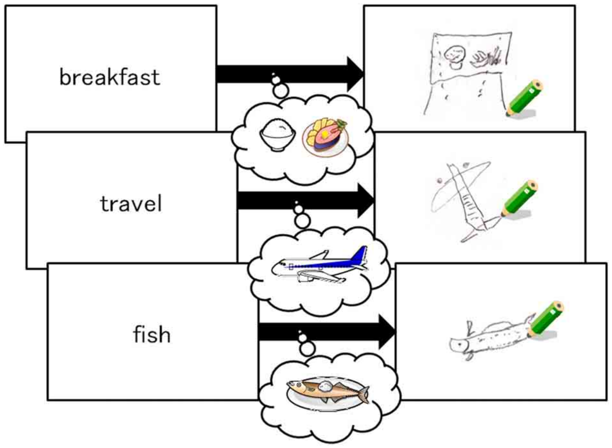


FIGURE 1 | Examples of participant's figure drawings (right) of the target word (left).

asked to form a mental image of a figure representing the target word, instead of drawing the image on paper. Both parts of the practice training consisted of 20 trials composed of three sentences each and took approximately one hour. Participants in the control group read the same sentences and remembered the target words without being instructed to draw or imagine any figures. The number of trials was the same as the training group. Time was limited to 30 min in the first training session and 25 min in the second training session. The control group spent the same amount of time with the material as the training group. None of the participants in the training or control groups exceeded the time allotted.

EXPERIMENTAL PROCEDURE

Each participant was tested on the RST task during both pre- and post-training experimental sessions. The post-training RST experimental session was performed after the training (control training for the control group). The post-training session began approximately 2 h after the end of the pre-training session for both groups. In the experimental sessions, only three sentence trials were performed, and participants read each sentence silently. In the preliminary screening RST, most participants could not remember three target words while they read three sentences aloud. Therefore, we adapted a recognition method to make task performance easier. Participants were instructed to make recognition judgments about probe stimuli (i.e., deciding whether each stimulus was a target word or not) rather than recall the words orally. They were required to read three sentences silently while concurrently remembering three target words in each of the sentences. Each sentence appeared within 6 s, and participants were required to push a button after they finished reading each sentence. After three sentences were presented for a given trial, three probe stimuli appeared at 6 s intervals in the same order as the sentences. Each probe stimulus comprised one word. When participants identified the target in the probe array, they pushed the left key to identify a word as "in set" and the right key to identify a word as "out of set." Half of the probe words were targets while the other half were not.

Figure 2 shows an RST experimental block (modified RST fitted to an fMRI experiment). One experimental RST block comprised three sentences in the reading phase (each lasting 18 s) and an 18 s recognition phase. RST blocks were repeated 16 times in one session. A control condition lasting 18 s was inserted between experimental RST blocks. During the control condition, participants pushed the right and left keys when the words "right" or "left" appeared on the screen. Stimulus sentences ranged from 18 to 24 character lengths (mean = 21.5) and were presented on the screen within a visual angle of 45°, using a mirror attached to a head coil. The target word that the participants were instructed to memorize was underlined in each sentence. The position of the target word in the sentence was counterbalanced across trials. RST blocks were repeated 16 times in post-training session, which was identical to the pre-training session. The RST and control conditions were repeated in the same order during the post-training session, but the stimulus sentences changed so that participants never read the same sentence twice.

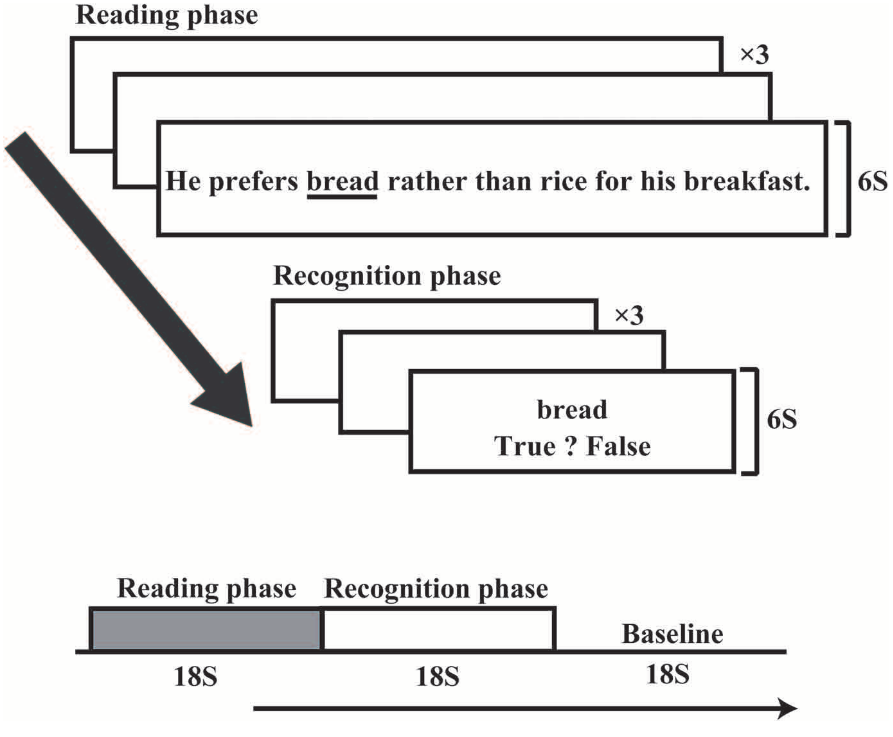


FIGURE 2 | Experimental time course of the RST. One experimental block comprised three sentences. After three sentences were presented, three probe stimuli appeared at 6 s intervals. Reading and recognition phases each took 18 s, followed by an 18 s baseline condition.

fMRI DATA ACQUISITION AND ANALYSIS

Whole brain imaging data were acquired on a 1.5-T MRI scanner (Shimazu-Marconi Magnex Eclipse), using a standard head coil. Head movements were minimized using a forehead strap. For functional imaging, a gradient-echo echo-planar imaging sequence with the following parameters was used: Repetition time (TR), 2000 ms; echo time (TE), 48 ms; flip angle, 80°; field of view (FOV), 22 × 22 cm; and matrix, 64 × 64 pixels. In one experimental session, 376 contiguous images and 20 slices each with a 6 mm thickness were obtained in the axial plane for each participant. After image collection, T1 anatomical images using a conventional spin echo pulse sequence (TR = 12 ms, TE = 4.5 ms, flip angle = 20°, FOV = 25.6 × 25.6 cm, and pixel matrix = 256 × 256) were collected for anatomical co-registration at the same locations as the functional images. Scanner sequences were synchronized with stimulus presentation using Presentation stimulus software (Neurobehavioral System, Inc.). Data were analyzed using SPM2 (Wellcome Department of Cognitive Neurology, London, UK) on Matlab (MathWorks, Sherborn, MA).

Analysis of fMRI data was initially performed for each individual participant for each pre- and post-training session. Six initial images from each scanning session were discarded from analysis in order to eliminate non-equilibrium effects of magnetization, leaving 369 total images for analysis. All functional images were realigned to correct for head movement. We selected images with less than 1 mm movement between scans. After realignment, the anatomical images were co-registered to the mean functional images. Functional images were then normalized with the anatomical image and spatially smoothed using a Gaussian filter (7 mm full width-half maximum). The box-car reference function was adopted for individual analysis to identify voxels under each task condition. Global activity for each scan was corrected using grand mean scaling. Low-frequency noise was modeled using hemodynamic response functions and the corresponding derivative. Single participant data were analyzed using a fixed-effects model, whereas group data from pre- and post-training sessions were analyzed using a random-effects model. Extraction of ROI data and computation of percent signal change of each ROI was performed using MarsBaR (Brett et al., 2002). As a comparison between post- and pre-training sessions, we performed a paired t-test to examine the effect of training (post-training vs. pre-training) in each ROI. A statistical threshold of $p < 0.05$ was used.

RESULTS

BEHAVIORAL PERFORMANCE

Recognition accuracy and response times during the experimental RST were calculated for both pre- and post-training phases. In the training group, the mean percentage recognition accuracies were higher in the post- than in the pre-training session (pre-, 90.33%, SD = 11.70; post-, 97.60%, SD = 5.80). However, in the control group, there was little apparent difference between pre-training and post-training sessions (pre-, 90.43%, SD = 10.70; post-, 91.20%, SD = 9.13). A Two-Way analysis of variance (ANOVA) of group × session showed a significant main effect of training, $F(1,48) = 6.26$, $p < 0.05$. The interaction between group and session was also significant, $F(1,48) = 9.59$, $p < 0.01$, such that a significant increase in the post-training session was found only in the training group (Tukey's HSD *post-hoc* analysis, $p < 0.01$). Mean response times for the training group were 1454 ms (SD = 418.29) in the pre-training session and 1350 ms (SD = 339.38) in the post-training session. The control group response times were 1543 ms (SD = 482.00) in the pre-training session and 1436 ms (SD = 409.00) in the post-training session. A Two-Way ANOVA showed a main effect of training, $F(1,48) = 19.60$, $p < 0.01$. There was a non-significant tendency toward an interaction between group and session, $F(1,48) = 3.06$, $p < 0.08$. However, there was a significant post-training RT decrease in the training group (Tukey's HSD *post-hoc* analysis, $p < 0.01$).

After the post-training session, all participants in the training group reported that they had used an imagery-based strategy to perform the post-training task. However, in the control group, most of the participants (20/25) used verbal rehearsal and five participants used both verbal rehearsal and visual imagery. These five participants did not show any performance increase.

fMRI RESULTS

Figure 3A shows areas of increased activation after training, compared using paired t-tests between the pre- and post-training sessions (voxel-level threshold uncorrected for multiple comparisons, $p < 0.0001$). We applied an uncorrected criterion to focus

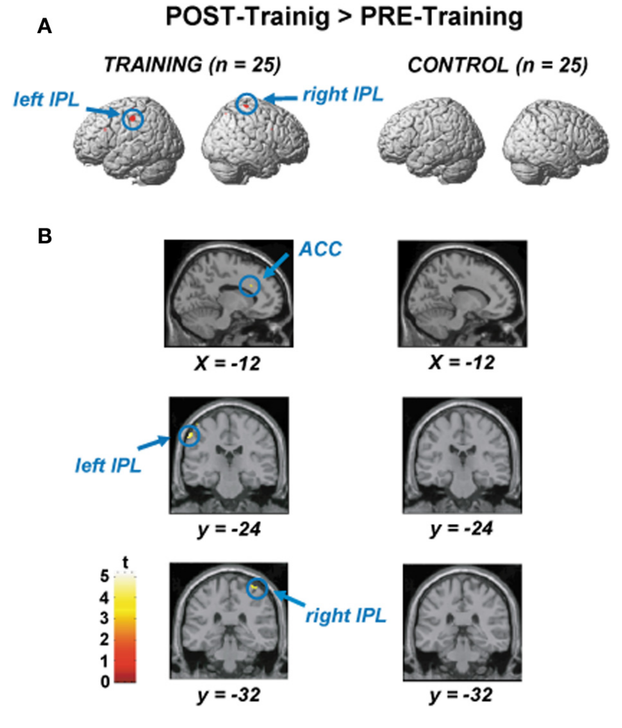


FIGURE 3 | (A) Areas of significantly increased activation after the training, as determined by paired t-test between the pre- and post-training session. The figures on the left side show activations for the training participants, and those on the right show those for the control participants.
(B) Significant ACC activation (upper), left IPL activation (middle), and right IPL activation (lower) by paired t-test between the pre- and post-training session.

upon specific regions where increased activation had previously been reported for younger adults performing the RST (Bunge et al., 2000; Osaka et al., 2003, 2004). The figures on the left side show activations for training participants, and those on the right show activations for control participants.

Figure 3B displays areas for which there was significantly increased activation after training (voxel-level threshold uncorrected for multiple comparisons, $p < 0.0001$). **Figure 3B** shows significantly increased activation in the ACC ($x, y, z = -12, 22, 28$, based on MNI coordinates, Z-score $= 3.33$, t-value $= 3.81$, $p < 0.0001$, uncorrected), which was observed only in the training group. Increased activation during the post-training session was also found in the left ($x, y, z = -60, -24, 48$, Z-score $= 4.22$, t-value $= 5.22$) and right inferior parietal lobule (IPL) ($x, y, z = 40, -32, 68$, Z-score $= 3.49$, t-value $= 4.03$). An activation increase was also found in the right superior parietal lobule (SPL) ($x, y, z = 30, -66, 56$, Z-score $= 3.14$, t-value $= 3.53$; not shown in the **Figure 3**). In the control participants, we could not find any areas showing significant differences between pre- and post-training sessions. Furthermore, we could not find any areas for which there was significantly increased activation after training in the recognition phase both for experimental and control groups.

SIGNAL INTENSITY

The ACC, bilateral DLPFC, bilateral IPL, and bilateral SPL were specified as the regions of interest (ROI) for detecting signal changes. Following the identification of activated areas, percent signal changes in each ROI were obtained for both training and control groups. Time-course data were obtained for the local-maxima of each ROI for each participant. Mean percent signal change was calculated by subtracting the mean blood oxygenation-level-dependent (BOLD) signal at baseline from signals observed during the pre- and post-training sessions.

Figure 4 shows mean signal changes in each ROI during the pre- and post-training sessions in the training and control groups. Two-Way ANOVAs of group × session were performed for the ROIs. Analyses were performed using the STATISTICA software package (StatSoft). A significant interaction between session and group was found for the ACC region, $F(1,48) = 6.93$, $p < 0.01$. Training effects were found in the training group (Tukey's HSD *post-hoc* analysis, $p < 0.05$). A significant main effect of session [$F(1,48) = 5.09$, $p < 0.05$] and an interaction between session and group [$F(1,48) = 5.80$, $p < 0.05$] were also found for the left and right IPL. For the right IPL, training effects in signal intensity increases were found in the training group (by Tukey's HSD *post-hoc* analysis, $p < 0.05$) but not the control group. Similarly, training effects were also found in the left IPL in the training group ($p < 0.05$) but not the control group. A significant main effect of session [$F(1,48) = 7.04$, $p < 0.05$] and an interaction between session and group [$F(1,48) = 4.80$, $p < 0.05$] were also found for the right SPL, with the signal change actually increasing after training (Tukey's HSD *post-hoc* analysis, $p < 0.05$). However, training effects in the recognition phase were not significant in these three regions for either the training group [$F(1,48) = 1.13 \sim 2.23$, $p = 0.26 \sim p = 0.13$] or the control group [$F(1,48) = 0.81 \sim 2.24$, $p = 0.37 \sim p = 0.14$].

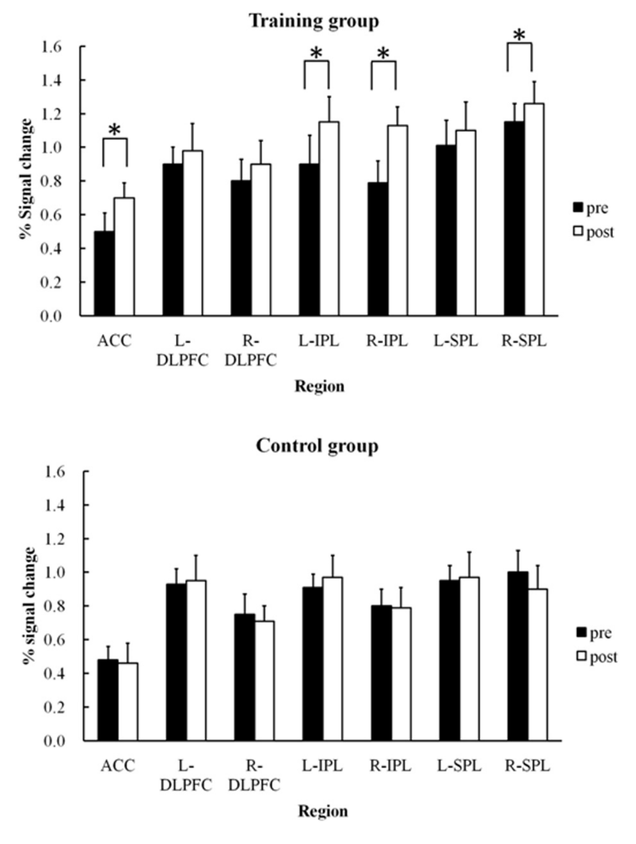


FIGURE 4 | Mean percent signal changes and SDs under training and control groups in each region of interest (ACC, bilateral DLPFC, bilateral IPL, bilateral SPL) during pre- and post-training RST ($*p < 0.05$**).**

DISCUSSION

INCREASING ACC RECRUITMENT

In the present experiment, we found relatively scarce ACC activation during the pre-training RST, and these results were in accordance with previous findings in older adults (Otsuka et al., 2006). Interestingly enough, ACC activation significantly increased in the post-training RST, likely due to the effects of training. The ACC has an attention coordination function, including inhibiting attention directed toward inappropriate stimuli, whereas the DLPFC more generally supports sustaining attention on task goals (MacDonald et al., 2000). Increased ACC activity is also reported to occur with errors on high-conflict trials such as go/no-go, oddball, and two-alternative forced-choice selections (Bush et al., 1998; Carter et al., 1998; Braver et al., 2001), suggesting that the ACC is active in conflict monitoring.

During the encoding phase of the RST in the scanner, participants needed to focus attention on the target words while inhibiting other irrelevant words in the sentence. Because the ACC potentially plays a role in perceiving conflicts between target and non-target words during the dual task, ACC impairment is associated with some deficits in exact executive control. Our participants likely did not have sufficient ACC activation in the pre-training session, leading to difficulties perceiving conflict between target and irrelevant words even if there was sufficient DLPFC activation to sustain attention to the target words. Therefore, without perceiving conflict during the encoding phase, participants face difficulty detecting differences between the target and non-target word during the recognition phase.

However, while participants performed the training task, they had to read one sentence at a time and then draw a figure representing the target word for each sentence. While they were drawing the figures (see **Figure 1**), participants could learn to switch attention from the target to its associated visual image code. Following such attention switching training (i.e., transferring from verbal to visual imagery), participant's likely experienced increased ACC activity as a function of increased attentional coordination. After the training, participants could more easily attend to targets and have more working memory resources available for both reading sentences and maintaining target words. With sufficient working memory resources, our participants could more easily assign the limited resources available for storage and processing during the RST.

Training appeared to promote ACC activation in the present study. With adequate coordination of attention, participants did not face a difficult conflict between the target and irrelevant filler words when they were required to recognize targets. Behavioral data also indicated better performance compared with performance during a pre-training RST. Moreover, response times decreased, which is seemingly indicative of increased confidence in correct responses after training.

INCREASED ACTIVATION IN THE RIGHT IPL AND SPL

The present fMRI data showed that right IPL and SPL activation was more apparent during the post-training RST compared to the pre-training RST. Retention of verbal information in the phonological loop itself activates the left IPL, while spatial information activates the right homologues (Smith et al., 1996). In post-training, all participants in the training group reported performing the task using verbal to visual code switching strategies. A similar result was reported by Carretti et al. (2007), who showed that strategic training using imagery to improve the memorization of word lists improved working memory performance in young and older individuals. According to Logie (2003), the visuo-spatial sketchpad can be subdivided into a visual cache, involved in storing visual images, and an inner scribe, involved in the retention of spatial information via active rehearsal (i.e., by drawing in the current study). The increased right IPL activation observed during the post-training RST may have been evoked by a visuo-spatial inner scribe that was used to sustain images of the target words (Logie and Pearson, 1997). Using the visuo-spatial inner scribe, participants can share the working memory load across two subsystems, providing them with a more efficient attentional coordination system.

It is interesting that providing a visually based imagery strategy improved performance on a verbal task in the present study. These findings suggest the intriguing possibility that when a verbal store within a multi-component model shows age-related performance declines, an alternative visual store can possibly serve to compensate for the deficits.

Right IPL and SPL activation is thought to be involved in the basic attentional process of executive function. A recent study demonstrated that right IPL activity is associated with disengagement and reorienting of attention to a relevant word presented outside the current focus of attention, whereas the intraparietal sulcus (IPS) contributes to an attentional set for the task goal (Corbetta et al., 2008). Furthermore, the right IPL (along with the SPL) likely plays a role in attentional filtering, which is a process whereby irrelevant information is screened out and relevant information prioritized (Minamoto et al., 2010).

Osaka et al. (2007) showed that increased left SPL activation coincided with attentional demands while performing a difficult RST requiring switching attention. While participants performed the difficult RST, the SPL appeared to play a role in reorienting attention and driving cooperative activity between the ACC and DLPFC. The SPL (including the lateral IPS area) is generally related to attentional processes (Culham and Kanwisher, 2001) and is specifically responsible for visuo-spatial attention and working memory (Wager and Smith, 2003). Activation in the posterior region supports the limited capacity of working memory (Linden, 2007). Thus, the SPL may potentially play a shared role in focusing attention along with DLPFC and ACC (Osaka and Osaka, 2007).

In the present study, bilateral activation was found in the IPL, which is often regarded as a mechanism that compensates for declining functional activity while performing working memory tasks (Cabeza et al., 2002; Reuter-Lorenz, 2002; Wager and Smith, 2003). It has also been suggested that reduced asymmetry in the PFC is compensatory when the homologous region can subserve a complementary role in task performance (Colcombe et al., 2005). When the attentional control system weakens, another system develops to compensate for any functional impairment. Bilateral IPL activation observed in the current study may serve as a compensatory system for impairments in working memory. In accord with previous suggestions (Wager and Smith, 2003; Colcombe et al., 2005), the observed right IPL activation possibly compensated for the declines in executive functioning with the help of the visuo-spatial inner scribe. This suggestion indicates that when one store becomes vulnerable to age-related decline, use of an alternative store, such as the visuo-spatial sketchpad, is a potentially reliable method to improve task performance in older adults. In the present experiment, we did not control the strategy use of the non-training group. If participants in the control group had adopted a visual imagery strategy, they might have shown a performance increase in behavioral and neural effects without training. Further investigation is needed to clarify the effects of strategy on working memory performance when only imagery instruction is provided. Moreover, the long-term effects of visual strategy training on working memory performance should be investigated in elderly populations. In addition, Grady et al. (2006) suggested that there is a gradual, age-related reduction in the ability to suspend non-task-related or default-mode activity. Further study will be required to test default-mode activity in older individuals.

In sum, we found that training in a specific verbal to visual code switching strategy improved complex working memory task performance in older adults. After training, fMRI data showed an increase in ACC, bilateral IPL, and right SPL activation in the encoding phase.

Based on these findings, we suggest that age-related impairment in working memory performance is partially caused by ineffective attention coordination, as supported by the ACC in connection with the IPL and SPL. Training in dual task paradigms, which require compensatory strategies such as code switching, could be one potentially effective approach for coping with working memory deficits in aging.

ACKNOWLEDGMENTS

This work was supported by grants from the Japan Society for the Promotion of Science (#23240036) and (#22220003) to Mariko Osaka and Naoyuki Osaka, respectively.

REFERENCES

Baddeley, A. D. (1986). *Working Memory.* Oxford: Oxford University Press.

Baddeley, A. D. (2003). Working memory: looking back and looking forward. *Nat. Rev. Neurosci.* 4, 829–839.

Baddeley, A. D., and Logie, R. H. (1999). “Working memory: the multiple component model,” in *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control*, eds A. Miyake and P. Shah (Cambridge, MA: Cambridge University Press), 28–61.

Braver, T. S., Barch, D. M., Gray, J. R., Molfese, D. L., and Snyder, A. (2001). Anterior cingulate cortex and response conflict: effects of frequency, inhibition and errors. *Cereb. Cortex* 11, 825–836.

Brett, M., Anton, J. L., Valabregue, R., and Poline, J. B. (2002). “Region of interest analysis using an SPM toolbox,” *The 8th International Conference on Functional Mapping of the Human Brain.* June 2–6, Sendai, Japan. Available on CD-ROM in *Neuroimage* 16, Abstract 974.

Bunge, S. A., Klingberg, T., Jacobsen, R. B., and Gabrieli, J. D. (2000). A resource model of the neural basis of executive working memory. *Proc. Natl. Acad. Sci. U.S.A.* 97, 3573–3578.

Bush, G., Whalen, P. J., Rosen, B. R., Jenike, M. A., Mcinerney, S. C., and Rauch, S. L. (1998). The counting Stroop: an interference task specialized for functional neuroimaging–validation study with functional MRI. *Hum. Brain Mapp.* 6, 270–282.

Cabeza, R. (2001). Cognitive neuroscience of aging: contributions of functional neuroimaging. *Scand. J. Psychol.* 42, 277–286.

Cabeza, R., Anderson, N. D., Locantore, J. K., and Mcintosh, A. R. (2002). Aging gracefully: compensatory brain activity in high-performing older adults. *Neuroimage* 17, 1394–1402.

Carretti, B., Borella, E., and De Beni, R. D. (2007). Does strategic memory training improve the working memory performance of young and older adults? *Exp. Psychol.* 54, 311–320.

Carter, C. S., Braver, T. S., Barch, D. M., Botvinick, M. M., Noll, D., and Cohen, J. D. (1998). Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science* 280, 747–749.

Cohen, J. D., Perlstein, W. M., Braver, T. S., Nystrom, L. E., Noll, D. C., Jonides, J., and Smith, E. E. (1997). Temporal dynamics of brain activation during a working memory task. *Nature* 386, 604–608.

Colcombe, S. J., Kramer, A. F., Erickson, K. I., and Scalf, P. (2005). The implications of cortical recruitment and brain morphology for individual differences in inhibitory function in aging humans. *Psychol. Aging* 20, 363–375.

Corbetta, M., Patel, G., and Shulman, G. L. (2008). The reorienting system of the human brain: from environment to theory of mind. *Neuron* 58, 306–324.

Craik, F. I. M., and Salthouse, T. A. (eds). (2000). *The Handbook of Aging and Cognition.* New Jersey: Lawrence Erlbaum.

Culham, J. C., and Kanwisher, N. G. (2001). Neuroimaging of cognitive functions in human parietal cortex. *Curr. Opin. Neurobiol.* 11, 157–163.

Dahlin, E., Neely, A. S., Larsson, A., Backman, L., and Nyberg, L. (2008a). Transfer of learning after updating training mediated by the striatum. *Science* 320, 1510–1512.

Dahlin, E., Nyberg, L., Backman, L., and Neely, A. S. (2008b). Plasticity of executive functioning in young and older adults: immediate training gains, transfer, and long-term maintenance. *Psychol. Aging* 23, 720–730.

Daneman, M., and Carpenter, P. A. (1980). Individual differences in working memory and reading. *J. Verb. Learn. Verb. Behav.* 19, 450–466.

D’Esposito, M., Detre, J. A., Alsop, D. C., Shin, R. K., Atlas, S., and Grossman, M. (1995). The neural basis of the central executive system of working memory. *Nature* 378, 279–281.

D’Esposito, M., Postle, B. R., Ballard, D., and Lease, J. (1999). Maintenance versus manipulation of information held in working memory: an event-related fMRI study. *Brain Cogn.* 41, 66–86.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., and Conway, A. R. (1999). Working memory, short-term memory, and general fluid intelligence: a latent-variable approach. *J. Exp. Psychol. Gen.* 128, 309–331.

Erickson, K. I., Colcombe, S. J., Wadhwa, R., Bherer, L., Peterson, M. S., Scalf, P. E., Kim, J. S., Alvarado, M., and Kramer, A. F. (2007). Training-induced plasticity in older adults: effects of training on hemispheric asymmetry. *Neurobiol. Aging* 28, 272–283.

Grady, C. L., Springer, M. V., Hongwanishkul, D., Mcintosh, A. R., and Winocur, G. (2006). Age-related changes in brain activity across the adult lifespan. *J. Cogn. Neurosci.* 18, 227–241.

Just, M. A., and Carpenter, P. A. (1992). A capacity theory of comprehension: individual differences in working memory. *Psychol. Rev.* 99, 122–149.

Li, S. C., Schmiedek, F., Huxhold, O., Rocke, C., Smith, J., and Lindenberger, U. (2008). Working memory plasticity in old age: practice gain, transfer, and maintenance. *Psychol. Aging* 23, 731–742.

Linden, D. E. (2007). The working memory networks of the human brain. *Neuroscientist* 13, 257–267.

Logie, R. H. (2003). Spatial and visual working memory: a mental workspace. *Psychol. Learn. Motiv.* 42, 37–78.

Logie, R. H., and Pearson, D. G. (1997). The inner eye and the inner scribe of visuo-spatial working memory: evidence from developmental fractionation. *Eur. J. Cogn. Psychol.* 9, 241–257.

MacDonald, A. W. 3rd, Cohen, J. D., Stenger, V. A., and Carter, C. S. (2000). Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science* 288, 1835–1838.

Madden, D. J., Whiting, W. L., Huettel, S. A., White, L. E., Macfall, J. R., and Provenzale, J. M. (2004). Diffusion tensor imaging of adult age differences in cerebral white matter: relation to response time. *Neuroimage* 21, 1174–1181.

McNamara, D. S., and Scott, J. L. (2001). Working memory capacity and strategy use. *Mem. Cognit.* 29, 10–17.

Minamoto, T., Osaka, M., and Osaka, N. (2010). Individual differences in working memory capacity and distractor processing: possible contribution of top-down inhibitory control. *Brain Res.* 1335, 63–73.

Nyberg, L., Sandblom, J., Jones, S., Neely, A. S., Petersson, K. M., Ingvar, M., and Backman, L. (2003). Neural correlates of training-related memory improvement in adulthood and aging. *Proc. Natl. Acad. Sci. U.S.A.* 100, 13728–13733.

Osaka, M., Komori, M., Morishita, M., and Osaka, N. (2007). Neural bases of focusing attention in working memory: an fMRI study based on group differences. *Cogn. Affect. Behav. Neurosci.* 7, 130–139.

Osaka, M., and Nishizaki, Y. (2000). “How working memory works in the central executive,” in *The Brain and Working Memory*, ed N. Osaka (Kyoto: Kyoto University Press), 203–223.

Osaka, M., and Osaka, N. (2007). “Neural bases of focusing attention in working memory: an fMRI study based on individual differences,” in *The Cognitive Neuroscience of Working Memory*, eds N. Osaka, R. H. Logie, and M. D’Esposito (Oxford: Oxford University Press), 99–118.

Osaka, M., Osaka, N., Kondo, H., Morishita, M., Fukuyama, H., Aso, T., and Shibasaki, H. (2003). The neural basis of individual differences in working memory capacity: an fMRI study. *Neuroimage* 18, 789–797.

Osaka, N., Osaka, M., Kondo, H., Morishita, M., Fukuyama, H., and Shibasaki, H. (2004). The neural basis of executive function in working memory: an fMRI study based on individual differences. *Neuroimage* 21, 623–631.

Otsuka, Y., Osaka, N., Morishita, M., Kondo, H., and Osaka, M. (2006). Decreased activation of anterior cingulate cortex in the working memory of the elderly. *Neuroreport* 17, 1479–1482.

Otsuka, Y., Osaka, N., and Osaka, M. (2008). Functional asymmetry of superior parietal lobule for working memory in the elderly. *Neuroreport* 19, 1355–1359.

Owen, A. M., Stern, C. E., Look, R. B., Tracey, I., Rosen, B. R., and Petrides, M. (1998). Functional organization of spatial and nonspatial working memory processing within the human lateral frontal cortex. *Proc. Natl. Acad. Sci. U.S.A.* 95, 7721–7726.

Petersen, S. E., Fox, P. T., Posner, M. I., Mintun, M., and Raichle, M. E. (1988). Positron emission tomographic studies of the cortical anatomy of single-word processing. *Nature* 331, 585–589.

Posner, M. I., and Petersen, S. E. (1990). The attention system of the human brain. *Annu. Rev. Neurosci.* 13, 25–42.

Raz, N. (2000). "Aging of the brain and its impact on cognitive performance: integration of structural and functional findings," in *The Handbook of Aging and Cognition*, eds F. I. M. Craik and T. A. Salthouse (New Jersey: Lawrence Erlbaum), 1–90.

Raz, N., Gunning, F. M., Head, D., Dupuis, J. H., Mcquain, J., Briggs, S. D., Loken, W. J., Thornton, A. E., and Acker, J. D. (1997). Selective aging of the human cerebral cortex observed *in vivo*: differential vulnerability of the prefrontal gray matter. *Cereb. Cortex* 7, 268–282.

Reuter-Lorenz, P. (2002). New visions of the aging mind and brain. *Trends Cogn. Sci.* 6, 394.

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychol. Rev.* 103, 403–428.

Smith, E. E., and Jonides, J. (1999). Storage and executive processes in the frontal lobes. *Science* 283, 1657–1661.

Smith, E. E., Jonides, J., and Koeppe, R. A. (1996). Dissociating verbal and spatial working memory using PET. *Cereb. Cortex* 6, 11–20.

Wager, T. D., and Smith, E. E. (2003). Neuroimaging studies of working memory: a meta-analysis. *Cogn. Affect. Behav. Neurosci.* 3, 255–274.

West, R. L. (1996). An application of prefrontal cortex function theory to cognitive aging. *Psychol. Bull.* 120, 272–292.

Yesavage, J. A. (1983). Imagery pretraining and memory training in the elderly. *Gerontology* 29, 271–275.

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: 26 October 2011; accepted: 05 February 2012; published online: 20 February 2012.

Citation: Osaka M, Otsuka Y and Osaka N (2012) Verbal to visual code switching improves working memory in older adults: an fMRI study. Front. Hum. Neurosci. ***6****:24. doi: 10.3389/fnhum.2012.00024*

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 03 December 2012
doi: 10.3389/fpsyg.2012.00525

Mental imagery for musical changes in loudness

Freya Bailes*, Laura Bishop, Catherine J. Stevens and Roger T. Dean

MARCS Institute, University of Western Sydney, Sydney, NSW, Australia

Edited by:
Joel Pearson, The University of New South Wales, Australia

Reviewed by:
Holly Bridge, University of Oxford, UK
Joel Pearson, The University of New South Wales, Australia

***Correspondence:**
Freya Bailes, Department of Drama and Music, University of Hull, Cottingham Road, Hull HU6 7RX, UK.
e-mail: f.bailes@hull.ac.uk

Musicians imagine music during mental rehearsal, when reading from a score, and while composing. An important characteristic of music is its temporality. Among the parameters that vary through time is sound intensity, perceived as patterns of loudness. Studies of mental imagery for melodies (i.e., pitch and rhythm) show interference from concurrent musical pitch and verbal tasks, but how we represent musical changes in loudness is unclear. Theories suggest that our perceptions of loudness change relate to our perceptions of force or effort, implying a motor representation. An experiment was conducted to investigate the modalities that contribute to imagery for loudness change. Musicians performed a within-subjects loudness change recall task, comprising 48 trials. First, participants heard a musical scale played with varying patterns of loudness, which they were asked to remember. There followed an empty interval of 8 s (nil distractor control), or the presentation of a series of four sine tones, or four visual letters or three conductor gestures, also to be remembered. Participants then saw an unfolding score of the notes of the scale, during which they were to imagine the corresponding scale in their mind while adjusting a slider to indicate the imagined changes in loudness. Finally, participants performed a recognition task of the tone, letter, or gesture sequence. Based on the motor hypothesis, we predicted that observing and remembering conductor gestures would impair loudness change scale recall, while observing and remembering tone or letter string stimuli would not. Results support this prediction, with loudness change recalled less accurately in the gestures condition than in the control condition. An effect of musical training suggests that auditory and motor imagery ability may be closely related to domain expertise.

Keywords: mental imagery, loudness, music, motor processing, melody, working memory

INTRODUCTION

Musicians imagine music during mental rehearsal (Holmes, 2005), when reading from a score (Brodsky et al., 2003), and while composing (Covington, 2005; Bailes, 2009; Bailes and Bishop, 2012). An important characteristic of music is its temporality, and among the parameters that vary through time is sound intensity, perceived as patterns of loudness. Desired increases in loudness can be notated in a score as "crescendi," while decreases can be notated as "decrescendi." There is anecdotal evidence that imagined music can vary in its overall "loudness" level (Trusheim, 1991; Sacks, 2007), but empirical evidence of imagery for loudness is inconclusive (Intons-Peterson, 1992; Pitt and Crowder, 1992; Wu et al., 2010). Moreover, imagery for changes in loudness remains a relatively neglected topic.

Mental representations of pitch and melody have been shown to involve auditory (Deutsch, 1970; Keller et al., 1995), verbal (Keller et al., 1995), and motor processing (Mikumo, 1994; Finney and Palmer, 2003). Yet, how we represent musical changes in loudness is unclear. In the current study, the modalities that contribute to imagery for loudness change were investigated. Theories suggest that our perceptions of loudness change relate to our perceptions of force or effort, implying that a motor representation is involved. For example, we have proposed a causal chain whereby the literal Force (F) with which a player activates an instrument is transmitted as the physical Energy (E) of the sound. For both player and listener, an appreciation of the Effort (E) involved in this contributes to the perception of Loudness (L) and Arousal (A; Dean and Bailes, 2008, 2010). This proposal (FEELA) was based on computational analyses of patterns of acoustic intensity, and it closely relates to ideas of Todd (1992). Some empirical evidence to corroborate this theory is provided by Eitan and Granot (2006). In an experiment in which participants were asked to imagine a figure moving to musical stimuli, an association was found between crescendi (increase in sound intensity) and the energy of the movements of the imagined figure. For example, figures were imagined moving from a walk to a run as the loudness increased. Following from this theoretical and empirical work, we hypothesize that motor representations contribute to imagery for musical changes in loudness.

The role of motor representations in auditory imagery is critical to distinctions that have been made between an "inner ear" (acoustic imagery) and an "inner voice" (subvocal rehearsal) (Smith et al., 1992). Both have been associated with the working memory subsystem known as the "phonological loop," involved in processing and memorizing verbal material (Kalakoski, 2001). Articulatory suppression has a negative impact on performance in tasks likely to elicit auditory mental rehearsal, suggesting that subvocalization can contribute to the generation of auditory imagery (Smith et al., 1995). Brodsky et al. (2003) used a musical task that is commonly performed by musicians, namely reading musical notation in silence, to examine the impact of concurrently performing auditory versus phonatory interference tasks on the successful imagining of a notated melody. Phonatory interference by way of concurrently singing or humming a different melody to that in the score proved the most distracting, pointing to acoustic imagery, and subvocal rehearsal in imagery for melodies. However, this research is limited in its focus on imagery for sequences of discrete events (notes), rather than on the dynamic properties of music. It is also shaped by the original concern of working memory models to describe verbal processing, with motor considerations in auditory imagery restricted to vocal production (see also Hubbard, 2010).

In the current study, an experiment was conducted to investigate the modalities that contribute to imagery for loudness change. Investigations of visual imagery and working memory have used an interference paradigm as the means to disrupt different types of processing. For example, in a study of movement imagery in rock climbing, Smyth and Waller (1998) trained participants on two routes, one vertical and the other horizontal. After training, participants imagined climbing the routes under control conditions and with one of three secondary tasks – dynamic visual noise, spatial tapping, or kinesthetic suppression. The secondary or interference tasks affected differentially the duration of horizontal and vertical routes leading the authors to conclude that there are multiple and complex forms of processing action and imaging movement. An investigation of memory span for ballet movements by professional dancers showed no effect of dynamic visual noise as a form of visual suppression on span and, by contrast, a significant effect of a motor interference task on span (Rossi-Arnaud et al., 2004). Pearson et al. (2008) manipulated background luminance during an imagery or a feature-based attention task. Differential effects of background luminance on the two tasks were used to distinguish effects attributable to imagery from those attributable to task instructions (Experiment 4). Following the tradition of an interference paradigm to probe working memory processes, we devised an interference paradigm in which a trial comprises two interleaved memory tasks, designed to test the interference of remembering material from one on the other. Rehearsal is generally required for maintenance of material in short-term memory (Berz, 1995), and the rationale of the current experiment is that such rehearsal will be variously disturbed by material of different modalities. Distractor tasks were designed to differentially place loads on verbal, auditory, and visuo-motor processing. The design requires the concurrent rehearsal of unfamiliar musical (melodies and loudness change scales) and distractor stimuli in working memory.

One of the challenges presented by ubiquitous real-world stimuli such as music is that it is time-varying. Prior studies of mental imagery have investigated more static material such as pictures, objects, or alphanumeric characters. Thus there is a need to evaluate contemporary accounts of imagery in the context of sequential and temporally structured and varied material. In turn, this requires the development of new methods of: stimulus presentation, on-line generative responding, and analysis of the resulting production (time-series) data. In short, investigation of imagery in music demands a method of responding that captures its temporal unfolding, and this may be best achieved by way of a production (rather than recognition) task. Accordingly, we used a continuous response paradigm that encouraged participants to imagine loudness change stimuli. Participants moved a volume slider to indicate increases and decreases in the "loudness" of the imagined stimuli. The advantages of such an approach are twofold. First, enacting the response in time is more likely to recruit a mental image of the stimulus than performing a stimulus recognition task. Second, movement is integral to this response mode, respecting our hypothesized link between representations of intensity and motor effort.

The principal hypothesis was that imagining changes in loudness would be disrupted by concurrently remembering movement sequences (presented visually). However, the loudness change stimuli in the current experiment comprised loudness changes in the sounding of ascending and descending scales, and pitch (the patterns of note ascent and descent) is integral to a representation of such stimuli, so it was possible that tone sequences would also impair mental imagery. Finally, if participants chose a strategy of labeling increases and decreases of intensity as "up" and "down" respectively in order to remember the loudness change scale stimuli, then a concurrent verbal task of remembering letters could be expected to interfere with the task of recreating the loudness change stimuli, perhaps suggestive of a verbal representation rather than a mental image. In line with previous research (e.g., Williamson et al., 2010), we describe letters as verbal stimuli due to their possible encoding in word form.

In experiments on working memory for actions that use an interference paradigm (e.g., Smyth and Pendleton, 1989) there is a problem, rarely discussed, of the similarity of intervening material with to-be-remembered material. For example, greater interference for recalling configurational movements of the body has been observed when intervening material consists of configurational movements of the body than spatial locations. The conclusion is then drawn that configurational movements are coded in working memory by a spatial plus kinesthetic system. However, there is also much greater similarity between the intervening and the to-be-remembered material in the configurational interference condition than in the spatial interference condition. This problem of similarity between to-be-remembered and interference material is addressed in the present experiment by having distractor stimuli that are all dissimilar from the to-be-remembered loudness change scale material.

Conductor gestures were used as visuo-spatial distractor stimuli of relevance to the communication and understanding of musical intensity, and to represent a motor sequence. Action-observation theories would suggest that observing a sequence of conductor gestures necessarily activates motor representations. Simulation theory (see Berthoz, 1996; Grush, 2004) also argues that we observe and understand the actions of others by covertly simulating them. Accordingly, observing a visual sequence of movements with a view to recalling them would involve simulating their production.

It also was important in the current study to separately determine whether these auditory, verbal, and motor distractor tasks would impact on imagery for *melodic* material, as suggested by past research. We hypothesized that imagining melodies would be disrupted by concurrently remembering tone sequences (presented aurally). However, we expected that imagining melody would involve motor processes too, such that having to remember a sequence of movements while performing a test of imagery for melody would also interfere. As when imagining changes in loudness, remembering visually presented letters while attempting to imagine melody could interfere if the letters were encoded sonically rather than visually.

MATERIALS AND METHODS

A within-subjects design comprised two different imagery tasks (melodies, loudness change scales), each with four different distractor conditions (control, letter sequence, tone sequence, movement sequence), generating eight different experiment conditions.

PARTICIPANTS

Participants ($N = 32$, 17 female, 15 male) able to read musical notation were recruited from universities and community music societies in greater Sydney. They received a small travel reimbursement (15 AU$). Ages ranged from 22 to 71 years ($M = 41.6$, $SD = 16$). Participants had a mean Ollen Musical Sophistication Index (OMSI; Ollen, 2006) of 595 (range 119–993, where a score >500 classifies the participant as "more musically sophisticated" and a score <500 as "less musically sophisticated"), with a mean of 8.7 years of musical training (range 1–16).

STIMULI

Melodies

For the melody imagery task, 28 melodies were selected from the Australian Music Examinations Board (AMEB) aural test syllabus for grades 2–3 (AMEB, 2002). In the current experiment, these melodies were designed to be retained in memory and to be related to a visual score. The melodies were monophonic, between eight and 12 notes in length, and written in a variety of different major and minor keys. All melodies ended on the tonic of the key. The audio files were generated and recorded through a Yamaha Disklavier 3 MIDI (Musical Instrument Digital Interfance) grand piano, controlled by Max/MSP. The velocity of each note was held constant, and each melody was made to span 8 s. Visual scores of the melodies were written in Sibelius.

Half of the melodies were altered to produce the "different" test stimuli, while half were unaltered for the "same" test stimuli. Three types of alteration were made: (1) the order of two consecutive pitches was reversed, as in the "Exchange" comparison of Mikumo (1994; four melodies), (2) a "step" was exchanged for a "leap" (four melodies), or (3) a "leap" was exchanged for a "step" (four melodies). As in Dowling (1978), a step was defined as an interval of three semitones or fewer, and a leap was an interval of four semitones or greater. Within the constraints of each alteration type, changes were designed to be visually non-obvious (e.g., no new accidentals or repeated notes) and avoid introducing dissonance into the melodic context. Contour was disrupted for five of the melodies. Alterations occurred evenly across beginning, middle, and end locations of the melody, but never occurred on the first or last notes.

Loudness change scales

For the loudness change imagery task, 16 different loudness change patterns comprising sequences of crescendi and decrescendi were produced. Eight loudness patterns were superimposed on an ascending/descending (in pitch) one octave major scale, while eight were superimposed on a descending/ascending (in pitch) scale. The audio files were generated and recorded through the Disklavier, controlled by Max/MSP. Each note in the scale was 500 ms, so that all the scales spanned 8 s. Half of each scale type (e.g., ascending/descending) began with a crescendo, while the other half began with a decrescendo. Each stimulus comprised between two and four loudness changes (crescendo or decrescendo), lasting between three and eight notes each. Loudness changes were implemented by manipulating the MIDI signal sent from Max/MSP to the Disklavier for the velocity at which each note should be played. The minimum and maximum note velocities were the same for all loudness changes (MIDI note velocity range from 20 to 60). No more than two consecutive notes shared the same note velocity.

For use in the test phase, visual scores of the scales, without loudness change markings, were written in Sibelius. Powerpoint and the screen capture software Capture Me were then used to record videos of each scale being gradually revealed at the rate of one note per 500 ms.

Distractor stimuli

Letter sequence. The same total set of six letters as used by Williamson et al. (2010) was used to construct four-letter visual sequences. Three letters from this set rhyme (B, D, G) and were expected to be easily confused in phonological memory, while three do not (M, Q, R). "Different" trials at test replaced a letter from the presentation sequence with either a rhyming letter (half) or a non-rhyming letter (half). Letter sequences were created in Powerpoint and recorded as videos using Capture Me. Each letter remained on the screen for 2 s with no gap in-between.

Tone sequence. Four-tone sequences were generated in Audacity. Pure sine tones were used, each being 2 s long, presented sequentially with no gap in-between. For each trial, the four tones were selected from outside the key[1] of the corresponding melody or loudness change scale stimulus. A set of possible tone sequences was constructed for each key prior to the experiment, from which one tone sequence was randomly selected once a melody or scale from the corresponding key was presented. "Different" trials at test replaced a tone from the presentation sequence with a tone also from outside the key of the melody or scale stimulus.

Movement sequence. A set of 10 clips of musical conducting were selected from the videos provided in "Expressive Conducting" (Wiens, 2002). The clips were selected to represent varied conducting gestures that ranged from slow to fast, and from small to large. No attempt was made to control the relationship between gestures and the pitch or loudness content of melodic and loudness change scale stimuli. While different gestures might be associated with the communication of different levels of musical sound intensity,

---

[1] A musical key is described by the scale to which most of the notes in a piece of music conform. For example, if most of the notes in a piece are in the scale of C major, with important pitches such as "C" (tonic) or "G" (dominant) occurring particularly often, its key is probably C major.

the subjective nature of this was beyond the scope of the present study, and so beyond choosing varied gestures, no attempt was made to control for level of expressed intensity in their selection. The movements were recorded from the back right of the conductor such that the face was not visible, and the white baton could be seen against the black background. The baton was visible at all times, and the left hand could not be seen. Original clips that were shorter than 2 s (the shortest original clip was 1.8 s) were stretched in Adobe Premiere Pro CS4 to bring them to the requisite length. Movement sequences were constructed in iMovie. They comprised three silent clips with a 1-s blank (black) screen in-between each clip. For half of the sequences, one of the three clips was replaced by another clip to create a corresponding "different" sequence. While letter and tone distractor sequences comprised four distinct events, pilot testing of gesture sequences suggested no difference in recognition accuracy between three and four gestures. However, in the pilot, participants appeared to be discouraged by the difficulty of remembering a longer sequence. This is comparable to observations from research in working memory for dance movements. Experiments on working memory span for body actions typically report a mean span of three actions for adult participants (e.g., Wood, 2007; Wachowicz et al., 2011).

APPARATUS

The experiment was run from a MacBook (OS X 10.5.8). Stimuli were presented and data were collected using a custom-made patch in Max/MSP. Participants wore Sennheiser HD 650 headphones, and data in the loudness change imagery task were collected by means of an I-CubeX push v1.1 slider facing away from the participant at a slight upwards incline (**Figure 1**).

PROCEDURE

The study was approved by the Human Research Ethics Committee of the University of Western Sydney. Written informed consent was first sought to participate in the study, and general instructions about the format of the experiment were provided. Participants began by filling out the OMSI questionnaire. Trials for each of the eight experiment conditions (two imagery tasks × four distractor tasks) were blocked, and instructions specific to the condition were provided at the start of that block. Participants then performed a practice trial for the block, and were given an opportunity to ask the experimenter any questions that they had before proceeding to the experiment trials. Presentation of the eight blocks was random. Each of the 24 melody stimuli was presented once without repeat across melody imagery blocks. Each of the 14 loudness change scale stimuli was presented once or twice across loudness change imagery blocks (the Max/MSP program randomly selected without replacement all 14 stimuli, then began the process again until 10 of the list had been presented a second time).

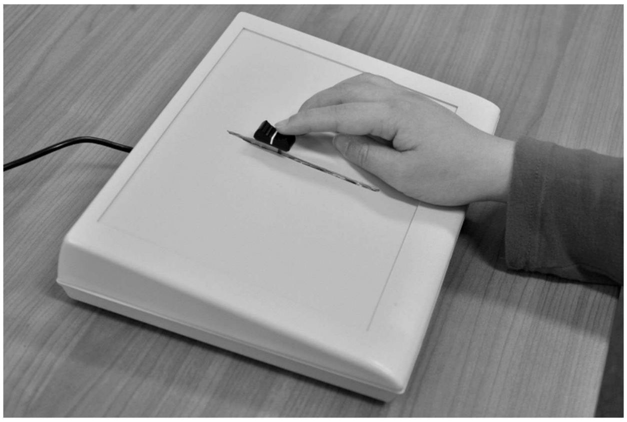


FIGURE 1 | Slider used to record imagery for loudness change. Photograph of an I-CubeX push v1.1 slider, used to measure the changes in loudness imagined by participants during the "Imagery for loudness change" task.

Imagery for melodies

In a melody trial, a melody was sounded, followed by presentation of the distractor stimulus (letter sequence, tone sequence, movement sequence, or control period of 8 s). Immediately after the distractor stimulus presentation, a visual score of the melody appeared on screen, and participants indicated as quickly as possible whether the score was the same or different to the melody that they had heard, by comparing their mental image of the melody with the score. "Same" and "Different" buttons appeared next to each other on the screen, and participants used a mouse to indicate their response. Following the melody test, the trial ended for the control condition, or a distractor recognition test appeared, in which participants were presented with the distractor stimulus (letter, tone, or movement sequence) and used the same buttons to indicate whether the distractor test sequence was the same or different to the distractor stimulus which had originally been presented. **Figure 2** shows the procedure.

Imagery for loudness change

In a loudness change scale trial, a scale modulated in acoustic intensity (loudness change) was sounded, followed by presentation of the distractor stimulus (letter sequence, tone sequence, movement sequence, or control period of 8 s). Immediately after the distractor stimulus presentation, an unfolding visual score of the notes of the scale was presented on the screen, and participants used a volume slider to indicate their mental image of the loudness change profile of the scale that they had heard. Notes appeared on the score at the same pace as they had been sounded at the start of the trial (i.e., one note per 500 ms), and participants were instructed to match the timing of their slider adjustments to the timing of the unfolding visual score. To ensure that slider movements began from the appropriate imagined loudness level at the start of the scale, a 2-s long orientation period was provided, visually marked by a yellow circle on the screen. During this time participants were to move the slider to the level that they thought best represented the opening loudness of the scale, before going on to indicate the loudness changes corresponding to the visually unfolding scale[2]. Following the loudness change

---

[2] Since moving a slider is a motor task, it was important to ensure that any impaired performance associated with movement sequence distractor conditions could not be attributable to physical motor production demands. A separate experiment required 12 participants (six from the current experiment and six new participants) to use the

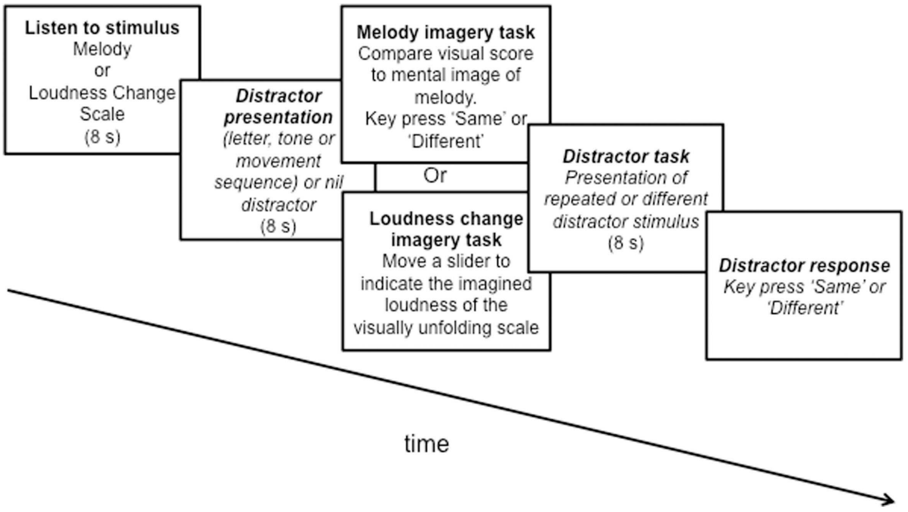


FIGURE 2 | Details of experimental procedure. Schematic of an experimental trial. In the imagery for melodies task, the procedure is represented in the upper panel of frame 3. In the imagery for loudness change task, the procedure is represented in the lower panel of frame 3.

scale test, the trial ended for the control condition, or a distractor recognition test appeared, in which participants were presented with the distractor stimulus (letter, tone, or movement sequence) and used "Same" and "Different" buttons to indicate whether the distractor test sequence was the same or different to the distractor stimulus which had originally been presented. **Figure 2** shows the procedure.

Each block comprised six trials and a practice. The experiment lasted approximately 45 min.

ANALYSIS

Loudness change scale recall scores

Participant responses for the loudness change scale recall task comprised the series of slider values produced by each participant on each trial. **Figure 3** illustrates sample participant response profiles and the corresponding reference scale key velocity profiles for four trials. Each participant's performance on the task was assessed by measuring the similarity between their response profiles and the corresponding scale key velocity profiles. Dynamic time warping (DTW; Giorgino, 2009) was used to compare participant response profiles and scale key velocity profiles. DTW is suitable for use with time-series data as it does not require independence of data points within the series. It identifies points along test (i.e., participant response) and reference (i.e., scale key velocity) data series that most likely correspond with each other. An average distance between profiles per event is then calculated that is independent of profile length. Participant profiles varied in length, since slider position was sampled continuously (every 100 ms) only when the slider was in motion. Scale key velocity profiles, originally 16 events in length, were therefore stretched so that they were continuous and had the same number of events as each individual participant profile. Temporal relationships between loudness changes were maintained through this step of the analysis. The DTW distance between each participant response profile and corresponding key velocity profile was then calculated. A total of 0.3% of the data from four trials (belonging to three participants) were excluded as outliers as DTW values were greater than 2.5 × SD from the participant's mean.

Melody recognition scores

Accuracy in the melody recognition was calculated as the proportion of correct responses (i.e., correct identification of a different or same stimulus) from all given responses per condition. Four participants were at chance performance only in the nil distractor (control) condition, and consequently they were excluded from analyses on the melody task.

Multi-level linear modeling

Multi-level linear modeling was used (lme4 in the statistical program "R") to determine how well the distractor condition was able to model the scores. One advantage of this approach over ANOVA is the possibility of modeling random effects so that different intercepts and gradients for individuals and block order can be included, thus controlling for intersubject variability or order effects. Models were developed stepwise, using interference condition as a predictor, and testing the impact of individuals, block order, OMSI, and years of musical training as random effects. Model selection used the Bayesian Information Criterion (BIC) to determine the most parsimonious fit. Confidence intervals (CI) were calculated as Highest Posterior Density estimates obtained by Markov Chain Monte Carlo sampling.

---

slider to indicate the loudness changes that they were hearing in the moment. Since this perceptual task did not require participants to rehearse or recall loudness change, we did not expect any interference from concurrent distractor conditions. Indeed, the participants' accuracy in marking loudness changes under letter and movement sequence conditions was not significantly different ($z = -0.235$, $p = 0.81$), suggesting that any relative impairment we might find from rehearsing movement sequences in the current imagery experiment could not be attributable to the motor production of the task.

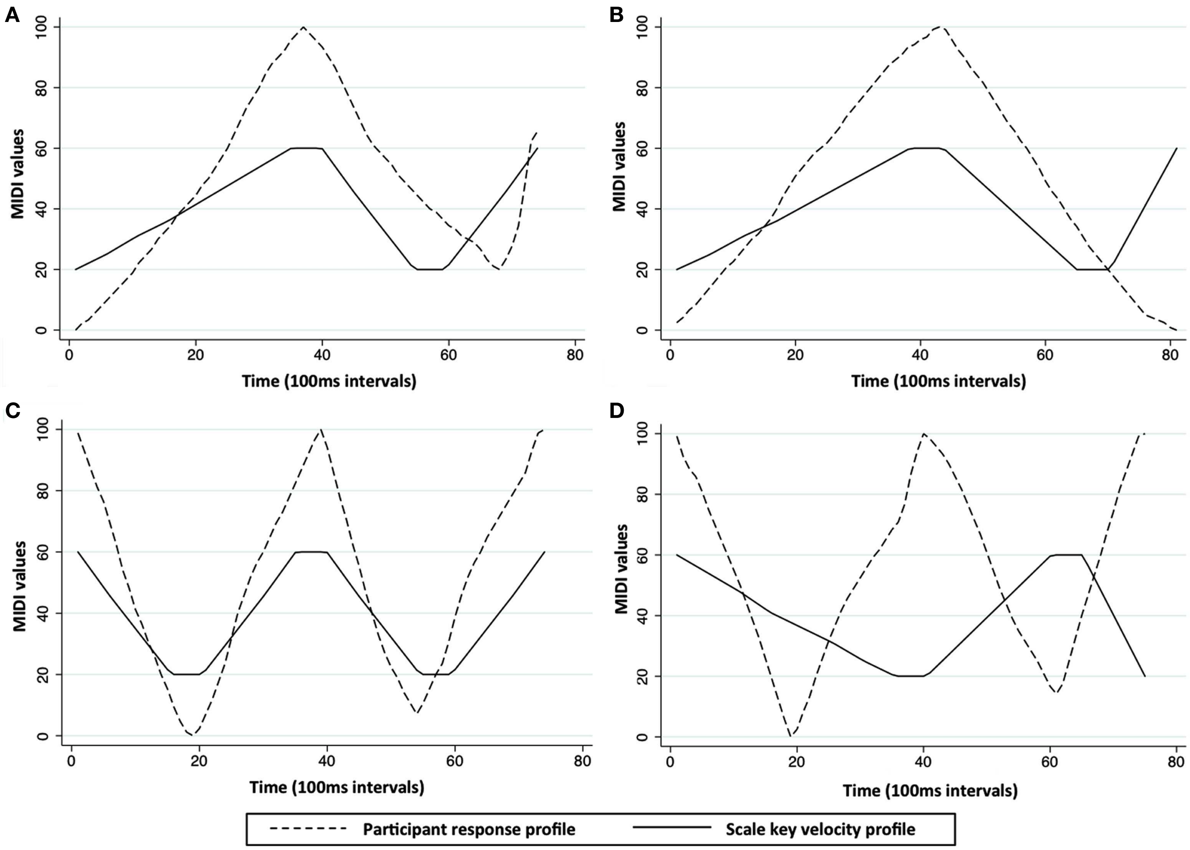


FIGURE 3 | Sample participant response profiles and scale key velocity profiles. Each plot shows one participant's response profile for a single trial and the corresponding scale key velocity profile that the participant was attempting to recall and map out while imagining the scale. DTW provides a measure of the average distance between participant response and scale key velocity profiles per 100 ms time interval. Plots **(A)** and **(C)** show trials in which loudness change profiles were imagined correctly and with precise timing; plots **(B)** and **(D)** show trials in which loudness change profiles were not imagined correctly.

RESULTS

IMAGERY FOR LOUDNESS CHANGE

In the best fit multi-level linear model, recall of loudness change after the movement sequence distractor was significantly worse than recall under the nil distractor condition ($\beta = 1.04$, $t = 2.07$, 95% CI: 0.04 to 2.07, $p = 0.04$). **Figure 4** displays the DTW distances in each of the distractor conditions. Neither recalling loudness change scales under letter sequence distraction ($\beta = 0.90$, $t = 1.86$, 95% CI: −0.07 to 1.84, $p = 0.06$), nor tone sequence distraction ($\beta = 0.42$, $t = 0.87$, 95% CI: −0.50 to 1.44, $p = 0.38$) was significantly worse than in the nil distractor control.

The optimized model also included a random intercept for each individual participant (SD = 2.39), for years of musical training (SD = 0.89), and for block order (SD = 0.67).

IMAGERY FOR MELODIES

In the model of accuracy in the melody recognition task, performance was not significantly different from the nil distractor condition under tone sequence distraction ($\beta = 0.08$, $t = 1.70$, 95% CI: −0.17 to 0.03, $p = 0.09$). Neither melody recognition under letter sequence distraction ($\beta = 0.06$, $t = 1.26$, 95% CI: −0.16 to 0.04, $p = 0.21$) nor movement sequence distraction ($\beta = 0.09$, $t = 1.91$, 95% CI: −0.19 to 0.01, $p = 0.06$) was significantly worse than in the nil distractor control. Results are displayed in **Figure 5**.

As in the model of DTW distances as an index of imagery for loudness change, the optimized model of accuracy in recognizing melodies included a random intercept for each individual participant (SD = 0.10). Here no significant contribution of years of musical training or block order was found.

Melody recognition in the nil distractor/control condition was significantly better than chance [$t(26) = 33.5, p < 0.0001$], making it unlikely that performance in this task was at floor.

MEMORY FOR DISTRACTOR STIMULI

It was also of interest to compare memory for the distractor stimuli following performance in each of the quite different loudness change scale recall and melody recognition tasks. **Table 1** summarizes the proportions of correctly remembered letter, tone, and movement sequences.

During the loudness change scale task, correct recall of the distractor sequences differed significantly by distractor type, as assessed by a repeated measures ANOVA of accuracy in

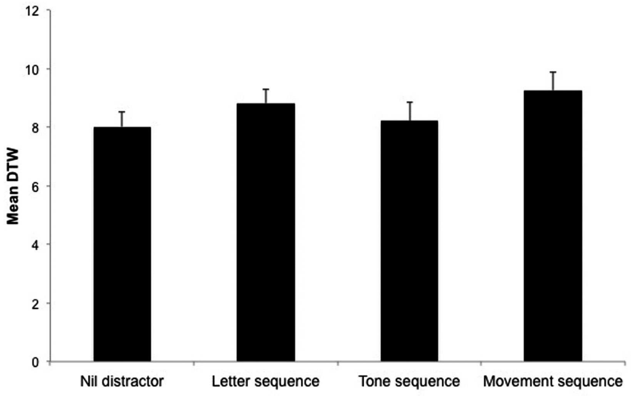


FIGURE 4 | Mean loudness change scale recall score (DTW). Mean DTW distances between slider response and stimulus intensity profiles under different distractor conditions. Small values resemble close loudness change reproduction, i.e., better task performance. Error bars represent standard error.

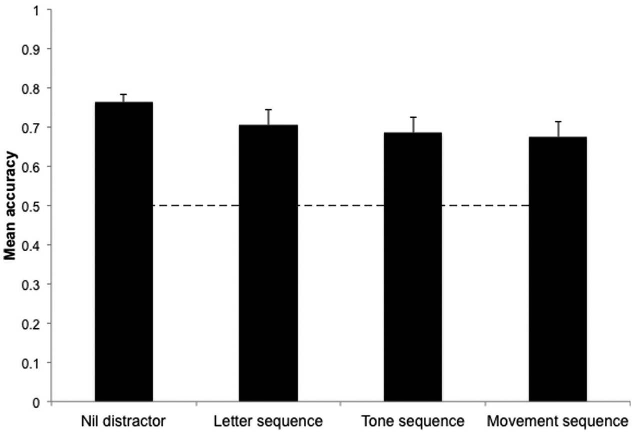


FIGURE 5 | Mean accuracy in melody recognition under different distractor conditions. Error bars represent standard error. Dashed line represents chance level accuracy.

Table 1 | Mean and SD (in parentheses) of accuracy in distractor stimulus recognition.

	Letter	Tone	Movement
Loudness change scale recall	0.87 (0.1)	0.76 (0.2)	0.55 (0.2)
Melody recognition	0.86 (0.1)	0.74 (0.2)	0.68 (0.2)

the distractor task, $F(2{,}94) = 28.11$, $p < 0.001$. Planned contrasts using a Bonferroni adjusted α of 0.02 show that letter sequences were recognized better than tone sequences $t(31) = 2.64$, $p < 0.01$, which were recognized better than movement sequences, $t(30) = 6.94$, $p < 0.001$.

The correct recall of distractor sequences in the melody recognition task also differed significantly by distractor type, $F(2{,}94) = 8.11$, $p < 0.001$. Once again, letter sequences were recognized better than tone, $t(31) = 3.16$, $p < 0.01$, and movement sequences $t(30) = 4.56$, $p < 0.001$, but accuracy recognizing tone and movement sequences did not differ, $t(30) = 1.16$, $p = 0.05$.

IMAGERY AND MUSICAL EXPERIENCE

No relationship was found between OMSI and score when modeling performance in either the melody recall or loudness change scale reproduction tasks. The OMSI is designed to categorize participants as more (>500) or less (<500) musically sophisticated. Comparing the imagery scores of participants categorized in this way confirmed the result from linear modeling that there were no significant differences on either imagery task along this dimension. However, years of musical training contributed to the model of accuracy in loudness change scale reproduction. Furthermore, a positive correlation of years of musical training with accuracy on the melody imagery task was found $r(30) = 0.4$, $p < 0.05$ (two-tailed).

DISCUSSION

This experiment aimed to determine the disruptive effects of rehearsing letter, tone, and movement sequences on mental imagery for changes in loudness. As predicted, rehearsing a movement sequence in mind significantly impaired the recall of loudness change scales. Rehearsing tone sequences did not, though rehearsing letter sequences, a task which could have involved subvocal motor rehearsal, came close to producing a significant impairment. Analyses of how well participants were able to remember the different distractor stimuli revealed that movement sequence recognition was consistently weaker than the recognition of the other distractor sequences. Equating the difficulty of tasks that are to be used in working memory experiments is a vexed issue that receives relatively little discussion. While recognizing conductor gestures might be regarded as more difficult, it is just as likely that the stimuli and task are less familiar than performing a task containing letters or musical tones. Familiarity refers to having knowledge of the material in long-term memory. Thus familiarizing participants with novel material such as gestures within an experiment is one way in the future that could strengthen task comparability. Alternatively, unfamiliar words and tones could be used to be more comparable with the novel conducting gestures.

It seems likely that retaining the movement sequences presented a substantial cognitive load during the performance of any concurrent memory task. However, memory for melodies was only marginally impaired by the movement distractor task, and so its impact primarily concerned the specific task of reproducing imagined changes in loudness. While a motor response was required to reproduce these imagined loudness changes, evidence from a separate experiment suggests that mentally rehearsing the movement sequences does not impair use of the slider *per se*. Consequently, our experiment provides evidence that imagery for musical loudness change can involve motor processing.

Contrary to expectations, the accurate imagining of melodies was not significantly impaired by the concurrent rehearsal of tone sequences. Perhaps participants ignored the intervening tone sequence, choosing to prioritize mental rehearsal of the melodies. Such a strategy would have been associated with poor performance on the subsequent tone recognition task, yet accuracy was better than chance [$t(31) = 5.32$, $p < 0.001$]. In addition, the verbal sequence did not impair melody recognition, a finding which is at odds with the results of Keller et al. (1995). Indeed, no differences in melody recognition were observed between the different distractor conditions.

The retention of letter sequences was significantly higher than the retention of other distractor sequences during both the loudness change scale and melody tasks. Yet this superior letter recognition did not come at the price of memory for the loudness change scale or melodies. The letters were presented visually, but they were selected in the knowledge that they might be encoded phonologically and rehearsed as an acoustic image or by subvocal rehearsal. An absence of interference from the letters in the melody recognition task might suggest a visual rehearsal strategy, while a lack of significant interference in the loudness change scale recall task might point to a similar approach, with the interesting corollary that if letter sequences were rehearsed visually, the successful imagining of loudness change scales must be achieved as a motor or auditory image, and not as a visual image of crescendo and decrescendo markings.

The finding that imagery for musical changes in loudness is disrupted by the concurrent rehearsal of a movement goes some way to answering the question of whether patterns of musical loudness are best described as a verbal, auditory, or motor representation. To be added to the list of potential modalities is vision, given that the presentation of the conductor gestures was visual, and participants might have been translating the changes in loudness that they heard into a visual code of what are called “hair pins” (score annotations to indicate crescendi and decrescendi). The absence of an impairment from rehearsing tones does not seem to suggest an exclusively auditory image. Similarly, the lack of a statistically significant impairment from rehearsing a letter sequence does not point to a uniquely verbal labeling of loudness change information such as “up,” “down.” The most likely scenario is that a balance of representation modality was involved. Such a view is consistent with current accounts of working memory that emphasize interference from process rather than content; these accounts recognize the influence of task demands, task relevant, and task irrelevant information, instructions, and context on performance (e.g., Marsh et al., 2009).

Working memory involves simultaneous short-term storage and processing of information (Oberauer, 2009) and is limited to three to five meaningful items in adults (Cowan, 2010). Individual differences in working memory capacity are thought to relate to differences in maintenance and retrieval. More specifically, working memory limitations arise from differences in both ability to actively maintain information and ability to retrieve task relevant information in the presence of highly interfering or irrelevant information (Unsworth and Engle, 2007). Research also points to considerable variation in imagery abilities (Mast and Kosslyn, 2002; Keller and Koch, 2006), and the current study is consistent with this, as models of performance on both types of imagery task (loudness change and melody) were improved by accounting for the variability introduced by individual participants.

Participants in this study were able to read music, suggesting at least a minimal amount of musical training. Not only has a link been established between auditory imagery abilities and musical training (Aleman et al., 2000), this has been extended to the particular context of action-effect anticipation (Keller and Koch, 2008). This suggests that a tight sensorimotor coupling results from extensively rehearsed associations between an action and its consequent sound. In the current study, the participants were not required to physically produce the test stimuli, and so had not explicitly learned an association between movement and melody or loudness change scale items. Nevertheless, musical experience has been found to enhance action-effect coupling quite broadly (Keller and Koch, 2008), and in the current experiment this might have reinforced the ability of participants to imagine motor and auditory components of loudness change. Indeed, the optimal model of performance in the imagery for loudness change task included years of musical training. A correlation between years of musical training and accuracy in the melody imagery task was also found.

Audio-motor coupling has been argued to be strong for musicians, and Baumann et al. (2007) have suggested that the activation of both auditory and motor areas of the brain while listening or playing even when participants attend to a distractor task is evidence for direct and automatic connections between auditory and motor areas in music. In an interview study of experienced musicians, Holmes (2005) found that motor imagery was a significant part of learning and memorizing music for performance. It is interesting that years of training was a significant factor, which amounts to an index of performance experience, while OMSI score was not, which is a composite measure of musical sophistication, taking into account listening and compositional experience. While the current study relied on musically literate participants, it would be interesting to investigate the modalities involved in imaging musical loudness change for a wider population. Some evidence points to a common metaphorical association of movement with musical loudness change (Eitan and Granot, 2006).

In conclusion, we have presented behavioral evidence for motor processing in the imagining of musical changes in loudness. Although concurrent verbal and auditory distractor tasks did not significantly impair participants’ ability to imagine loudness change stimuli, we should not conclude that a uniquely motor representation drives imagery for musical loudness change. These verbal and auditory distractor tasks failed to impair performance on a melody imagery task, in spite of previous research to suggest that melodic material is rehearsed as an auditory image. Future work is needed to determine how rehearsing the tonal and letter sequences employed in the current study should have impaired auditory and verbal imagery for musical stimuli. Individual differences in imagery ability were evident, and it remains to be seen whether these individual differences are reflected in the processing modalities preferred when imagining musical stimuli. Our ongoing research is studying the use of mental imagery for loudness change by expert musicians in performance. The current experiment has provided an effective interference task for loudness change imagery in the guise of conductor gestures, allowing us to examine the strategies used by musicians when they cannot consciously plan (imagine) their expressive intentions.

ACKNOWLEDGMENTS

Our thanks to Stephen Fazio for assistance with editing video clips, Yvonne Leung for data collection, and to members of the Amateur Chamber Music Society for their generous help collecting data.

REFERENCES

Aleman, A., Nieuwenstein, M. R., Böcker, K. B. E., and De Haan, E. H. F. (2000). Music training and mental imagery ability. *Neuropsychologia* 38, 1664–1668.

AMEB. (2002). *Aural Tests: Graded Exercises in Aural Skills.* Melbourne: Allans.

Bailes, F. (2009). “Translating the musical image: case studies of expert musicians,” in *Sounds in Translation: Intersections of Music, Technology and Society*, eds A. Chan and A. Noble (Canberra: ANU E-Press), 41–59.

Bailes, F., and Bishop, L. (2012). “Musical imagery in the creative process,” in *The Act of Musical Composition: Studies in the Creative Process*, ed. D. Collins (Ashgate).

Baumann, S., Koeneke, S., Schmidt, C. F., Meyer, M., Lutz, K., and Jancke, L. (2007). A network for audio-motor coordination in skilled pianists and non-musicians. *Brain Res.* 1161, 65–78.

Berthoz, A. (1996). The role of inhibition in the hierarchical gating of executed and imagined movements. *Brain Res. Cogn. Brain Res.* 3, 101–113.

Berz, W. L. (1995). Working memory in music: a theoretical model. *Music Percept.* 12, 353–364.

Brodsky, W., Henik, A., Rubinstein, B., and Zorman, M. (2003). Auditory imagery from musical notation in expert musicians. *Percept. Psychophys.* 65, 602–612.

Covington, K. (2005). The mind’s ear: I hear music and no one is performing. *Coll. Music Symp.* 45, 25–41.

Cowan, N. (2010). The magical mystery four: how is working memory capacity limited, and why? *Curr. Dir. Psychol. Sci.* 19, 51–57.

Dean, R. T., and Bailes, F. (2008). Is there a ‘rise-fall temporal archetype’ of intensity in electroacoustic music? *Can. Acoust.* 36, 112–113.

Dean, R. T., and Bailes, F. (2010). A rise-fall temporal asymmetry of intensity in composed and improvised electroacoustic music. *Organised Sound* 15, 147–158.

Deutsch, D. (1970). Tones and numbers: specificity of interference in immediate memory. *Science* 168, 1604–1605.

Dowling, W. J. (1978). Scale and contour: two components of a theory of memory for melodies. *Psychol. Rev.* 85, 341–354.

Eitan, Z., and Granot, R. Y. (2006). How music moves: musical parameters and listeners’ images of motion. *Music Percept.* 23, 221–247.

Finney, S., and Palmer, C. (2003). Auditory feedback and memory for music performance: sound evidence for an encoding effect. *Mem. Cognit.* 31, 51–64.

Giorgino, T. (2009). Computing and visualizing dynamic time warping alignments in R: the dtw package. *J. Stat. Softw.* 31, 1–24.

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

Holmes, P. (2005). Imagination in practice: a study of the integrated roles of interpretation, imagery and technique in the learning and memorisation processes of two experienced solo performers. *Br. J. Music Educ.* 22, 217–235.

Hubbard, T. L. (2010). Auditory imagery: empirical findings. *Psychol. Bull.* 136, 302–329.

Intons-Peterson, M. J. (1992). “Components of auditory imagery,” in *Auditory Imagery*, ed. D. Reisberg (Hillsdale, NJ: Lawrence Erlbaum Associates), 45–71.

Kalakoski, V. (2001). “Musical imagery and working memory,” in *Musical Imagery*, eds R. I. Gødoy and H. Jørgensen (Lisse: Swets & Zeitlinger), 43–56.

Keller, P. E., and Koch, I. (2006). Exogenous and endogenous response priming with auditory stimuli. *Adv. Cognit. Psychol.* 2, 269–276.

Keller, P. E., and Koch, I. (2008). Action planning in sequential skills: relations to music performance. *Q. J. Exp. Psychol.* 61, 275–291.

Keller, T. A., Cowan, N., and Saults, J. S. (1995). Can auditory memory for tone pitch be rehearsed? *J. Exp. Psychol. Learn. Mem. Cogn.* 21, 635–645.

Marsh, J. E., Hughes, R. W., and Jones, D. M. (2009). Interference by process, not content, determines semantic auditory distraction. *Cognition* 110, 23–38.

Mast, F. W., and Kosslyn, S. M. (2002). Visual mental images can be ambiguous: insights from individual differences in spatial transformation abilities. *Cognition* 86, 57–70.

Mikumo, M. (1994). Motor encoding strategy for pitches and melodies. *Music Percept.* 12, 175–197.

Oberauer, K. (2009). Interference between storage and processing in working memory: feature overwriting, not similarity-based competition. *Mem. Cognit.* 37, 346–357.

Ollen, J. (2006). *A Criterion-related Validity Test of Selected Indicators of Musical Sophistication Using Expert Ratings.* Ph.D., Ohio State University, Columbus.

Pearson, J., Clifford, C. W. G., and Tong, F. (2008). The functional impact of mental imagery on conscious perception. *Curr. Biol.* 18, 982–986.

Pitt, M. A., and Crowder, R. G. (1992). The role of spectral and dynamic cues in imagery for musical timbre. *J. Exp. Psychol. Hum. Percept. Perform.* 18, 728–238.

Rossi-Arnaud, C., Cortese, A., and Cestari, V. (2004). Memory span for movement configurations: the effects of concurrent verbal, motor and visual interference. *Curr. Psychol. Cognit.* 22, 335–349.

Sacks, O. (2007). *Musicophilia: Tales of Music and the Brain.* London: Picador.

Smith, J. D., Reisberg, D., and Wilson, M. (1992). “Subvocalization and auditory imagery: interactions between the inner ear and inner voice,” in *Auditory Imagery*, ed. D. Reisberg (Hillsdale, NJ: Lawrence Erlbaum Association), 95–117.

Smith, J. D., Wilson, M., and Reisberg, D. (1995). The role of subvocalization in auditory imagery. *Neuropsychologia* 33, 1433–1454.

Smyth, M. M., and Pendleton, L. R. (1989). Working memory for movements. *Q. J. Exp. Psychol.* 41A, 235–250.

Smyth, M. M., and Waller, A. (1998). Movement imagery in rock climbing: patterns of interference from visual, spatial and kinaesthetic secondary tasks. *Appl. Cogn. Psychol.* 12, 145–157.

Todd, N. P. M. (1992). The dynamics of dynamics: a model of musical expression. *J. Acoust. Soc. Am.* 91, 3540–3550.

Trusheim, W. H. (1991). Audiation and mental imagery: implications for artistic performance. *Q. J. Music Teach. Learn.* 2, 138–147.

Unsworth, N., and Engle, R. W. (2007). The nature of individual differences in working memory capacity: active maintenance in primary memory and controlled search from secondary memory. *Psychol. Rev.* 114, 104–132.

Wachowicz, F., Stevens, C. J., and Byron, T. P. (2011). Effects of balance cues and experience on serial recall of human movement. *Dance Res.* 29, 450–468.

Wiens, P. W. (2002). “Expressive Conducting.” [Audiovisual material].

Williamson, V. J., Baddeley, A. D., and Hitch, G. J. (2010). Musicians’ and nonmusicians’ short-term memory for verbal and musical sequences: comparing phonological similarity and pitch proximity. *Mem. Cognit.* 38, 163–175.

Wood, J. N. (2007). Visual working memory for observed actions. *J. Exp. Psychol. Gen.* 136, 639–652.

Wu, J., Yu, Z., Mai, X., Wei, J., and Luo, Y. (2010). Pitch and loudness information encoded in auditory imagery as revealed by event-related potentials. *Psychophysiology* 48, 1–5.

Received: 30 April 2012; accepted: 06 November 2012; published online: 03 December 2012.

Citation: Bailes F, Bishop L, Stevens CJ and Dean RT (2012) Mental imagery for musical changes in loudness. Front. Psychology ***3****:525. doi: 10.3389/fpsyg.2012.00525*

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

frontiers in
HUMAN NEUROSCIENCE

REVIEW ARTICLE
published: 05 September 2012
doi: 10.3389/fnhum.2012.00247

Imagining is not doing but involves specific motor commands: a review of experimental data related to motor inhibition

Aymeric Guillot[1,2]*, Franck Di Rienzo[1], Tadhg MacIntyre[3], Aidan Moran[4] and Christian Collet[1]

[1] Centre de Recherche et d'Innovation sur le Sport (EA 647), équipe Performance Motrice, Mentale et du Matériel, Université de Lyon, Université Claude Bernard Lyon 1, Villeurbanne, France
[2] Institut Universitaire de France, Paris, France
[3] Physical Education and Sport Sciences Department, University of Limerick, Limerick, Ireland
[4] School of Psychology, University College Dublin, Dublin, Ireland

Edited by:
Joel Pearson, The University of New South Wale, Australia

Reviewed by:
Stephanie D. Preston, University of Michigan, USA
Yoshikazu Ugawa, Fukushima Medical Universtiy, Japan

***Correspondence:**
Aymeric Guillot, Centre de Recherche et d'Innovation sur le Sport, Université Claude Bernard Lyon 1, 27-29 Boulevard du 11 Novembre 1918, 69622 Villeurbanne Cedex, France.
e-mail: aymeric.guillot@univ-lyon1.fr

There is now compelling evidence that motor imagery (MI) and actual movement share common neural substrate. However, the question of how MI inhibits the transmission of motor commands into the efferent pathways in order to prevent any movement is largely unresolved. Similarly, little is known about the nature of the electromyographic activity that is apparent during MI. In addressing these gaps in the literature, the present paper argues that MI includes motor execution commands for muscle contractions which are blocked at some level of the motor system by inhibitory mechanisms. We first assemble data from neuroimaging studies that demonstrate that the neural networks mediating MI and motor performance are not totally overlapping, thereby highlighting potential differences between MI and actual motor execution. We then review MI data indicating the presence of subliminal muscular activity reflecting the intrinsic characteristics of the motor command as well as increased corticomotor excitability. The third section not only considers the inhibitory mechanisms involved during MI but also examines how the brain resolves the problem of issuing the motor command for action while supervising motor inhibition when people engage in voluntary movement during MI. The last part of the paper draws on imagery research in clinical contexts to suggest that some patients move while imagining an action, although they are not aware of such movements. In particular, experimental data from amputees as well as from patients with Parkinson's disease are discussed. We also review recent studies based on comparing brain activity in tetraplegic patients with that from healthy matched controls that provide insights into inhibitory processes during MI. We conclude by arguing that based on available evidence, a multifactorial explanation of motor inhibition during MI is warranted.

Keywords: motor imagery, motor command inhibition, motor performance, mental processes, electromyography, sensorimotor control

INTRODUCTION

One of the most remarkable capacities of the mind is its ability to simulate sensations, movements, and other types of experience. In most occasions, mentally imagining is like perceiving in the absence of the corresponding sensory information. In other words, imagining involves "seeing" with the "mind's eye," "hearing" with the "mind's ear," and so on for each sensory modality (Kosslyn, 2010). Accordingly, mental imagery is a multimodal construct which consists of either recalling previously perceived images or feelings, or envisaging forthcoming events. Within this construct, "motor imagery" (MI) refers to the mental representation of an action without engaging in its actual execution. MI involves an integrated covert simulation of physical movement, and may be defined as a dynamic mental state during which the representation of a given motor act is rehearsed in working memory without any overt motor output (Decety and Grezes, 1999; Collet and Guillot, 2010).

The vast majority of experimental investigations dealing with the MI experience primarily focused on visual and kinesthetic imagery (KI). While visual imagery (VI) refers as to the visualization of an action, KI involves the sensations of how it feels to perform an action, including the force and effort perceived during movement and balance (Callow and Waters, 2005), hence suggesting to consider the body as a generator of forces (Jeannerod, 1994). Interestingly, previous data showed that KI was close to motor execution, with an extensive overlap of the corresponding neural networks (Solodkin et al., 2004). Other researchers introduced the concept of imagery perspective and further distinguished between first- and third-person VI perspectives. During the first-person perspective, performers visualize the action as how would happen in the real-life situation, while in the third-person perspective, they imagine, as spectators, the action that somebody is performing, regardless of the agency of the movement (i.e., whether they, "see" themselves or others). Although there is some confusion in the usage of these terms in many studies (Moran et al., 2012), a considerable amount of experimental research suggests that MI is a valuable and cost-effective technique to improve motor performance and to enhance motor recovery (see reviews by Driskell et al., 1994; de Vries and Mulder, 2007; Guillot and Collet, 2008; Munzert et al., 2009).

Despite the preceding evidence on the efficacy of MI, several unresolved issues have emerged with regard to the neural underpinnings of this construct. More precisely, little agreement exists among imagery researchers as to the extent to which the neural substrates of MI overlap with those of actual practice. Also, little evidence exists on the question of whether or not motor commands are inhibited during imagined movements (Kasess et al., 2008). Nevertheless, certain trends exist in the relevant literature. For example, the pattern of electromyographic (EMG) activity during MI generally supports the hypothesis of residual muscle activity which might originate from an incomplete inhibition of the motor command (Jeannerod, 1994). Similarly, transcranial magnetic stimulation (TMS) experiments also support this assumption by delineating the features of corticospinal facilitation during MI (Stinear, 2010). Finally, clinical studies show that some patients with specific brain damage fail to inhibit the motor action associated with its mental representation, and thus fully "execute the imagined action," hence highlighting uninhibited movements during mental rehearsal. Data in amputees, patients with spinal cord injury (SCI) and Parkinson's disease (PD) further contribute to theoretical understanding of how the motor command is inhibited during MI.

Against this background, the present paper aims to examine these issues and provide deeper evidence that MI includes motor commands for muscle contractions, which are blocked at some level of the motor system by inhibitory mechanisms. On some occasions (e.g., in sport settings), individuals may retain the potential to move while imagining an action, e.g., miming unambiguously some parts of movement execution while rehearsing the imagined movement. These situations raise an important but neglected question: How does the brain resolve the paradox whereby it is required to issue the motor command needed for action when MI is performed, while concurrently issuing an inhibitory command when the person is moving during MI?

NEURAL CORRELATES OF MOTOR IMAGERY

Understanding the neural correlates of motor performance and MI has been an important purpose of brain studies since the advent of neuroimaging techniques. Considerable experimental evidence has accumulated to suggest that movement execution and MI share substantial overlap of active brain regions (for review, see Guillot et al., 2012). Such apparent functional equivalence supports the hypothesis that MI draws on the similar neural networks that are used in actual perception and motor control (Jeannerod, 1994; Grezes and Decety, 2001; Holmes and Collins, 2001). Moreover, MI can also activate neural circuits used during tasks investigating memory and emotion (Kosslyn et al., 2001). As we will see, however, the neural networks underlying these behavior are not strictly identical. This is because when performing MI, participants are aware that movement will not be performed, and therefore that motor commands must be inhibited.

Neuroimaging studies provided preliminary evidence that motor-related areas of the brain (e.g., the ventral and dorsal parts of the premotor cortex, as well as the supplementary motor area – SMA) and subcortical areas including the cerebellum and the basal ganglia, are active during MI of both simple and complex movements (e.g., Lotze and Halsband, 2006; Guillot et al., 2008; Munzert et al., 2009). Furthermore, research suggests that MI activates a subset of areas required for movement execution (Macuga and Frey, 2012), thus leading to a partial overlap in the corresponding neural networks. The contribution of the contralateral primary motor cortex (cM1) to imagined actions is more controversial, however (for reviews, see Lotze and Halsband, 2006; Munzert et al., 2009; Lotze and Zentgraf, 2010). Whereas some researchers did not report cM1 activations during MI (e.g., Gerardin et al., 2000; Hanakawa et al., 2008), others found fleeting involvement (Dechent et al., 2004) or significant activation (Lotze et al., 1999b; Porro et al., 2000; Solodkin et al., 2004; Guillot et al., 2008; Sharma et al., 2008). Such discrepancies may be due to methodological differences and difficulties in monitoring compliance with MI instructions (Sharma et al., 2006). Interestingly, Ehrsson et al. (2003) showed that the content of MI was reflected in the pattern of motor cortical activation, as MI of hand, foot, and tongue movements specifically activated the corresponding hand, foot, and tongue sections of cM1. Additional evidence indicates that activation of cM1 might be differentially influenced by MI instructions, MI ability, and motor expertise (Lotze and Zentgraf, 2010). Taken together, the bulk of neuroimaging studies suggest that cM1 is activated during MI – but more weakly than during actual movement. Interestingly, Kasess et al. (2008) reported that SMA may substantially contribute to inhibit activity of cM1 during MI.

Activation of parietal areas including the inferior and superior parietal lobules, as well as the precuneus, was also frequently reported during MI (Gerardin et al., 2000; Hanakawa et al., 2003; Guillot et al., 2009; Munzert et al., 2009). Experimental studies in patients with parietal lesions further support that these structures are critically involved in the generation and guidance of mental images, including the ability to achieve temporal congruence between MI and motor performance (Sirigu et al., 1996; Malouin et al., 2004).

The patterns of neural activity underlying imagery types (e.g., VI vs. KI) and imagery perspectives (first-person vs. third-person imagery) are partially mediated through separate neural systems. For instance, Solodkin et al. (2004) investigated neural networks associated with physical execution, VI, and KI of hand movements. Although some shared neural substrates were evident between these processes, differences were found in the inputs received from the superior parietal lobule. Specifically, inputs from SMA to cM1 were lower to those observed during motor execution (Gao et al., 2011). In a single group of participants with high MI abilities, Guillot et al. (2009) showed that VI activated predominantly the visual pathways including the occipital regions and the precuneus, whereas KI involved mainly motor-associated structures and the inferior parietal lobule.

Neuroimaging studies have also shown that the neural networks underlying MI differ as a function of both individual expertise level and imagery ability. For instance, an inverse relationship between the pattern of brain activity and expertise level has been reported (Ross et al., 2003; Milton et al., 2007) with decreased activations in the SMA, the cerebellum, and the basal ganglia. More recent work by Guillot et al. (2008) confirms that the neural networks mediating MI partially differ as a function of imagery ability. Specifically, whereas strong imagers tended to show activation in the parietal and ventrolateral premotor regions, weaker imagers tended to recruit the cerebellum, the orbito-frontal, and posterior cingulate cortices. In both cases, however, dynamic brain changes were found to become more refined and circumscribed with imagery/physical practice (PP) – a trend which was also evident during the learning process of motor tasks.

In summary, neuroimaging research clearly demonstrates that MI and motor performance of the same task share certain neural substrates but the overlap is incomplete. Furthermore, data challenge the assumption that neural activity is described as being in all of the same areas as execution, albeit to a lesser extent. As expected, some areas that are active during motor performance are not involved during MI. Although less common, the converse is also true, with some regions being more strongly and/or selectively activated during MI compared to actual execution of the same movement (e.g., the pre-SMA – e.g., Hanakawa et al., 2008). From the preceding evidence, we conclude that the main difference between MI and motor performance is probably that MI involves the *inhibition* of some motor commands triggering movements – although the neural *level* at which the motor command is stopped has not yet been clearly identified. To address this latter issue, we believe that EMG recordings may provide a reliable means to investigate whether or not the brain activation recorded during MI actually originate from mental representation – as opposed to the potential motor activity that could have accompanied the task. In a similar vein, TMS can be used to explore the degree to which MI modulates both corticomotor excitability and intracortical inhibition.

MUSCLE AND TMS ACTIVITY DURING IMAGINED ACTIONS

EMG CORRELATES OF MOTOR IMAGERY

A great amount of experimental data has been collected on the physiological operations involved during MI (**Table 1**) – notably the peripheral muscular activity which may occur during the mental representation of an action (Guillot et al., 2010). Since the pioneering work of Jacobson (1930, 1932), who provided the first scientific evidence that MI of bending the arm produced small contractions of the flexor arm muscles, debate has existed on whether or not MI is accompanied by subliminal muscle contractions.

Muscle quiescence during MI has been reported in many experimental studies (e.g., Yue and Cole, 1992; Decety et al., 1993; Lotze et al., 1999b; Mulder et al., 2005). Interestingly, the lack of EMG activity during MI was sometimes considered a precondition prior to engaging in MI practice (Michelon et al., 2006). On some occasions, EMG data were even recorded during scanning sessions *per se*, to demonstrate that variations of cerebral blood flow were directly related to the mental work and not to any concomitant movement (e.g., Gerardin et al., 2000; Jackson et al., 2003; Lotze et al., 2003; Hanakawa et al., 2008). In sum, these results seem to suggest that MI involves the intention of *not* executing the movement, with strong inhibitory mechanisms blocking the motor command.

Conversely, similar EMG activity has been observed during overt motor execution and MI conditions, with a reduced magnitude in the simulated action (e.g., Jowdy and Harris, 1990; Hashimoto and Rothwell, 1999). Gandevia et al. (1997) further demonstrated that not only MI did activate the alpha motoneurons, but that the skeletomotor discharge was also accompanied by recruitment of spindle afferents when the covert contraction was sufficiently strong. Furthermore, there is evidence of increasing EMG activity accompanying imagined mental effort (Boschker, 2001; Slade et al., 2002; Guillot et al., 2007). Finally, EMG activity has been observed not only in agonistic muscles, but also in antagonistic muscles, as a function of both the weight to be lifted (Bakker et al., 1996) and the muscle contraction type (Guillot et al., 2007). In this latter study, the authors found that the subliminal muscle responses during MI of concentric, isometric, and eccentric contractions typically mirrored the configuration of the EMG activity recorded during actual practice. These data support the hypothesis that muscle activity recorded during MI is not a general tonic activation but reflects the content of the specific motor command of the movement that is mentally rehearsed. According to Jeannerod (1994), an incomplete inhibition of the motor command could provide a valid explanation for these muscle discharges. So, it seems that imagined movements produce a qualitatively similar, but quantitatively smaller, drive to muscles compared with actual motor execution, thereby suggesting that a small part of the motor command is actually sent to the effectors during MI. Interestingly, Solodkin et al. (2004) argued that both the supplementary motor area and the lateral premotor cortex might also play a role in increasing muscle tone during MI – especially during KI. This speculation is plausible given that these brain regions have direct projections to spinal cord through the internal capsule, adjacent to the well-known corticospinal path originating from cM1 (Luppino et al., 1994; Morecraft et al., 2002).

Inconsistencies in the reports of concomitant EMG activity in the muscles participating in the movement during MI might be explained by differences in the experimental designs, as well as by the nature of the EMG recordings (Guillot et al., 2010). For instance, EMG activity might not be discernible due to the use of surface EMG electrodes and intramuscular electrodes should ideally be preferred, although intrusive and thus rarely used in MI experiments (except in the case of Gandevia et al., 1997). Analogously, the effect of the muscle contraction type, the intensity of the mental effort, and the intrinsic nature of MI may also contribute to understand why EMG activity was not systematically reported. Finally, with few exceptions, studies reporting a lack of EMG activity primarily investigated laboratory movements, whereas those experiments providing evidence of a muscle activity during imagery included more goal-related movements (e.g., skills in sport). Experimental data reporting muscle activity only in a part of the tested sample (Li et al., 2004a; Dickstein et al., 2005) lend support to the fact that muscle activity was not systematically discernible due to such confounding factors. Interestingly, the pattern of muscle activation has never been found to match

Table 1 | **Studies investigating the EMG activity during motor imagery.**

Study	Number of participants	Motor task	Main findings
LACK OF MUSCLE ACTIVITY DURING MOTOR IMAGERY			
Decety et al. (1993)	$n=6$	Leg contraction to press and release a loaded footplate	No change in phosphocreatine concentration or in pH during motor imagery
Demougeot and Papaxanthis (2011)	$n=17$	Vertical arm movements	No arm muscle activation during motor imagery
Gentili et al. (2006)	$n=40$	Pointing arm movement	No EMG activity during motor imagery
Gerardin et al. (2000)	$n=8$	Auditory-cued hand movements	Surface EMG did not detect any muscle activity during motor imagery
Gueugneau et al. (2008)	$n=9$	Pointing arm movement	No EMG activity during motor imagery
Hanakawa et al. (2008)	$n=13$	Finger tapping sequence	Surface EMG was monitored during fMRI recordings to confirm the absence of muscle activity during motor imagery
Jackson et al. (2003)	$n=9$	Foot sequence task	No significant difference in the EMG signal between imagery and baseline conditions, showing that the patterns of cerebral activation during fMRI recordings are not due to movements
Kleber et al. (2007)	$n=16$	Singing of an Italian aria	No difference between baseline and imagined singing
Lafleur et al. (2002)	$n=9$	Foot sequence task	EMG recordings showed no change in muscle activity during scans compared to baseline levels
Lim et al. (2006)	$n=13$	Arm movement	No EMG activity during motor imagery
Lotze et al. (1999b)	$n=10$	Making a fist	Low EMG activity, which did not differ from the baseline, was a precondition before fMRI recordings
Lotze et al. (2003)	$n=16$	Performance of Mozart's violin concerto in G major	No observable differences between motor imagery and rest
Mulder et al. (2004)	$n=37$	Abduction of the big toe	No EMG activity during motor imagery
Mulder et al. (2005)	$n=31$	Squat movements with additional weights	EMG activity recorded during motor imagery did not differ from baseline
Naito et al. (2002)	$n=10$	Palmar flexion and dorsiflexion of the wrist	No EMG activity in the motor imagery condition
Personnier et al. (2008)	$n=28$	Arm movements in the sagittal plane	Muscle activation patterns are very similar between motor imagery and rest conditions
Ranganathan et al. (2004)	$n=30$	Isometric little finger abduction and elbow flexion	Muscle activity during motor imagery was near zero
Roosink and Zijdewind (2010)	$n=20$	Finger tapping sequence	No EMG activity during motor imagery
Shick (1970)	$n=10$	Volleyball serve skill	No EMG activity during motor imagery
Yahagi et al. (1996)	$n=7$	Wrist flexion	No EMG activity during motor imagery
Yue and Cole (1992)	$n=30$	Isometric little finger abduction	No EMG activity during motor imagery

(Continued)

Table 1 | Continued

Study	Number of participants	Motor task	Main findings
Zijdewind et al. (2003)	$n = 29$	Ankle plantar-flexion	No EMG activity during motor imagery in the majority of the participants. When little EMG activity was recorded, participants were asked to concentrate until being able imagining the movement without muscle activation
MUSCLE ACTIVITY DURING MOTOR IMAGERY			
Bird (1984)	$n = 5$	Motor imagery of a past athletic event (including riding, rowing, swimming, water skiing, and basketball)	The EMG configuration during motor imagery mirrored that observed during actual practice
Bonnet et al. (1997)	$n = 26$	Foot pressure on a pedal	EMG activity weakly increased during motor imagery
Boschker (2001) Bakker et al. (1996)	$n = 39$	Arm lifting movements (biceps curls)	Significant EMG activity is recorded in the muscles contributing to the contraction. Greater muscle activity in the active than in the passive arm, and greater biceps activity when imagining lifting a heavy compared to a light weight
Dickstein et al. (2005)	$n = 15$	Rising on tiptoes	EMG activity was recorded in six participants in at least one of the target muscles
Gandevia et al. (1997)	$n = 12$	Range of simple and complex movements (e.g., flexions/extensions, handwriting, walking, threading a needle…)	Imagery increased background EMG in the involved muscles. In some occasions, spindle discharge also increased
Guillot et al. (2007)	$n = 30$	Biceps dumbbell curls	The magnitude of EMG activity is correlated to the mental effort required to imagine the movement. EMG patterns during imagery of concentric, isometric, and eccentric contractions mirror those observed during actual movements. EMG activity is recorded in agonist, antagonist, synergist, and fixator muscles
Hale (1982)	$n = 48$	Biceps dumbbell curls	Internal imagery perspective produced greater biceps activity than the external imagery perspective
Harris and Robinson (1986)	$n = 36$	Arm lifting	Significant EMG activity is recorded in the muscles contributing to the contraction. Greater EMG activity during the first-person than during the third-person perspective
Hashimoto and Rothwell (1999)	$n = 9$	Wrist flexion and extension	Larger EMG responses in flexor and extensor muscles during imagined flexions and extensions, respectively
Jacobson (1930, 1932)	The number of participants varied among tasks	Biceps dumbbell curls, bending the forearm, sweeping, climbing a rope	EMG activity was recorded in the specific muscle involved with the imagined activity
Jowdy and Harris (1990)	$n = 38$	Juggling task	Increased muscle activity during motor imagery. No effect of the imagery ability on the magnitude of muscle activity
Lebon et al. (2008)	$n = 30$	Biceps dumbbell curls	The median frequency of EMG power spectrum in agonist and antagonist muscles was significantly higher during motor imagery than during baseline
Li et al. (2004a)	$n = 9$	Flexion and extension movements of the fingers	EMG activity was recorded in the finger flexors in four participants
Livesay and Samaras (1998)	$n = 30$	Tightly squeezing a hand-size rubber ball	Increased EMG activity in the dominant forearm

(Continued)

Table 1 | Continued

Study	Number of participants	Motor task	Main findings
Lutz and Linder (2001)	$n = 160$	Dart throwing	Greater biceps EMG activity was recorded when imagery instructions included assertions about behavior, such as motor actions and visceral responses
Shaw (1938)	The number of participants varied among tasks	Range of complex movements (e.g., flexions/extensions, handwriting, walking, threading a needle. . .)	Increased EMG activity during motor imagery was distributed across different muscle groups including those not directly related to the corresponding movement
Slade et al. (2002)	$n = 60$	Biceps dumbbell and manipulandum curls	EMG activity was significantly greater for both curls in the active arm during motor imagery when compared to baseline
Suinn (1980)	$n = 1$	Skiing a downhill race	Recorded muscle patterns were strikingly similar to those observed during actual practice
Wehner et al. (1984)	$n = 27$	Contour tracking arm task	Similar frequency distribution in the power spectrum during actual practice and motor imagery

the usual triphasic sequence generated during actual motor performance (Murphy et al., 2008). Furthermore, although traditional psychoneuromuscular theory (Carpenter, 1894) postulated that muscle activity recorded during MI might generate slight neuromuscular feedback, strong enough to improve subsequent motor performance through priming of the motor pathways, there is no direct evidence that muscle activation during MI is associated with improved motor performance. In other words, research has not yet demonstrated that the increase in muscle activity fully contributes to performance enhancement, but EMG recordings support that the motor command is actually prepared, and then blocked by inhibitory processes, during MI.

CORTICOSPINAL FACILITATION PATTERNS DURING MI

Transcranial magnetic stimulation activates neurons transsynaptically (Rothwell, 1991) and allows the study of corticospinal facilitation (i.e., the level of excitability of the corticomotor pathway). Typically, the motor evoked potential (MEP) elicited by a suprathreshold TMS pulse delivered to cM1 is recorded at the peripheral level, using surface EMG. The development of repetitive pulse TMS protocols has allowed the study of intracortical facilitation and inhibition, and also helps to delineate excitatory and inhibitory interactions between different brain regions mediating motor control (Reis et al., 2008). In this section, we review some TMS studies suggesting that MI involves elaboration of motor command signals at the CNS level. We also discuss the extent to which MI activates the descending somatic motor pathways originating from pyramidal neurons and projecting toward the alpha motoneurons pool. Finally, we consider the recent hypotheses emerging from TMS findings regarding motor inhibition during MI.

CNS activity during MI and PP enables researchers to understand whether MI effectively involves motor commands processing. TMS studies provided converging evidence that MI increases the corticomotor excitability (Stinear, 2010). Excitability changes within motor cortical areas during MI, including reduced intracortical inhibition, are analogous to those observed during motor preparation and execution (Abbruzzese et al., 1999; Kumru et al., 2008). Intracortical inhibitory interneurons are known to play an essential role in the shaping of motor commands (Stinear and Byblow, 2003a). This phenomenon is thought to mirror an analogous motor activity at the cortical level during both tasks. Contrary to PP, intracortical inhibition of cM1 is not entirely removed during MI (Kumru et al., 2008). In general, whilst MI involves improved cortical facilitation and reduced intracortical inhibition, it does so with reduced amplitude compared to PP (Clark et al., 2004; Leonard and Tremblay, 2007). It has been suggested that the CNS manages to keep corticospinal facilitation below the motor threshold for activating the alpha motor neurons pool during MI (Stinear, 2010).

Does cortical activity during MI effectively reflect shaping of motor output – which may require motor inhibition? Firstly, corticospinal facilitation during MI is effector-specific according to MI content (Kasai et al., 1997; Stinear and Byblow, 2003a, 2004). The motor threshold in muscles involved into MI content is lower, while the amplitude of the subsequent MEP is higher (Facchini et al., 2002). By contrast, MEPs elicited in non-involved muscles remain unaffected (Facchini et al., 2002; Stinear and Byblow, 2003b; Quartarone et al., 2005). Furthermore, Leonard and Tremblay (2007) demonstrated that the muscle-specific pattern of corticospinal facilitation during MI was altered in aging populations, thus reflecting individual ability to shape motor output coding for isolated finger movements. Corticospinal facilitation during MI is also graded upon the extent to which the muscle is actually recruited during the corresponding motor performance (Yahagi and Kasai, 1998). Liang et al. (2007) further reported that corticospinal facilitation during MI of wrist flexions mirrored the synergic pattern of muscle activity produced by PP, a key feature of the motor command (see also Kasai et al., 1997, for results suggesting preservation of agonist-antagonist patterns of muscle activations during MI). Likewise, using TMS evoked muscle twitch, Li et al. (2004b) reported that MI of finger flexion preserved the unintended functional coupling in strength response between the digits, described as the "enslaving effect" (Zatsiorsky et al., 2000). Further, Facchini et al. (2002) observed that only MI of contralateral thumb abduction, but not ipsilateral, facilitated MEPs in the contralateral effector. This result supports the idea that MI reproduces the hemispheric specificity with regards to cM1 enrollment during imagination of lateralized movements. Finally, Fadiga et al. (1999) observed that MI of elbow flexion/extension increased MEPs in the biceps brachii merely during the timing portions of MI corresponding to the arm flexion, thus suggesting that corticospinal facilitation during MI matched the temporal features of physical performance.

To summarize, the accumulating evidence appears to suggest that the corticospinal facilitation is highly specific to the motor task. Increased corticomotor excitability during MI may not be a result of a general state of arousal due to execution of cognitive operations (Rossini et al., 1999; Clark et al., 2004; Fourkas et al., 2006), but instead due to the demands of the internal processing of motor output (Stinear, 2010).

As mentioned previously, most features of corticospinal facilitation during MI suggest internal elaboration of neural signals for muscle contractions. It is generally assumed that increased corticospinal facilitation during MI reflects analogous involvement of cM1 between MI and PP (Stinear, 2010). However, such increased corticomotor excitability may also reflect more general changes in the balance between excitatory and inhibitory impulses, which can occur at different stages of the somatic motor system. Understanding how the centrally shaped motor output is inhibited during MI initially requires further analysis of the extent to which CNS excitability changes (due to MI) affect the descending motor pathways. For instance, one may question whether corticospinal facilitation during MI involves the motor system at the spinal level. Yahagi et al. (1996) addressed this issue and observed that whilst MI of wrist flexions facilitated the MEPs in the flexor carpi radialis, no change was recorded in the H-reflex surface EMG traces evoked by electrical stimulation, thus revealing that corticospinal facilitation during MI occurred without any change in spinal excitability. This finding was replicated in several studies investigating changes in H-reflex during MI, in combination with TMS (Kasai et al., 1997; Hashimoto and Rothwell, 1999; Patuzzo et al., 2003). F-waves elicited by peripheral nerve transcutaneous electrical stimulation provide an objective measurement of spinal excitability, without interference from descending neural impulses of cerebral origin (e.g., cerebral spontaneous regulation of spinal reflexes). Rossini et al. (1999) reported a 9.8–14% increase in F-waves amplitude during MI of index and little finger abduction. These robust results challenge previous observations where no change in F-waves was recorded during MI of finger actions (Facchini et al., 2002; Stinear and Byblow, 2003a; Stinear et al., 2006). However, in the two movements investigated by Rossini et al. (1999), only MI of finger abduction elicited a slight increase by 5.9% in the TMS evoked MEP with no change recorded in the MEP latencies when compared to a non-motor mental activity (mental arithmetic). Rossini et al. (1999) stated that while MI may have increased spinal motoneuronal excitability, corticospinal facilitation during MI primarily reflected changes of cortical origin. Consequently, TMS data suggest the analogous involvement of cM1 into motor command processing during both MI and PP. Both tasks elicit excitability changes at the cortical level (Kasai et al., 1997; Abbruzzese et al., 1999; Patuzzo et al., 2003; Stinear and Byblow, 2004), whereas there is a paucity of robust TMS evidences of excitability changes at the spinal level. Consequently, TMS results indicate that inhibition during MI might intervene during the early stages of motor processing. Several neuroimaging findings support this contention, revealing that specific cortical and subcortical sites could contribute to prevent overt motor processing during MI (Lotze et al., 1999b; Kasess et al., 2008).

Challenges to these TMS-based accounts of motor inhibition during MI come from some EMG studies. For example, Bonnet et al. (1997) reported the sharp increase of both H- and T-reflexes during MI of strong foot pressure above a pedal. In this study, T-reflexes displayed a highly specific pattern of facilitation (i.e., lateralized and graded depending on the stimulated movement) which was not observed in H-reflexes facilitation. Bonnet et al. (1997) argued that MI elicited both spinal and spindle activation in the task-relevant corresponding effectors. This finding was replicated in studies reporting increased H-reflexes excitability during MI (Hale et al., 2003). As mentioned above, Gandevia et al. (1997) reported increased activity from spindle afferents using microneurographic recordings from the relevant muscles. The authors concluded that MI recruited both motor units and afferent spindles. Their results further demonstrated that, in some cases, the motor commands built up during MI might reach the muscle level and elicit neural feedback from muscle receptors. Gandevia et al. (1997) therefore stated that MI may consist of "*unintentional performance of (. . .) planned motor task*," hence suggesting that somatic activity during MI might account for the observed effects of MI training on motor performance, through reinforcement of motor output conduction throughout the neuromuscular system (Gandevia, 1999).

Experimental studies also support the central elaboration of motor commands during MI. For example, both EMG and TMS findings support for the role of concurrent indirect information concerning motor inhibition during MI. Firstly, EMG data indicate that a residual motor command can be partially addressed to peripheral effectors during MI. Secondly, TMS findings suggest that the motor system keeps the facilitation of the corticomotor pathways below the motor threshold, in spite of a highly action-specific pattern of arousal. These two ways of understanding motor inhibition during MI could also represent different ways of analyzing a multimodal process: specific interactions between cerebral regions could result in the transmission of a residual motor command toward the descending volleys, whilst interactions between cerebral sites and/or spinal influences could keep corticospinal excitability below the motor threshold. For instance, a recent TMS finding asserts that the ipsilateral inferior parietal lobe might exert an inhibitory influence on cM1 during MI (Lebon et al., 2012).

MOTOR INHIBITION

Earlier in this paper, we showed that motor performance and MI are mediated by distinct neural networks, despite an extensive overlap between KI and PP. In particular, while mental operations of motor planning and programming are actually performed during MI, motor commands must be inhibited before being sent to peripheral effectors within the descending pathways. This inhibition process aims at preventing the performer from engaging in any movement during mental rehearsal. However, we previously underlined significant differences in cerebral activations elicited by MI and actual movement (Hanakawa et al., 2008), some of which being probably responsible of the inhibition of motor commands. Interestingly, Schwoebel et al. (2002) described unique behavior from a patient with bilateral parietal lesions. When imagining hand movements, the patient simultaneously executed the imagined motor act but without being aware of the movements. Surprisingly, these movements were also significantly more accurate than volitional movements. The findings from this clinical case study are consistent with previous accounts suggesting that MI may normally involve the inhibition of movements.

Most studies dealing with MI generally give information about the inhibition of motor commands by providing EMG recordings. This is a reliable means to ensure that brain activation recorded during MI actually comes from mental representation and not from potential motor activity that could have accompanied the mental task. Thus, EMG recordings during MI should be comparable to those that occur during rest. However, a common challenge is that across most investigations, MI and rest conditions are rarely accurately compared. Guillot et al. (2007) observed significant increased pattern of EMG activity in all muscles of the arm, forearm, and even shoulder during MI of forearm flexion, when compared to the rest condition, while goniometric data did not reveal any movement. The magnitude of this activation was correlated with the mental effort required to imagine lifting a weight. Indeed, MI of heavy concentric contraction (80% of the best mark) resulted in greater pattern of EMG activity than during MI of light concentric condition (50% of the best mark). The intensity of the imagined contraction was thus paralleled by the magnitude of the subliminal EMG activity, thereby highlighting a close link between the central nervous system and the periphery during MI. Bakker et al. (1996) and Boschker (2001) had previously found that mentally lifting a 9 kg dumbbell resulted in a larger EMG activity than lifting 4 kg 1/2. Jeannerod (1994) and Bonnet et al. (1997) attributed changes in EMG activity during MI to an incomplete inhibition of the motor command. This hypothesis was emphasized by differential muscle activity associated with the contraction type. Interestingly, different types of mentally rehearsed contractions elicited specific changes in EMG activity that closely corresponded to those observed during actual contraction.

The preceding evidence shows that MI might recruit the same movement pattern as the actual motor command, although at subliminal intensity, hence involving the same neural substrate. Thus, EMG activity during MI seems to mirror that observed during actual motor execution. Importantly, this was not a tonic nonspecific activity as the patterns of EMG distinguished among the different types of contraction to the same extent as actual execution would have done.

Arising from the argument so far, two questions need to be addressed. First, how useful is this specific residual motor command? Second, what are the neural substrates of partial inhibition of the motor command? One of the most plausible outcomes is that sensory afferent information provided to the CNS should serve as feedback in the hypothesis of a forthcoming actual movement. Secondly, the cerebellum might be involved in the inhibition of movement execution during MI (Lotze et al., 1999b).

The question of inhibiting movement execution, after the motor commands have been prepared, has often been asked (e.g., see de Jong et al., 1990). Motor inhibition is usually tested with the "Go/No-Go" paradigm. Here, participants are requested to give a motor response when a specific stimulus is presented and to withhold the response occasionally when another stimulus is triggered. Reaction time to the "Go" signal is recorded, thus facilitating the study of how the motor system inhibits the response when the "No-Go" signal is randomly given. Typically, Go/No-Go paradigms elicit a race between response activation and response inhibition processes. de Jong et al. (1990) postulated the existence of two inhibitory mechanisms: inhibition of central activation processes and inhibition of transmission of motor commands from central to peripheral structures. Unfortunately, the way in which these inhibitory processes work may not directly be applied to MI. For example, the processes of motor command inhibition during MI may not work exactly as during those elicited by the Go/No-Go paradigms because, in the latter, participants do not know in advance whether they will have to act or to inhibit action. By contrast, when participants are requested to mentally represent an action, they implicitly know that they will restrict their cerebral activity to covert movement only. Thus, motor command inhibition should be integrated into the process of movement preparation through motor representation. As postulated by de Jong et al. (1990), a possible mechanism for response inhibition (which could be applied to MI), involves the inhibition of central response activation processes. Thus, response initiation might be inhibited by preventing central response activation from reaching the targeted muscles. In this way, the interruption of an already initiated response can be achieved by discontinuing the output from central to peripheral motor structures. This speculation was experimentally attested by large fronto-central positivity when the response was successfully inhibited (de Jong et al., 1990). The inhibitory mechanisms are effective before sending the information, elaborated within the associate cortices, to the primary motor cortex. Especially, this inhibition may originate from the prefrontal cortical areas associated with limbic structures and cingulate cortices (the behavioral inhibition system early postulated by Gray, 1990). This behavioral system has other connections with the parietal associative cortices involved in No-Go performance (Watanabe et al., 2002). However, response inhibition could also come from active mechanisms at different subcortical levels including the spinal cord (Bonnet et al., 1997), the brainstem, and the cerebellum (Lotze et al., 1999b). A particular example is when the programming of a movement is not entirely well-adapted to its expected goal and requires changing one or several parameters, such as movement direction or amplitude. In this case, we do not need to fully inhibit actual performance but only to better adapt the programming of movement to the environmental constraints under which it occurs. Thus, a flexible central inhibitory mechanism may become crucial when selective motor inhibition is required. Many results from neuroimaging research suggest that the right inferior frontal gyrus is integrated within a fronto-basal-ganglia network (Aron et al., 2007), which could intercept the "Go" process and stop the motor responses (Lenartowicz et al., 2011). This function is also consistent with a role in reprogramming of action plans, which may comprise inhibition, and its activity can be triggered through automated, bottom-up processing.

If central inhibition processes do not succeed in preventing central motor outflow, the overt response can be inhibited by preventing the transmission of motor commands to peripheral motor structures. This possibility is consistent with the hypothesis that motor commands could be inhibited at any time (de Jong et al., 1990). In other words, the inhibition of overt movements may still be achieved by means of peripheral mechanisms. By contrast with central inhibitory mechanisms, peripheral inhibitory mechanisms may be useful only when actions have to be inhibited or interrupted unselectively. Brunia (2003) proposed that there are several inhibitory mechanisms at work in the periphery of the motor system, all depending upon activity of local propriospinal interneurons, situated at the same or neighboring segments of the spinal cord as that of the motoneurons of agonist or synergist muscles. Normally, corticospinal fibers contact alpha motoneurons mono-synaptically. However, while this organization works for hands and fingers, it probably does not hold for other body segments. To result in a movement, the influence of propriospinal neurons upon motoneurons has to be excitatory. The intrinsic organization of the spinal cord enables movement production including several inhibitory systems such as the short feedback system from the Renshaw cells or the reciprocal inhibition reflex system. These processes are beyond the scope of the present paper, however. We thus hypothesize that motor inhibition during MI is mainly related to the *central* but not to the peripheral system.

Logan (1983) conducted several studies on the degree to which people inhibit the thoughts that underlie their actions when they inhibit action. Participants were requested to make category and rhyme judgments about words and were given stop signals that required them to inhibit the actions they executed to express their judgments. They pressed one key if the word was a member of the category or rhymed with a target, and pressed another key if the word was not a member of the category or did not rhyme. If a stop signal occurred, they were supposed to inhibit the response. These researchers then presented the materials again to test participants' memory for words whose responses were inhibited, and used recognition memory judgments: they presented both words for which that they had made or not judgments about, and asked them to indicate whether words had been presented before. Subsequently, Logan (1985) used repetition priming to test memory by presenting similar kind of words, and then asking whether response time was faster for old words than for new ones. Both studies revealed that thoughts went on to completion when actions were inhibited, suggesting that mental activity was independent from the motor response. As far as we consider MI, mental activity is directly related to action, thus suggesting two related processes differing from the relationships highlighted by Logan (1983, 1985).

As previously mentioned, de Jong et al. (1990) described two inhibitory mechanisms that could work to withdraw actual motor command: inhibition of central activation and inhibition of transmission of motor commands from central to peripheral structures. If we consider that motor planning and programming are central processes, we can assume that these are common to actual execution and MI. Hence, the difference between these two behaviors would be the existence of an active process of motor command inhibition. Now, the question of how motor command inhibition is neurally implemented remained unresolved. But what exactly does research tell us about motor command inhibition and how can this knowledge be applied to MI? Most experiments on motor command inhibition were conducted using stop signal paradigms, early formalized by Logan and Cowan (1984). The key component of response inhibition depends upon the relative finishing time between Go and No-Go operations. In other words, the Go response is inhibited by the activation of a stop-process. The major difference between inhibition of motor command in the context of action execution vs. MI is mainly related to *uncertainty*. Uncertainty is emphasized in an updated model by Boucher et al. (2007). A Go response may also be inhibited by the preparation of an alternative go response. In this case, response inhibition would depend on the relative finishing time of the primary-task response and the alternative response (Verbruggen and Logan, 2009). In both models of response inhibition, the participant does not know in advance whether he/she will have to withhold the motor command in the Go/No-Go paradigm. When requested to mentally imagine a movement, the participant is clearly aware that no command should be transferred to peripheral effectors when he/she is requested to mentally imagine the action. Inhibition does not rely on the same process in both conditions. In the stop paradigm, response inhibition depends on triggering No-Go signal. Once the Go or No-Go stimulus is perceived, the participant should decide to act or to inhibit action, taking into account the information provided. Although action vs. inhibition could simultaneously be prepared as an alternative response, this is an all or nothing process. There is no such uncertainty during MI. However, the inhibition of movement may be total or partial and may also take even several intermediate degrees. In other words, the participant could nevertheless accompany the mental representation by residual execution, e.g., miming partially some significant steps of execution or movement rhythm. The other main difference is related to the *time course* of these processes. On the basis of event-related brain potentials, EMG recordings, and continuous behavioral response measures, experimental data from de Jong et al. (1990) evidenced that responses could be interrupted at any time. Thus, actual movement is inhibited as early as the stop signal is triggered whereas MI could accept several conditions from no movement at all until residual movements related to actual movement that accompany and facilitate mental representation. This may explain why MI could also keep some elements of motor execution during mental representation. However, and with reference to the casual definition of MI, we should wonder whether we could still call this process "imagery" when associated with residual parts of movements.

Finally, central inhibition processes are well-summarized by Garavan et al. (2002) who postulated two main neural networks mediating inhibition. Right dorsolateral prefrontal and right inferior parietal areas are associated with response inhibition while a region of the cingulate cortex is involved in "difficult" inhibitions. Left dorsolateral prefrontal cortex was activated when subjects adjusted their ongoing behavior in response to an error or to unexpected changes in the environment. With regard to selective inhibitory mechanisms, it is therefore not surprising that residual muscle activity remains observable during MI, through EMG recordings, as described in preceding paragraphs. This is compatible with both hypotheses previously described: On the one hand, motor commands could be inhibited at any time and in different ways. On the other hand, MI could be dependent upon the central process of inhibition only. Overall, we should point out again that inhibition of the actual command, based on Go/No-Go designs, does not exactly correspond to that during MI, mainly because this inhibition is not under the control of decision making under time pressure. Therefore, a specific cerebral organization might control motor commands inhibition during MI, nevertheless sharing most features of the central processes inhibition we previously described.

In summary, the issue of inhibiting motor performance can be explained by two theoretical models. We first assume that MI results in a subliminal activity of the motor system. As postulated by Jeannerod and Frak (1999), the motor system is involved not only in the production of movements, but also in the mental representation of action. The authors extended its function to the process of learning by observation, even until understanding the behavior of others. Therefore, if we consider MI as a subliminal motor command, it will not cause muscle activity and there is no need for active inhibition process. The functional similarity between actual movement and MI comes from the identity of the motor structures that are believed to control them. Thus, the only difference between actual execution and its mental representation would be the degree of mobilization of motor commands: the preparatory phase would be common to actual action, its mental representation, and the consequences of action both in terms of sensations generated and knowledge of result (did the action reach its goal?). In this regard, Macuga and Frey (2012) recently postulated that the neural representations of observed, imagined, and imitated actions were dissociable and hierarchically organized. The differential activity among these three conditions favored an alternative hierarchical model in which these behaviors rely on partially independent mechanisms. This result might challenge the hypothesis of complete similarity between actual movement and MI, and therefore favor the second hypothesis.

More pragmatically, it is easy to experience motor representation in association with movements or sequences of movements, more or less related to the imagined action and supposed to accompany and facilitate MI. One of the most remarkable examples is when observing some elite athletes during the preparation phase, just before competing. For example, some skiers prepare for races by closing their eyes and mentally rehearsing the course that they are about to traverse (Louis et al., 2012). Although they probably do not experience the entire course, some portions are nevertheless mentally rehearsed using symbolic limb movements – which enable them to mime the represented action. These movements of arms and hands symbolize the turns and the timing at which these should be done. Finally, as revealed by Lorey et al. (2009), we are all familiar with pictures of athletes moving while imagining their subsequent performance during pre-performance routines. Theoretically, such phenomena raise an interesting question. Strictly speaking, is it valid to describe MI performed with associated movement as "MI"? Although this question goes beyond the scope of the present paper, it is important to remember that specific subliminal muscle activity is detectable during MI of any given movement (Guillot et al., 2007). We would also point out that the theoretical mechanisms we described above have the potential to explain how inhibition works during MI. In particular, the inhibition process could occur at *every stage* of the represented action: complete inhibition during MI would mean that actual movement is entirely removed from MI (this corresponds to the usual definition which is often given, "*MI is the mental representation of an action without any overt movement*"). The hypothesis of partial motor inhibition could also be invoked, and as previously mentioned, we may combine movements to their mental representation, even if they are only partially outlined.

How does MI affect motor commands? Performing MI might activate somatic and autonomic motor commands differently. From the intention to act, direct voluntary commands are normally transmitted through the pyramidal tract to elicit movement. The process of an incomplete inhibition that accompanies MI may be viewed at organizing peripheral effectors during the preparation phase of the forthcoming actual execution. Duclos et al. (2008) provided evidence of anticipatory changes in patterns of human motoneurons discharge during motor preparation. This may also be observed during MI. Conversely, Bonnet et al.'s (1997) view is that MI should be compared to action, rather than to motor preparation, hence considering MI as the intention to avoid movement execution, although MI might be more closely related to pre-executive processes of a movement than its actual execution itself (Hanakawa et al., 2008). Michelon et al. (2006) claimed that the MI process does not necessarily require a motor simulation which would integrate the mapping of the effector-specific commands required to achieve the movement. This would be a quite different MI as that followed by actual execution. The close relationship (temporal, structural) between MI and actual execution appears to favor an upstream organization of inhibition, implying the behavioral inhibition system. This would also explain why the most automated parts of movement commands are not inhibited, as they are controlled at the subcortical level. Nevertheless, Bonnet et al. (1997) and Jeannerod (2006) stated that the inhibitory mechanisms may also be localized downstream of the motor cortex, perhaps at the spinal cord, or brainstem level. In addition, Lotze et al. (1999b) postulated that the posterior cerebellum might also play a crucial role in the inhibition of the motor command. There are probably several systems and processes of motor inhibition, coordinated at different levels of the central nervous system from the premotor cortex to the spinal cord. The question of a selective inhibition remains to be asked. It could explain the subliminal muscle activity and even the somatic commands addressed to the low levels of the CNS (e.g., controlling postural regulations).

Several experimental data provide evidence of incomplete inhibition of the motor commands addressed to the different effectors. This finding should be used in the field of clinical rehabilitation (whatever the nature of the neurological damage either peripheral or central), MI use is known to benefit to functional recovery (e.g., Braun et al., 2006; Zimmermann-Schlatter et al., 2008). Further research should also investigate the processes of somatic and autonomic motor commands inhibition during MI. So far, two different mechanisms are thought being involved in the inhibition of a voluntary action. The first is related to central programming processes whereas the second is responsible for central motor command transmission from central structures to peripheral effectors (de Jong et al., 1990). As early hypothesized by Jeannerod (1994), the neural commands for muscle contractions may be blocked at some level of the motor system by active inhibitory mechanisms. This purpose is, however, associated with incomplete inhibition of the motor command that would provide a consistent explanation for the recording of muscle activity during MI.

CLINICAL EXPERIMENTS

Earlier in this paper, we reviewed evidence that motor commands are involved during MI before being inhibited. In this section, we discuss the extent to which clinical data obtained from patients with stroke, PD and also from amputees and those with SCI provide further evidence on this issue. Firstly, analyzing central activity during MI in patients suffering from central nervous system disorders reinforces the postulate of analogous central processing between MI and motor performance (**Table 2**). MI interventions can thus be used to improve motor processing after various cases of neurological disorders (Sharma et al., 2006; Dickstein and Deutsch, 2007). Secondly, motor impairments due to neurological diseases are reflected by changes in MI ability. This conclusion holds in most clinical populations, and therefore validates the assumption that MI reproduces actual motor performance states at the CNS level – yet without going to completion, hence the hypothesis of motor inhibition. Eventually, changes in MI ability after neurological disorder provide new perspectives to the study of motor command inhibition during MI.

STROKE

Motor imagery ability seems preserved in most cases following stroke (Johnson et al., 2002; Sabate et al., 2007; Malouin et al., 2008). However, MI accuracy impairments were described in stroke and brain injured patients. These changes support the assumption that, during MI, the CNS reproduces a state of actual motor processing. Indeed, MI ability changes mirror those observed during actual motor performance. For instance, the time required to perform mental rehearsal of actions involving impaired limbs increased compared to that of actions performed with spared effectors (e.g., Decety and Boisson, 1990; Sirigu et al., 1995; Wu et al., 2010; Dettmers et al., 2012). Stroke patients also reported decreased imagery vividness during imagination of movements performed with the affected side in the case of lateralized brain lesions (Malouin et al., 2008). Malouin et al. (2004) suggested that temporal uncoupling between MI and PP could also occur during mental simulation of actions involving the stroke-unaffected side. Nonetheless, most findings in stroke patients support the assumption that MI mirrors motor impairments resulting from cerebral damage (**Table 2**). Inconsistencies regarding MI ability changes after stroke (i.e., specific or non-specific to the motor impairment) might account for the nature and localization of the stroke lesion (Liepert et al., 2012).

Imagery studies on stroke patients are largely consistent with results obtained from studies in healthy subjects – showing that MI is a dynamic state of motor processing, reproducing the features of CNS activity in a similar way to that during actual motor performance. Examining MI ability after brain lesions can thus contribute to understanding of the neural processes mediating actual motor performance. Sirigu et al. (1996) reported that mismatching between MI and PP times characterized the parietal brain lesion in stroke patients. The authors inferred that the parietal cortex might play a key role in elaboration of movement representation during motor preparation. More recently, Stinear et al. (2007) found that right hemisphere stroke patients *overestimated* MI duration as compared to PP, while left hemisphere stroke patients achieved more accurate temporal congruence between actual and imagined time. These data support hemispheric specificity with regard to internal generation of the temporal parameters of actual execution. As the temporal characteristics of MI are also affected by lesions, these results corroborate previous findings with regard to the neural substrates mediating movement preparation during MI.

To summarize, studies of MI ability in stroke patients show that MI and PP share common neural substrates and involve similar motor commands. But to what extent do imagery data obtained from stroke patients shed light on the issue of motor inhibition during MI? As MI ability is usually preserved following trauma to the nervous system, one should assume that inhibition remain possible even after cerebral damage. As several studies have reported that MI ability is preserved after brain damage affecting both cortical and subcortical motor networks, we may postulate that cerebral structures mediating motor control do not play a critical inhibitory role. This assumption would be congruent with some TMS findings assuming that there is no specific suppressive mechanism occurring at the brain level to inhibit motor output during MI, which would rather be caused by an incomplete level of CNS facilitation during MI (i.e., the level of excitability would not reach the motor threshold during MI, contrary to during PP). Nonetheless, Schwoebel et al. (2002) reported the case of a stroke hemiparetic patient with bilateral parietal brain lesions around the primary somatosensory cortex, who fully executed the "imagined" actions. As the patient performed more efficiently the demanded motor act during MI than during PP, authors argued that overt movements during MI reflected overt processing of the forward models for overt actions. These would be preserved and recalled during MI. Further, the authors nicely demonstrated that sensory integration, mediated by the primary somatosensory cortex, distinguished between overt and covert performance. Impaired sensory integration due to brain lesion thus explained both accurate actual executions during MI and altered voluntary motor performance during PP. Critically, this case report indicated that the motor command was effectively built up during MI, and may normally be actively inhibited throughout motor processing by specific interactions between sensorimotor brain regions. While the precise mechanisms underlying this effect remain unclear, the hypothesis that sensory feedback integration may be a key component for efficient motor suppression during MI is supported by neuroimaging findings (Solodkin et al., 2004; Alkadhi et al., 2005).

Table 2 | Studies looking at the motor command during motor imagery.

Study	Sample of patients and controls	Methodology	Result regarding MI ability changes and potential implications for understanding motor command processing and/or subsequent motor inhibition
Alkadhi et al. (2005)	SCI ($n = 8$) Controls ($n = 8$)	fMRI	MI recruited the neural networks subserving MI and actual movements in healthy controls. Primary motor cortex activity during MI in patients was activated to the same extent than during actual practice in healthy controls, suggesting weakened inhibition
Cramer et al. (2005)	SCI ($n = 12$) Controls ($n = 12$)	fMRI	No task modulation in cerebral activity between MI and PP. Reduced activation volumes in the primary sensorimotor cortex and increased activity within the primary sensorimotor cortex during MI reflect brain function changes after SCI
Gustin et al. (2010)	SCI ($n = 11$) Controls ($n = 19$)	fMRI	Contrary to controls, MI elicited activity within the primary motor area and several brain regions included in the pain neuromatrix. Activity correlated to pain perception during MI
Hotz-Boendermaker et al. (2008)	SCI ($n = 9$) Controls ($n = 12$)	fMRI	Cerebral activity during attempted and imagined movement supports motor program preservation. Recruitment of additional brain regions during MI (compared to healthy controls) reflects altered sensorimotor integration
Lacourse et al. (1999)	SCI ($n = 19$) Controls ($n = 10$)	EEG	Isomorphic electrophysiological correlates during MI and attempted execution in SCI patients, but not in healthy controls. Weakened inhibitory mechanisms as a consequence of SCI, due to deafferentation
Olsson (2012)	SCI ($n = 1$) Controls ($n = 8$)	fMRI	Changes in MI ability according to the remaining capabilities of the motor system
Battaglia et al. (2006)	Stroke ($n = 8$) Controls ($n = 10$)	TMS	Reduced corticospinal facilitation supporting that unilateral stroke patients have lateralized MI deficits
Daprati et al. (2010)	Stroke ($n = 32$) Controls ($n = 12$)	Mental rotation	Impaired MI ability. Patients may have developed MI strategies independently from the actual state of the motor system
Decety and Boisson (1990)	SCI ($n = 4$) Brain injury ($n = 6$)	Mental chronometry	Contrary to SCI patients, stroke patients presented longer MI times when engaging the paralyzed upper/lower limb, comparing to MI of actions with unaffected limbs. For movements that could be physically executed, patients achieved the temporal congruence between MI and executed actions
Dettmers et al. (2012)	Stroke ($n = 31$) Tetraparetic ($n = 10$)	Mental chronometry KVIQ VMIQ	MI ability is impaired on the affected side of the lesion, specifically after stroke eliciting deafferentation. In both clinical populations, the features of MI ability reflect the actual state of the motor system
Gonzalez et al. (2005)	Stroke ($n = 11$) Controls ($n = 11$)	Mental chronometry	Higher MI and PP times in patients who recovered from stroke than in healthy controls
Kagerer et al. (1998)	Brain injury ($n = 4$) Controls ($n = 4$)	Mental chronometry	Patients exhibited longer MI and PP times than for actions involving the more affected side with preserved temporal congruence between MI and PP
Kimberley et al. (2006)	Stroke ($n = 10$) Controls ($n = 10$)	fMRI	Cerebral activity during MI reflects the ipsilateral control of the stroke-affected hand, a common plastic brain change after lateralized stroke lesions
Liepert et al. (2012)	Stroke ($n = 20$)	TMS Mental chronometry	MI ability impaired for movements involving the stroke-affected hand, but only in patients suffering from a somatosensory brain lesion compared to patients with "*pure motor strokes*"
Malouin et al. (2008)	Stroke ($n = 32$) Controls ($n = 32$)	KVIQ	Patients obtained better scores when MI concerned the unaffected side of the lesion, but only for MI of lower limb actions
Sabate et al. (2004)	Stroke ($n = 9$) Controls ($n = 10$)	Mental chronometry	Decreased movement velocity during PP also observed during MI. The hemispheric-dependent effects of lateralized stroke on the actual motor performance of each hand (affected/non-affected) was reproduced during MI
Sabate et al. (2007)	Stroke ($n = 33$) PD ($n = 8$) Controls ($n = 18$)	Mental chronometry	Strong correlation between MI and PP times after stroke. Mismatches between MI and PP times support that changes in MI ability reflect the actual state of the motor system

(Continued)

Table 2 | Continued

Study	Sample of patients and controls	Methodology	Result regarding MI ability changes and potential implications for understanding motor command processing and/or subsequent motor inhibition
Schwoebel et al. (2002)	Stroke ($n=1$)	Motor tasks Mental rotation	A patient with bilateral parietal brain lesion fully executed the mentally rehearsed actions. Inhibition during MI was impaired, presumably due to disturbances within a fronto-parietal circuit mediating motor inhibition
Sirigu et al. (1995)	Stroke ($n=1$)	Mental chronometry	Temporal parameters of MI predicted that of PP in a variety of situations, altogether reflecting hypokinesia after unilateral stroke
Sharma et al. (2009a)	Stroke ($n=8$) Controls ($n=13$)	fMRI	Abnormal functional connectivity patterns within the motor network during MI correlated with motor outcome after stroke recovery
Sharma et al. (2009b)	Stroke ($n=20$) Controls ($n=17$)	fMRI	During MI of the affected hand, activation of the anterior subdivision of cM1 was similar to that during PP, and activity of the ipsilesional posterior subdivision of M1 correlated with motor performance. The result support that MI reveals the actual state of the motor system after stroke
Stinear et al. (2007)	Stroke ($n=12$) Controls ($n=8$)	Mental chronometry TMS	Absence of corticospinal facilitation during MI in the stroke-affected hand
Szameitat et al. (2012)	Stroke ($n=5$) Controls ($n=21$)	fMRI	Cortical activations during MI resemble that during attempted overt execution within sensorimotor and premotor cortices. Potential analogous involvement of the sensorimotor system in the two tasks
Vromen et al. (2011)	Stroke ($n=21$)	Mental rotation	Stroke patients ($n=20$) without spatial neglect outperformed a patient ($n=1$) with spatial neglect during a visual mental rotation task involving the upper limb
Wu et al. (2010)	Stroke ($n=18$)	Mental chronometry	Longer times required to imagine upper limb actions involving stroke-affected effectors
Cohen et al. (2011)	PD ($n=24$) Controls ($n=10$)	Mental chronometry	Temporal discrepancies between times required to imagine and actually walk through a narrow doorway characterized PD patients with freezing of gait syndrome
Cunnington et al. (1997)	PD ($n=14$) Controls ($n=10$)	EEG	Impaired motor preparation, while potentials associated with motor execution seemed relatively preserved
Cunnington et al. (2001)	PD ($n=6$) Controls ($n=3$)	PET	Reduced pre-SMA activation and compensatory brain activity during MI altogether characterized the motor deficit in PD patients
Dominey et al. (1995)	PD ($n=7$) Controls ($n=7$)	Mental chronometry	Asymmetrical slowing of MI times according to the affected side in lateralized PD patients, hence supporting that MI and PP shared common neural structures
Helmich et al. (2007)	PD ($n=19$)	fMRI	MI involving the affected side in lateralized PD patients recruited additional cognitive resources compared to MI involving the unaffected side
Helmich et al. (2012)	PD ($n=38$) Controls ($n=19$)	fMRI	Distinct sensorimotor processing at the subcortical level during MI characterized patients with and without resting state tremor
Heremans et al. (2011)	PD ($n=14$) Controls ($n=14$)	MI questionnaires Mental chronometry	MI ability was preserved in PD patients, but was performed more slowly than in healthy controls
Kuhn et al. (2006)	PD ($n=8$)	EEG	Analogous contribution of subthalamic nucleus to feedforward organization during MI and PP of wrist actions. Electrophysiological correlates of MI within these structures support its role in sensory feedback integration for overt motor and postural regulations after PD
Samuel et al. (2001)	PD ($n=6$) Controls ($n=6$)	PET	MI yielded decreased activity in frontal areas (dorsolateral frontal cortex), hence reflecting impaired motor preparation in PD patients as compared to healthy controls
Thobois et al. (2000)	PD ($n=8$) Controls ($n=8$)	PET	MI elicited reduced activations for movement with the affected side in lateralized PD patients. MI of the unaffected side was impaired, but to a lesser extent

(Continued)

Table 2 | Continued

Study	Sample of patients and controls	Methodology	Result regarding MI ability changes and potential implications for understanding motor command processing and/or subsequent motor inhibition
Thobois et al. (2002)	PD ($n = 7$)	PET	Subthalamic nucleus stimulation analogously improved cerebral activity during MI and PP
Diers et al. (2010)	Amputees ($n = 14$) Controls ($n = 9$)	fMRI	MI activated different neural substrates depending on whether amputee patients experienced phantom limb pain or not. MI activated the contralateral primary sensorimotor cortex only in non-pain patients
MacIver et al. (2008)	Amputees ($n = 13$) Controls ($n = 6$)	fMRI	MI training elicited reversed sensorimotor plasticity in amputee that corresponded to decreased phantom limb pain symptoms
Marconi et al. (2007)	Amputees ($n = 8$) Controls ($n = 9$)	TMS	MI mirrored sensorimotor reorganizations in the patients. Upper/lower limb inhibitory relationships within cM1 might be removed after amputation
Nico et al. (2004)	Amputees ($n = 16$) Controls ($n = 7$)	Mental rotation task	MI was affected by amputation in patients as compared to healthy controls. Selective MI impairments were observed according to whether amputation affected the dominant/non-dominant limb
Raffin et al. (2012b)	Amputees ($n = 14$)	fMRI	Partially overlapping, albeit non-identical, neural networks mediating MI, and attempted physical practice with the phantom limb

PARKINSON'S DISEASE

MI ability changes in patients suffering from PD are consistent with those observed in stroke patients. Central processing during MI is selective according to the limbs affected by PD and mirrors the actual motor impairment (Dominey et al., 1995; Helmich et al., 2007), even though MI of non-affected body regions may also be disturbed to a lesser extent (Thobois et al., 2000). In spite of basal ganglia dysfunction, MI ability is preserved in early and mid-stage PD patients (Heremans et al., 2011). However, several neuroimaging studies have discovered abnormal brain activation patterns during MI in PD patients as compared to healthy controls (Cunnington et al., 1997; Samuel et al., 2001). Specifically, reduced premotor and sensorimotor activations, as well decreased cerebellum activation during MI were reported (Thobois et al., 2000; Cunnington et al., 2001; Samuel et al., 2001). Compensatory activations occurring during MI of actions involving affected effectors were also reported in PD patients, which could reflect the actual motor deficit (Cunnington et al., 2001; Thobois et al., 2002; Helmich et al., 2007). Of great interest, however, is the finding that subthalamic nucleus electrical stimulation enabled reduction of compensatory activations during MI in PD patients (Thobois et al., 2002). These data provide meaningful evidence that elaboration of motor command *does*, in fact, occur during MI after PD, as the central MI activity seems to mirror the effect of PD on actual motor performance (Cunnington et al., 1997; Samuel et al., 2001; Tremblay et al., 2008). Spontaneous eye movements occur during MI and resemble those occurring during PP (Heremans et al., 2008). Heremans et al. (2012) observed that external visual cueing even reinforced MI accuracy in PD patients, thus confirming that the central processing of somatic motor signals was an intrinsic component of the MI experience. Finally, several studies evidenced that PD patients may benefit from MI training in the rehabilitation of motor disorders, which suggests skill transfer from MI to PP with regards to actual motor processing (Mannix et al., 1999; Lim et al., 2006; Braun et al., 2011). Tamir et al. (2007) reported the adjunctive benefits of MI training in motor rehabilitation following PD. MI practice contributed to reduce bradykinesia (Subramanian et al., 2011). PD typically refers to basal ganglia dysfunction. These structures are known to play a role in overt motor performance inhibition in healthy subjects, through specific neural interactions with cM1 during the early stages of motor processing (Stinear et al., 2009). However, it remains unclear whether these structures also participate in motor inhibition during MI, but recent findings suggest that PD patients lose the ability to elicit corticospinal facilitation during MI (Tremblay et al., 2008). The authors assumed that the patients failed to involve the motor system during MI. Nonetheless as the MI ability seems preserved and to elicit sensorimotor activity at the brain level in most neuroimaging studies, it can be hypothesized that changes in basal ganglia activity during MI could increase inhibitory interactions during MI. This postulate remains a working hypothesis awaiting experimental proofs.

AMPUTEES

Consistent reorganizations of the cortical sensorimotor map occur after limb amputation, due to neuroplasticity (i.e., the capability of synapses to adapt their structure and function in response to environmental and behavioral demand). The expansion within primary sensorimotor cortices of the cortical surface corresponding to unaffected body parts toward the adjacent deafferented and deefferented areas (i.e., corresponding to the missing limb) is now well-established (Knecht et al., 1996; Pascual-Leone et al., 1996; Karl et al., 2001; Ramachandran et al., 2010). Nonetheless, the motor system preserves the ability to process central commands controlling the missing limb following amputation, as suggested in several neuroimaging studies reporting similar sensorimotor activations during actions performed either with the phantom limb or with the contralateral unaffected one (Ersland et al., 1996; Roux et al., 2001, 2003). EMG recordings at the level of the stump revealed specific motor commands suggesting preservation of motor programs controlling the amputated limb (Reilly et al., 2006). Are motor commands processed during MI of actions with the phantom limb? Recent findings suggest that the neural networks mediating MI and PP of actions with the phantom limb in amputee patients consistently overlap, in spite of significant differences (Raffin et al., 2012b). Most findings regarding imagery in these patients are derived from studies investigating the therapeutic management of phantom limb pain (i.e., a frequent disabling consequence following amputation; Shukla et al., 1982; Ehde et al., 2000). A causal relationship was found between sensorimotor reorganizations and phantom limb pain (Flor et al., 1995; MacIver et al., 2008). Expansion of homunculus regions corresponding to spared body parts toward deafferented areas (e.g., hand region "invaded" by regions corresponding to face body parts), in both primary somatosensory and motor cortices, characterize phantom limb pain patients as compared to pain-free amputees (Lotze et al., 1999a, 2001; Karl et al., 2001). Interestingly, involving the motor system into overt motor processing (i.e., attempted executed actions) with the phantom limb (e.g., using visual feedback) tends to produce pain relief (Chan et al., 2007; Mercier and Sirigu, 2009). An interesting practical implication of this finding is that if MI reproduces overt motor processing states at the CNS level, then MI *training* might also be efficient in the management of neuropathic pain after amputation. MacIver et al. (2008) investigated the effects of a 6-week MI training program on phantom limb pain. fMRI scanning sessions were performed before and after the experimental intervention. MI training significantly reduced pain symptoms. fMRI investigations highlighted a reversed plasticity as compared to plastic changes due to amputation observed in the patients during the pretest. These results promote the therapeutic relevance of MI in the rehabilitation of pain disorders, presumably due to actual motor processing at the CNS level. Indeed, change in the cortical representation of a body segment is usually achieved through repetitive motor practice (Wolf et al., 2002). In the study by MacIver et al. (2008), MI elicited reversed plasticity, hence supporting the assumption that amputees are able to process motor commands during MI, even for actions with their injured limb. In this regard, however, the fact that phantom limb pain patients and those with and non-phantom limb pain present *different* brain responses to MI is of special interest (Diers et al., 2010). Only non-pain patients activated the contralateral primary sensorimotor cortex during MI. Some authors have reported that neuropathic pain after amputation could be due to the mismatching between motor output and sensory feedback (Mayer et al., 2008), in agreement with the model of pathological pain by Harris (1999). MI might therefore reproduce a pathological state where the motor system fails to elicit the primary sensorimotor activity demand of the cognitive task. These findings concur with a large body of literature supporting the role of differential neuroplasticity in the generation of pain symptoms, which seems reproduced during MI. Further, it can be assumed that amputees no longer require to inhibit the motor command during MI of actions involving the phantom, as no muscle activity could occur (a similar observation can be made in SCI patients – see below). If MI *does* involve active motor suppression at the CNS level, these mechanisms will no longer be relevant after amputation and might potentially be reshaped by the emergence of a new body schema, prompting considerable neural reorganization. Challenging considerations to this approach comes from the fact that amputee patients usually preserve a perceptual representation of their missing limb. Further, no study yet reported weakened inhibition during MI after amputation and no EMG activity was recorded at the level of the stump (Raffin et al., 2012a). This is an interesting perspective for future studies: If there is no changes in the ability to inhibit the motor commands during MI, conclusions regarding the neural underpinnings of motor suppression during MI might be drawn.

SCI PATIENTS

Motor activity during MI can be inferred from the effects of MI practice in neuropathic pain SCI patients. Gustin et al. (2008) observed that MI increased neuropathic pain intensity in six SCI patients out of seven during mental rehearsal of actions involving infra-lesional effectors. Gustin et al. (2010) delineated the neural substrates mediating this pain response to MI. They observed several activations within the pain neuromatrix network, correlated to increased pain perception when SCI patients performed MI. Both the ipsilateral premotor cortex and the SMA participated to actual motor processing. When compared to healthy controls, SCI patients elicited greater activation within cM1 attesting that motor output during MI reached the circuitry underlying pain response, due to cerebral reorganizations after the neurological lesion. Therefore, it seems reasonable to assume that motor signals *do* occur during MI, and might act as a triggering stimulus for pain in specific clinical cases of neurological disorders (see above for considerations in amputee patients).

As in the case of stroke, PD, and amputees patients, research investigating MI ability after SCI provides further insight into central processing of motor signals during MI. Using mental chronometry, Decety and Boisson (1990) observed that SCI patients achieved close temporal congruence between MI and PP, whereas brain injured patients failed to do so. These data support the theory that MI accuracy is preserved following SCI because, contrary to what happens after stroke, SCI does not result in cerebral damage. However, SCI elicits consistent reorganization of the sensorimotor cortical maps, with changes in neuronal excitability (Topka et al., 1991; Curt et al., 2002; Dunlop, 2008; Kokotilo et al., 2009) due to deafferentation and deefferentation (Bruehlmeier et al., 1998). As MI is mediated by cerebral substrates overlapping with motor-related regions reorganized consecutively to the lesion, plastic changes after SCI are likely to affect MI ability. Cramer et al. (2005) reported altered MI processing in SCI patients as compared to healthy controls during MI of right foot movements. In this study, task modulation in central activity between MI and PP was absent in SCI patients whereas different brain activation patterns mediated the two tasks in healthy controls. Similarly, Alkadhi et al. (2005) reported that during MI, SCI patients recruited those neural networks that typically mediate both MI and PP in healthy controls. As already mentioned, MI usually results in lower activation intensities than PP in healthy participants, which is typically imputable to inhibitory processes (Porro et al., 1996). Therefore, as early hypothesized by Lacourse et al. (1999), deafferentation and deefferentation following SCI may result in weakened motor command inhibition during MI. SCI patient no longer require to inhibit the mentally rehearsed movement due to the spinal cord lesion, which prevents neural transmissions toward peripheral effectors. This clinical topic is different than that in amputees (see above), in that after SCI patients no longer feel their deafferented and deefferented body regions. Alkadhi et al. (2005) suggested that deafferentation may be a key component for adaptive brain changes regarding inhibition during MI, hence matching the conclusions by Schwoebel et al. (2002). In a recent magnetoencephalographic control-case study (Di Rienzo et al., submitted), a SCI patient presented remarkably similar activations within cM1 during both MI and PP, hence suggesting weakened inhibition of this area during MI. This pattern of activation was associated with disturbed functional network of interrelations between cM1, S1, and SMA. By contrast, a healthy age-matched control participant presented a significant reduction in cM1 activation during MI, with significant inter-relationships in neural activities between cM1 and both SMA and S1. Interestingly, both S1 and SMA are thought to play a key role in motor suppression during MI (Schwoebel et al., 2002; Solodkin et al., 2004; Alkadhi et al., 2005; Kasess et al., 2008). Hotz-Boendermaker et al. (2008) discovered that patients remained nonetheless able to subjectively distinguish MI from attempted PP for actions involving infra-lesional effectors, but the authors mentioned that MI elicited activation of additional brain regions in SCI patients in comparison to controls – presumably to assist motor commands processing. Neuroimaging studies in SCI patients support the theory that central reorganization after SCI results in increased congruence between MI and PP, due to weakened inhibitory processes as a consequence of deafferentation. Accordingly, motor inhibition during MI may have a cortical component involving specific interactions between perirolandic sites and cM1. These results are complementary to those from the TMS approach to understanding inhibition (i.e., intracortical inhibition within M1 and/or incomplete state of CNS facilitation). However, central activity in SCI patients suggests that MI recruits motor programs for overt movements with paralyzed effectors (Sabbah et al., 2002), thereby confirming that these patients preserve the ability to process motor command signals during MI.

A discussed previously, neuroimaging studies support the likelihood of weakened motor inhibition during MI after SCI. However, whether the reductions in sensorimotor activity between MI and PP result in greater transmission of neural signals to the descending pathways during MI remains unknown. Roy et al. (2011) obtained evidence of the downregulation of intracortical inhibition during MI after SCI using paired-pulse TMS. These data support previous findings reporting increased excitability after complete SCI in spared neural pathways (Topka et al., 1991). Reduced intracortical inhibition after SCI is thought to enable unmasking of latent synaptic connections at the cortical level, providing a possible causal mechanism for cerebral plasticity after SCI (Saturno et al., 2008). As MI and PP are assumed to be functionally equivalent, it is plausible that changes in intracortical inhibition of the sensorimotor cortex (allowing the reshaping of actual motor performance) may be reflected during MI, hence confirming some shared neural substrates between these two processes.

Taken together, clinical data from SCI, stroke, amputees, and PD patients converge to suggest that central patterns elicited during MI effectively reflect the internal elaboration of motor commands, although specific clinical impairment provides different and complementary insights to our current knowledge regarding motor inhibition during MI.

CONCLUSION

In this paper, we reviewed data that support the emerging hypothesis that both central and peripheral neurophysiological correlates of MI tightly resemble those elicited by actual practice of the same task, even in the absence of overt movement production. Such data also suggest that motor commands are involved during MI. Furthermore, we highlighted evidence that the isomorphism between the representation of imagined and executed actions is preserved, if not strengthened, in specific cases of neurological disorders. Neuroimaging studies clearly support the involvement of both primary and secondary motor-related areas during MI, hence suggesting that neural impulses for motor commands may be elaborated at the cerebral level and addressed, at least partially, from cM1 to the anterior part of the spinal cord via the descending pathways. TMS studies of MI also confirmed this assumption and elucidated our understanding of the neurophysiological processes mediating the involvement of cM1 during MI. Accordingly, corticospinal facilitation during MI might result from changes in excitability at the cortical level. Some authors have also reported that residual EMG activity during MI reflects the features of the motor commands.

Based on these issues, we should now consider the unresolved question of how motor commands are inhibited throughout the motor system to prevent overt execution during MI. We postulate that motor inhibition during MI may not result from a parallel neurophysiological process, concomitant to MI, designed to prevent muscle contractions. First, TMS studies have shown that MI produces opposite effects to those elicited by voluntary relaxation of peripheral effectors regarding corticospinal facilitation (Taniguchi et al., 2008). Secondly, neuroimaging studies have failed to highlight specific neural structures mediating motor inhibition during MI, while TMS data support the idea of increased neuronal excitability and reduced intracortical inhibition within cM1 during MI. A notable finding from neuroimaging research is that secondary motor-related areas like the cerebellum and SMA might play a key role in motor output suppression during MI. Also, impaired sensory feedback integration following deafferentation or brain lesions around the primary somatosensory cortex result in weakened inhibition during MI, thus promoting the role of sensory sites in motor output suppression during MI. Therefore, inhibition during MI may be a functional process resulting from the specific contribution of neural sites usually dedicated to overt motor processing. This theoretical stance might account for the fact that MI activates the motor system in a lesser extent to actual practice.

If we assume that motor inhibition may be intrinsic to the motor command during MI, and that it is potentially mediated by a

Table 3 | Studies addressing the question of the motor inhibition during motor imagery.

Authors	Type of study	Method	Participants	Potential inhibitory regions
Alkadhi et al. (2005)	Motor imagery of foot movement	fMRI	Healthy ($n = 8$) Patients ($n = 8$)	Motor command suppression but no clear inhibitory regions
Bonnet et al. (1997)	Motor imagery of a foot pressure on a pedal	Reflex stimulation	Healthy ($n = 20$)	Inhibitory spinal influences
Di Rienzo et al. (submitted)	Case study with a C6–C7 quadriplegic patient	MEG	Patients ($n = 1$)	Primary sensory area and supplementary motor area
Jeannerod (2001, 2006)	Review papers	–	–	prefrontal cortical areas and/or brainstem and spinal influences
Kasess et al. (2008)	Motor imagery of finger movements	fMRI	Healthy ($n = 8$)	Supplementary motor area
Lotze et al. (1999b)	Motor imagery of hand movements	fMRI	Healthy ($n = 10$)	Posterior cerebellum
Schwoebel et al. (2002)	Case study with a patient suffering from bilateral parietal lesions	Psychophysic experiment	Patients ($n = 1$)	Fronto-parietal network
Solodkin et al. (2004)	Motor imagery of a finger-to-thumb opposition task	fMRI	Healthy ($n = 18$)	Superior parietal lobule and supplementary motor area
Deiber et al. (1998)	Motor imagery of finger movements	TEP	Healthy ($n = 10$)	Inferior frontal cortex
Lebon et al. (2012)	Motor imagery and mental rotation of a pinching movement	TMS	Healthy ($n = 11$)	Inferior parietal lobule

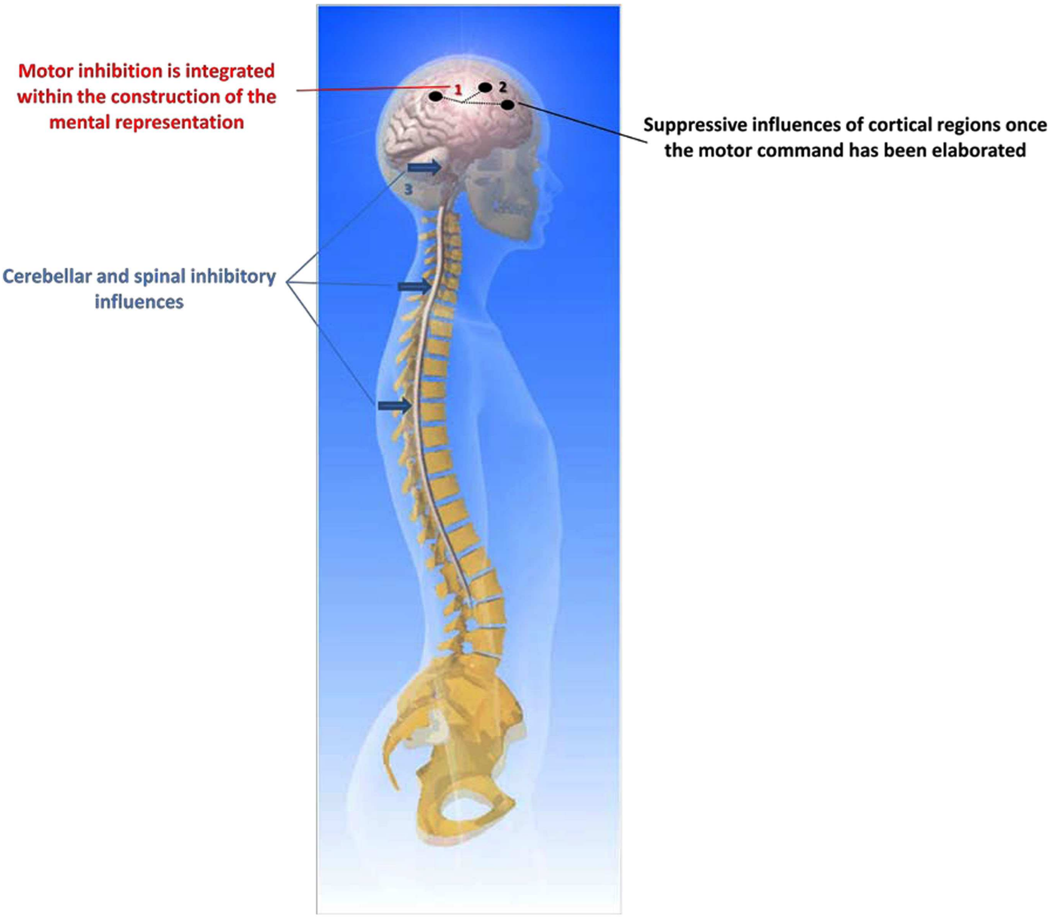


FIGURE 1 | The three possible routes of motor command inhibition during motor imagery.

highly specific interplay between motor-related neural structures, a promising focus of future research may be identified. Specifically, priority for future investigators will be to explore the extent to which cM1 is subjected to active inhibition during MI. Using advanced statistical modeling, several authors have shown that reduced cM1 activation during MI is related to the suppressive influence of other motor-related brain regions, thus suggesting that motor inhibition during MI may also intervene at the early stages of motor planning (Kasess et al., 2008).

Based on current understandings and literature evidence, we promote a multifactorial explanation of motor inhibition during MI that might involve both cerebral and spinal mechanisms (**Table 3**). From these findings, we postulate that there are three possible routes of motor command inhibition during MI (**Figure 1**). One may first hypothesize that motor inhibition is a part of the imagery experience, hence only *subthreshold* motor commands are sent to the effectors to prevent movement execution. A second alternative is the possibility that the inhibitory cerebral regions progressively weaken the motor command during the time course of the MI process, so that only a *residual* activity is sent and can be recorded in the corresponding muscles (for a review on a similar chain of processes during inhibitory motor control in No-Go paradigms, see Band and van Boxtel, 1999). Finally, it is possible that downstream regions including brainstem and spinal influences contribute to motor inhibition at a later stage than in the case of the other two possibilities. There is still, however, one key element requiring further experimental investigation. Specifically, research should establish the degree to which cM1 is inhibited by suppressive neural impulses of cerebral origin during MI. Similarly, it is vital to investigate whether or not neural impulses elicited by MI are blocked by spinal mechanisms triggered by descending input of cerebral origin. Interestingly, inhibition of actual actions in decision making experiments (i.e., Go/No-Go paradigms) or during motor control of complex skills that require real-time adaptation to changing environmental constraints, can happen in the very late stages, and even during overt motor processing. We should also investigate to a greater extent how these inhibitory processes exactly work as these do not operate under a principle of all or nothing. As mentioned by Band and van Boxtel, 1999, p. 190), "*A crucial question is whether it is most important to define the locus of* [motor] *inhibition by the source of inhibitory activity (agent), by the process at which inhibition was exerted (site), or by the location where reduction of response activity can be recorded (manifestation)*." An important aspect is that imagined movement may be accompanied by reduced motor activity at a level which the subject is unaware of, such as when voluntarily performing movements simultaneously with its own action representation for improving MI vividness. Therefore, as the functional equivalence between imagined and executed actions has mainly been considered at the scope of neurophysiological correlates in activation, it might also be observable through the neural processes mediating motor output suppression during these two behaviors.

ACKNOWLEDGMENTS

We thank Patrice Thiriet, Alix Poulot, and Emmanuel Bettler (*Innovation Conception et Accompagnement en Pédagogie* – ICAP – Université Claude Bernard Lyon 1) for the Figure template.

REFERENCES

Abbruzzese, G., Assini, A., Buccolieri, A., Marchese, R., and Trompetto, C. (1999). Changes of intracortical inhibition during motor imagery in human subjects. *Neurosci. Lett.* 263, 113–116.

Alkadhi, H., Brugger, P., Boendermaker, S. H., Crelier, G., Curt, A., Hepp-Reymond, M. C., and Kollias, S. S. (2005). What disconnection tells about motor imagery: evidence from paraplegic patients. *Cereb. Cortex* 15, 131–140.

Aron, A. R., Durston, S., Eagle, D. M., Logan, G. D., Stinear, C. M., and Stuphorn, V. (2007). Converging evidence for a fronto-basal-ganglia network for inhibitory control of action and cognition. *J. Neurosci.* 27, 11860–11864.

Bakker, F. C., Boschker, M. S. J., and Chung, T. (1996). Changes in muscular activity while imagining weight lifting using stimulus or response propositions. *J. Sport Exerc. Psychol.* 18, 313–324.

Band, G. P. H., and van Boxtel, G. J. M. (1999). Inhibitory motor control in stop paradigms: review and reinterpretation of neural mechanisms. *Acta Psychol. (Amst.)* 101, 179–211.

Battaglia, F., Quartarone, A., Ghilardi, M. F., Dattola, R., Bagnato, S., Rizzo, V., Morgante, L., and Girlanda, P. (2006). Unilateral cerebellar stroke disrupts movement preparation and motor imagery. *Clin. Neurophysiol.* 117, 1009–1016.

Bird, E. I. (1984). EMG quantification of mental rehearsal. *Percept. Mot. Skills* 59, 899–906.

Bonnet, M., Decety, J., Jeannerod, M., and Requin, J. (1997). Mental simulation of an action modulates the excitability of spinal reflex pathways in man. *Brain Res. Cogn. Brain Res.* 5, 221–228.

Boschker, M. S. J. (2001). *Action-Based Imagery: On the Nature of Mentally Imagined Motor Actions.* Amsterdam: University of Amsterdam.

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.

Braun, S. M., Beurskens, A. J., Borm, P. J., Schack, T., and Wade, D. T. (2006). The effects of mental practice in stroke rehabilitation: a systematic review. *Arch. Phys. Med. Rehabil.* 87, 842–852.

Braun, S., Beurskens, A., Kleynen, M., Schols, J., and Wade, D. (2011). Rehabilitation with mental practice has similar effects on mobility as rehabilitation with relaxation in people with Parkinson's disease: a multicentre randomised trial. *J. Physiother.* 57, 27–34.

Bruehlmeier, M., Dietz, V., Leenders, K. L., Roelcke, U., Missimer, J., and Curt, A. (1998). How does the human brain deal with a spinal cord injury? *Eur. J. Neurosci.* 10, 3918–3922.

Brunia, C. H. M. (2003). "How is stopping realized?" in *Errors, Conflicts, and The Brain. Current Opinions on Performance Monitoring*, eds M. Ullsperger and M. Falkenstein (Liepzig: MPI of Cognitive Neuroscience), 96–103.

Callow, N., and Waters, A. (2005). The effect of kinesthetic imagery on the sport confidence of flat-race horse jockeys. *Psychol. Sport. Exerc.* 6, 443–459.

Carpenter, W. B. (1894). *Principles of Mental Physiology.* Appleton: D. Appleton.

Chan, B. L., Witt, R., Charrow, A. P., Magee, A., Howard, R., Pasquina, P. F., Heilman, K. M., and Tsao, J. W. (2007). Mirror therapy for phantom limb pain. *N. Engl. J. Med.* 357, 2206–2207.

Clark, S., Tremblay, F., and Ste-Marie, D. (2004). Differential modulation of corticospinal excitability during observation, mental imagery and imitation of hand actions. *Neuropsychologia* 42, 105–112.

Cohen, R. G., Chao, A., Nutt, J. G., and Horak, F. B. (2011). Freezing of gait is associated with a mismatch between motor imagery and motor execution in narrow doorways, not with failure to judge doorway passability. *Neuropsychologia* 49, 3981–3988.

Collet, C., and Guillot, A. (2010). "Autonomic nervous system activities during imagined movements," in *The Neurophysiological Foundations of Mental and Motor Imagery*, eds A. Guillot and C. Collet (New York: Oxford University Press), 95–108.

Cramer, S. C., Lastra, L., Lacourse, M. G., and Cohen, M. J. (2005). Brain motor system function after chronic, complete spinal cord injury. *Brain* 128, 2941–2950.

Cunnington, R., Egan, G. F., O'Sullivan, J. D., Hughes, A. J., Bradshaw, J. L., and Colebatch, J. G. (2001). Motor imagery in Parkinson's disease: a PET study. *Mov. Disord.* 16, 849–857.

Cunnington, R., Iansek, R., Johnson, K. A., and Bradshaw, J. L. (1997). Movement-related potentials in Parkinson's disease. Motor imagery and movement preparation. *Brain* 120(Pt 8), 1339–1353.

Curt, A., Alkadhi, H., Crelier, G. R., Boendermaker, S. H., Hepp-Reymond, M. C., and Kollias, S. S. (2002). Changes of non-affected upper limb cortical representation in paraplegic patients as assessed by fMRI. *Brain* 125, 2567–2578.

Daprati, E., Nico, D., Duval, S., and Lacquaniti, F. (2010). Different motor imagery modes following brain damage. *Cortex* 46, 1016–1030.

de Jong, R., Coles, M. G., Logan, G. D., and Gratton, G. (1990). In search of the point of no return: the control of response processes. *J. Exp. Psychol. Hum. Percept. Perform.* 16, 164–182.

de Vries, S., and Mulder, T. (2007). Motor imagery and stroke rehabilitation: a critical discussion. *J. Rehabil. Med.* 39, 5–13.

Decety, J., and Boisson, D. (1990). Effect of brain and spinal cord injuries on motor imagery. *Eur. Arch. Psychiatry Clin. Neurosci.* 240, 39–43.

Decety, J., and Grezes, J. (1999). Neural mechanisms subserving the perception of human actions. *Trends Cogn. Sci. (Regul. Ed.)* 3, 172–178.

Decety, J., Jeannerod, M., Durozard, D., and Baverel, G. (1993). Central activation of autonomic effectors during mental simulation of motor actions in man. *J. Physiol. (Lond.)* 461, 549–563.

Dechent, P., Merboldt, K. D., and Frahm, J. (2004). Is the human primary motor cortex involved in motor imagery? *Brain Res. Cogn. Brain Res.* 19, 138–144.

Deiber, M. P., Ibanez, V., Honda, M., Sadato, N., Raman, R., and Hallett, M. (1998). Cerebral processes related to visuomotor imagery and generation of simple finger movements studied with positron emission tomography. *Neuroimage* 7, 73–85.

Demougeot, L., and Papaxanthis, C. (2011). Muscle fatigue affects mental simulation of action. *J. Neurosci.* 31, 10712–10720.

Dettmers, C., Benz, M., Liepert, J., and Rockstroh, B. (2012). Motor imagery in stroke patients, or plegic patients with spinal cord or peripheral diseases. *Acta Neurol. Scand.* (in press).

Dickstein, R., and Deutsch, J. E. (2007). Motor imagery in physical therapist practice. *Phys. Ther.* 87, 942–953.

Dickstein, R., Gazit-Grunwald, M., Plax, M., Dunsky, A., and Marcovitz, E. (2005). EMG activity in selected target muscles during imagery rising on tiptoes in healthy adults and poststroke hemiparetic patients. *J. Mot. Behav.* 37, 475–483.

Diers, M., Christmann, C., Koeppe, C., Ruf, M., and Flor, H. (2010). Mirrored, imagined and executed movements differentially activate sensorimotor cortex in amputees with and without phantom limb pain. *Pain* 149, 296–304.

Dominey, P., Decety, J., Broussolle, E., Chazot, G., and Jeannerod, M. (1995). Motor imagery of a lateralized sequential task is asymmetrically slowed in hemi-Parkinson's patients. *Neuropsychologia* 33, 727–741.

Driskell, J. E., Cooper, C., and Moran, A. (1994). Does mental practice enhance performance? *J. Appl. Psychol.* 79, 481–492.

Duclos, Y., Schmied, A., Burle, B., Burnet, H., and Rossi-Durand, C. (2008). Anticipatory changes in human motoneuron discharge patterns during motor preparation. *J. Physiol. (Lond.)* 586, 1017–1028.

Dunlop, S. A. (2008). Activity-dependent plasticity: implications for recovery after spinal cord injury. *Trends Neurosci.* 31, 410–418.

Ehde, D. M., Czerniecki, J. M., Smith, D. G., Campbell, K. M., Edwards, W. T., Jensen, M. P., and Robinson, L. R. (2000). Chronic phantom sensations, phantom pain, residual limb pain, and other regional pain after lower limb amputation. *Arch. Phys. Med. Rehabil.* 81, 1039–1044.

Ehrsson, H. H., Geyer, S., and Naito, E. (2003). Imagery of voluntary movement of fingers, toes, and tongue activates corresponding body-part-specific motor representations. *J. Neurophysiol.* 90, 3304–3316.

Ersland, L., Rosen, G., Lundervold, A., Smievoll, A. I., Tillung, T., Sundberg, H., and Hugdahl, K. (1996). Phantom limb imaginary fingertapping causes primary motor cortex activation: an fMRI study. *Neuroreport* 8, 207–210.

Facchini, S., Muellbacher, W., Battaglia, F., Boroojerdi, B., and Hallett, M. (2002). Focal enhancement of motor cortex excitability during motor imagery: a transcranial magnetic stimulation study. *Acta Neurol. Scand.* 105, 146–151.

Fadiga, L., Buccino, G., Craighero, L., Fogassi, L., Gallese, V., and Pavesi, G. (1999). Corticospinal excitability is specifically modulated by motor imagery: a magnetic stimulation study. *Neuropsychologia* 37, 147–158.

Flor, H., Elbert, T., Knecht, S., Wienbruch, C., Pantev, C., Birbaumer, N., Larbig, W., and Taub, E. (1995). Phantom-limb pain as a perceptual correlate of cortical reorganization following arm amputation. *Nature* 375, 482–484.

Fourkas, A. D., Avenanti, A., Urgesi, C., and Aglioti, S. M. (2006). Corticospinal facilitation during first and third person imagery. *Exp. Brain Res.* 168, 143–151.

Gandevia, S. C. (1999). Mind, muscles, and motoneurones. *J. Sci. Med. Sport.* 2, 167–180.

Gandevia, S. C., Wilson, L. R., Inglis, J. T., and Burke, D. (1997). Mental rehearsal of motor tasks recruits alpha-motoneurones but fails to recruit human fusimotor neurones selectively. *J. Physiol. (Lond.)* 505(Pt 1), 259–266.

Gao, Q., Duan, X., and Chen, H. (2011). Evaluation of effective connectivity of motor areas during motor imagery and execution using conditional Granger causality. *Neuroimage* 54, 1280–1288.

Garavan, H., Ross, T. J., Murphy, K., Roche, R. A., and Stein, E. A. (2002). Dissociable executive functions in the dynamic control of behavior: inhibition, error detection, and correction. *Neuroimage* 17, 1820–1829.

Gentili, R., Papaxanthis, C., and Pozzo, T. (2006). Improvement and generalization of arm motor performance through imagery practice. *Neuroscience* 137, 761–772.

Gerardin, E., Sirigu, A., Lehericy, S., Poline, J. B., Gaymard, B., Marsault, C., Agid, Y., and Le Bihan, D. (2000). Partially overlapping neural networks for real and imagined hand movements. *Cereb. Cortex* 10, 1093–1104.

Gonzalez, B., Rodriguez, M., Ramirez, C., and Sabate, M. (2005). Disturbance of motor imagery after cerebellar stroke. *Behav. Neurosci.* 119, 622–626.

Gray, J. A. (1990). Brain systems that mediate both emotion and cognition. *Cogn. Emot.* 4, 269–288.

Grezes, J., and Decety, J. (2001). Functional anatomy of execution, mental simulation, observation, and verb generation of actions: a meta-analysis. *Hum. Brain Mapp.* 12, 1–19.

Gueugneau, N., Crognier, L., and Papaxanthis, C. (2008). The influence of eye movements on the temporal features of executed and imagined arm movements. *Brain Res.* 1187, 95–102.

Guillot, A., and Collet, C. (2008). Construction of the motor imagery integrative model in sport: a review and theoretical investigation of motor imagery use. *Int. Rev. Sport Exerc. Psychol.* 1, 31–44.

Guillot, A., Collet, C., Nguyen, V. A., Malouin, F., Richards, C., and Doyon, J. (2008). Functional neuroanatomical networks associated with expertise in motor imagery. *Neuroimage* 41, 1471–1483.

Guillot, A., Collet, C., Nguyen, V. A., Malouin, F., Richards, C., and Doyon, J. (2009). Brain activity during visual versus kinesthetic imagery: an fMRI study. *Hum. Brain Mapp.* 30, 2157–2172.

Guillot, A., Di Rienzo, F., and Collet, C. (2012). "The neurofunctional architecture of motor imagery," in *Functional Magnetic Resonance Imaging/Book 1*, eds T. D. Papageorgiou, G. Christopoulos, and S. Smirnakis (InTech) (in press).

Guillot, A., Lebon, F., and Collet, C. (2010). "Electromyographic activity during motor imagery," in *The Neurophysiological Foundations of Mental and Motor Imagery*, eds A. Guillot and C. Collet (New York: Oxford University Press), 83–93.

Guillot, A., Lebon, F., Rouffet, D., Champely, S., Doyon, J., and Collet, C. (2007). Muscular responses during motor imagery as a function of muscle contraction types. *Int. J. Psychophysiol.* 66, 18–27.

Gustin, S. M., Wrigley, P. J., Gandevia, S. C., Middleton, J. W., Henderson, L. A., and Siddall, P. J. (2008). Movement imagery increases pain in people with neuropathic pain following complete thoracic spinal cord injury. *Pain* 137, 237–244.

Gustin, S. M., Wrigley, P. J., Henderson, L. A., and Siddall, P. J. (2010). Brain circuitry underlying pain in response to imagined movement in people with spinal cord injury. *Pain* 148, 438–445.

Hale, B. D. (1982). The effects of internal and external imagery on muscular and ocular concomitants. *J. Sport Psychol.* 4, 379–387.

Hale, B. S., Raglin, J. S., and Koceja, D. M. (2003). Effect of mental imagery of a motor task on the Hoffmann reflex. *Behav. Brain Res.* 142, 81–87.

Hanakawa, T., Dimyan, M. A., and Hallett, M. (2008). Motor planning, imagery, and execution in the distributed motor network: a time-course study with functional MRI. *Cereb. Cortex* 18, 2775–2788.

Hanakawa, T., Immisch, I., Toma, K., Dimyan, M. A., Van Gelderen, P., and Hallett, M. (2003). Functional properties of brain areas associated with motor execution and imagery. *J. Neurophysiol.* 89, 989–1002.

Harris, A. J. (1999). Cortical origin of pathological pain. *Lancet* 354, 1464–1466.

Harris, D. V., and Robinson, W. J. (1986). The effects of skill level on EMG activity during internal and external imagery. *J. Sport Psychol.* 8, 105–111.

Hashimoto, R., and Rothwell, J. C. (1999). Dynamic changes in corticospinal excitability during motor imagery. *Exp. Brain Res.* 125, 75–81.

Helmich, R. C., Bloem, B. R., and Toni, I. (2012). Motor imagery evokes increased somatosensory activity in parkinson's disease patients with tremor. *Hum. Brain Mapp.* 33, 1763–1779.

Helmich, R. C., De Lange, F. P., Bloem, B. R., and Toni, I. (2007). Cerebral compensation during motor imagery in Parkinson's disease. *Neuropsychologia* 45, 2201–2215.

Heremans, E., Feys, P., Nieuwboer, A., Vercruysse, S., Vandenberghe, W., Sharma, N., and Helsen, W. (2011). Motor imagery ability in patients with early- and mid-stage Parkinson disease. *Neurorehabil. Neural Repair* 25, 168–177.

Heremans, E., Helsen, W. F., and Feys, P. (2008). The eyes as a mirror of our thoughts: quantification of motor imagery of goal-directed movements through eye movement registration. *Behav. Brain Res.* 187, 351–360.

Heremans, E., Nieuwboer, A., Feys, P., Vercruysse, S., Vandenberghe, W., Sharma, N., and Helsen, W. F. (2012). External cueing improves motor imagery quality in patients with Parkinson disease. *Neurorehabil. Neural Repair* 26, 27–35.

Holmes, P. S., and Collins, D. J. (2001). The PETTLEP approach to motor imagery: a functional equivalence model for sport psychologists. *J. Appl. Sport Psychol.* 13, 60–83.

Hotz-Boendermaker, S., Funk, M., Summers, P., Brugger, P., Hepp-Reymond, M. C., Curt, A., and Kollias, S. S. (2008). Preservation of motor programs in paraplegics as demonstrated by attempted and imagined foot movements. *Neuroimage* 39, 383–394.

Jackson, P. L., Lafleur, M. F., Malouin, F., Richards, C. L., and Doyon, J. (2003). Functional cerebral reorganization following motor sequence learning through mental practice with motor imagery. *Neuroimage* 20, 1171–1180.

Jacobson, E. (1930). Electrical measurements of neuromuscular states during mental activities. *Am. J. Physiol.* 91, 567–608.

Jacobson, E. (1932). Electrophysiology of mental activities. *Am. J. Psychol.* 44, 677–694.

Jeannerod, M. (1994). The representing brain: neural correlates of motor intention and imagery. *Behav. Brain Sci.* 17, 187–202.

Jeannerod, M. (2001). Neural simulation of action: a unifying mechanism for motor cognition. *Neuroimage* 14, S103–S109.

Jeannerod, M. (2006). *Motor Cognition.* New York: Oxford University Press.

Jeannerod, M., and Frak, V. (1999). Mental imaging of motor activity in humans. *Curr. Opin. Neurobiol.* 9, 735–739.

Johnson, S. H., Sprehn, G., and Saykin, A. J. (2002). Intact motor imagery in chronic upper limb hemiplegics: evidence for activity-independent action representations. *J. Cogn. Neurosci.* 14, 841–852.

Jowdy, D. P., and Harris, D. V. (1990). Muscular responses during mental imagery as a function of motor skill level. *J. Sport Exerc. Psychol.* 12, 191–201.

Kagerer, F. A., Bracha, V., Wunderlich, D. A., Stelmach, G. E., and Bloedel, J. R. (1998). Ataxia reflected in the simulated movements of patients with cerebellar lesions. *Exp. Brain Res.* 121, 125–134.

Karl, A., Birbaumer, N., Lutzenberger, W., Cohen, L. G., and Flor, H. (2001). Reorganization of motor and somatosensory cortex in upper extremity amputees with phantom limb pain. *J. Neurosci.* 21, 3609–3618.

Kasai, T., Kawai, S., Kawanishi, M., and Yahagi, S. (1997). Evidence for facilitation of motor evoked potentials (MEPs) induced by motor imagery. *Brain Res.* 744, 147–150.

Kasess, C. H., Windischberger, C., Cunnington, R., Lanzenberger, R., Pezawas, L., and Moser, E. (2008). The suppressive influence of SMA on M1 in motor imagery revealed by fMRI and dynamic causal modeling. *Neuroimage* 40, 828–837.

Kimberley, T. J., Khandekar, G., Skraba, L. L., Spencer, J. A., Van Gorp, E. A., and Walker, S. R. (2006). Neural substrates for motor imagery in severe hemiparesis. *Neurorehabil. Neural Repair* 20, 268–277.

Kleber, B., Birbaumer, N., Veit, R., Trevorrow, T., and Lotze, M. (2007). Overt and imagiend singing of an Italian aria. *Neuroimage* 36, 1238–1246.

Knecht, S., Henningsen, H., Elbert, T., Flor, H., Hohling, C., Pantev, C., and Taub, E. (1996). Reorganizational and perceptional changes after amputation. *Brain* 119(Pt 4), 1213–1219.

Kokotilo, K. J., Eng, J. J., and Curt, A. (2009). Reorganization and preservation of motor control of the brain in spinal cord injury: a systematic review. *J. Neurotrauma* 26, 2113–2126.

Kosslyn, S. M. (2010). "Multimotal images in the brain," in *The Neurophysiological Foundations of Mental and Motor Imagery*, eds A. Guillot and C. Collet (New York, NY: Oxford University Press), 3–16.

Kosslyn, S. M., Ganis, G., and Thompson, W. L. (2001). Neural foundations of imagery. *Nat. Rev. Neurosci.* 2, 635–642.

Kuhn, A. A., Doyle, L., Pogosyan, A., Yarrow, K., Kupsch, A., Schneider, G. H., Hariz, M. I., Trottenberg, T., and Brown, P. (2006). Modulation of beta oscillations in the subthalamic area during motor imagery in Parkinson's disease. *Brain* 129, 695–706.

Kumru, H., Soto, O., Casanova, J., and Valls-Sole, J. (2008). Motor cortex excitability changes during imagery of simple reaction time. *Exp. Brain Res.* 189, 373–378.

Lacourse, M. G., Cohen, M. J., Lawrence, K. E., and Romero, D. H. (1999). Cortical potentials during imagined movements in individuals with chronic spinal cord injuries. *Behav. Brain Res.* 104, 73–88.

Lafleur, M. F., Jackson, P. L., Malouin, F., Richards, C. L., Evans, A. C., and Doyon, J. (2002). Motor learning produces parallel dynamic functional changes during the execution and imagination of sequential foot movements. *Neuroimage* 16, 142–157.

Lebon, F., Byblow, W. D., Collet, C., Guillot, A., and Stinear, C. M. (2012). The modulation of motor cortex excitability during motor imagery depends on imagery quality. *Eur. J. Neurosci.* 35, 323–331.

Lebon, F., Rouffet, D., Collet, C., and Guillot, A. (2008). Modulation of EMG power spectrum frequency during motor imagery. *Neurosci. Lett.* 435, 181–185.

Lenartowicz, A., Verbruggen, F., Logan, G. D., and Poldrack, R. A. (2011). Inhibition-related activation in the right inferior frontal gyrus in the absence of inhibitory cues. *J. Cogn. Neurosci.* 23, 3388–3399.

Leonard, G., and Tremblay, F. (2007). Corticomotor facilitation associated with observation, imagery and imitation of hand actions: a comparative study in young and old adults. *Exp. Brain Res.* 177, 167–175.

Li, S., Kamper, D. G., Stevens, J. A., and Rymer, W. Z. (2004a). The effect of motor imagery on spinal segmental excitability. *J. Neurosci.* 24, 9674–9680.

Li, S., Latash, M. L., and Zatsiorsky, V. M. (2004b). Effects of motor imagery on finger force responses to transcranial magnetic stimulation. *Brain Res. Cogn. Brain Res.* 20, 273–280.

Liang, N., Ni, Z., Takahashi, M., Murakami, T., Yahagi, S., Funase, K., Kato, T., and Kasai, T. (2007). Effects of motor imagery are dependent on motor strategies. *Neuroreport* 18, 1241–1245.

Liepert, J., Greiner, J., Nedelko, V., and Dettmers, C. (2012). Reduced upper limb sensation impairs mental chronometry for motor imagery after stroke: clinical and electrophysiological findings. *Neurorehabil. Neural Repair* 26, 470–478.

Lim, V. K., Polych, M. A., Hollander, A., Byblow, W. D., Kirk, I. J., and Hamm, J. P. (2006). Kinesthetic but not visual imagery assists in normalizing the CNV in Parkinson's disease. *Clin. Neurophysiol.* 117, 2308–2314.

Livesay, J. R., and Samaras, M. R. (1998). Covert neuromuscular activity of the dominant forearm during visualization of a motor task. *Percept. Mot. Skills* 86, 371–374.

Logan, G. D. (1983). On the ability to inhibit simple thoughts and actions: I. Stop signal studies of decision and memory. *J. Exp. Psychol. Learn. Mem. Cogn.* 9, 585–606.

Logan, G. D. (1985). On the ability to inhibit simple thoughts and actions: II. Stop-signal studies of repetition priming. *J. Exp. Psychol. Learn. Mem. Cogn.* 11, 675–691.

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.

Lorey, B., Bischoff, M., Pilgramm, S., Stark, R., Munzert, J., and Zentgraf, K. (2009). The embodied nature of motor imagery: the influence of posture and perspective. *Exp. Brain Res.* 194, 233–243.

Lotze, M., Flor, H., Grodd, W., Larbig, W., and Birbaumer, N. (2001). Phantom movements and pain. An fMRI study in upper limb amputees. *Brain* 124, 2268–2277.

Lotze, M., Grodd, W., Birbaumer, N., Erb, M., Huse, E., and Flor, H. (1999a). Does use of a myoelectric prosthesis prevent cortical reorganization and phantom limb pain? *Nat. Neurosci.* 2, 501–502.

Lotze, M., Montoya, P., Erb, M., Hulsmann, E., Flor, H., Klose, U., Birbaumer, N., and Grodd, W. (1999b).

Activation of cortical and cerebellar motor areas during executed and imagined hand movements: an fMRI study. *J. Cogn. Neurosci.* 11, 491–501.

Lotze, M., and Halsband, U. (2006). Motor imagery. *J. Physiol. Paris* 99, 386–395.

Lotze, M., Scheler, G., Tan, H. R., Braun, C., and Birbaumer, N. (2003). The musician's brain: functional imaging of amateurs and professionals during performance and imagery. *Neuroimage* 20, 1817–1829.

Lotze, M., and Zentgraf, K. (2010). "Contribution of the primary motor cortex to motor imagery," in *The Neurophysiological Foundations of Mental and Motor Imagery*, eds A. Guillot and C. Collet (New York: Oxford University Press), 31–45.

Louis, M., Collet, C., Champely, S., and Guillot, A. (2012). Differences in motor imagery time when predicting task duration in skiers and equestrian riders. *Res. Q. Exerc. Sport* 83, 86–93.

Luppino, G., Matelli, M., Camarda, R., and Rizzolatti, G. (1994). Corticospinal projections from mesial frontal and cingulate areas in the monkey. *Neuroreport* 5, 2545–2548.

Lutz, R., and Linder, D. E. (2001). Does electromyographic activity during motor imagery predict performance? A test of bioinformational theory and functional equivalences. *J. Sport Exerc. Psychol.* 23, s63.

MacIver, K., Lloyd, D. M., Kelly, S., Roberts, N., and Nurmikko, T. (2008). Phantom limb pain, cortical reorganization and the therapeutic effect of mental imagery. *Brain* 131, 2181–2191.

Macuga, K. L., and Frey, S. H. (2012). Neural representations involved in observed, imagined, and imitated actions are dissociable and hierarchically organized. *Neuroimage* 59, 2798–2807.

Malouin, F., Richards, C. L., Desrosiers, J., and Doyon, J. (2004). Bilateral slowing of mentally simulated actions after stroke. *Neuroreport* 15, 1349–1353.

Malouin, F., Richards, C. L., Durand, A., and Doyon, J. (2008). Clinical assessment of motor imagery after stroke. *Neurorehabil. Neural Repair* 22, 330–340.

Mannix, L. K., Chandurkar, R. S., Rybicki, L. A., Tusek, D. L., and Solomon, G. D. (1999). Effect of guided imagery on quality of life for patients with chronic tension-type headache. *Headache* 39, 326–334.

Marconi, B., Pecchioli, C., Koch, G., and Caltagirone, C. (2007). Functional overlap between hand and forearm motor cortical representations during motor cognitive tasks. *Clin. Neurophysiol.* 118, 1767–1775.

Mayer, A., Kudar, K., Bretz, K., and Tihanyi, J. (2008). Body schema and body awareness of amputees. *Prosthet. Orthot. Int.* 32, 363–382.

Mercier, C., and Sirigu, A. (2009). Training with virtual visual feedback to alleviate phantom limb pain. *Neurorehabil. Neural Repair* 23, 587–594.

Michelon, P., Vettel, J. M., and Zacks, J. M. (2006). Lateral somatotopic organization during imagined and prepared movements. *J. Neurophysiol.* 95, 811–822.

Milton, J., Solodkin, A., Hlustik, P., and Small, S. L. (2007). The mind of expert motor performance is cool and focused. *Neuroimage* 35, 804–813.

Moran, A., Guillot, A., MacIntyre, T., and Collet, C. (2012). Re-imagining motor imagery: building bridges between cognitive neuroscience and sport psychology. *Br. J. Psychol.* 103, 224–247.

Morecraft, R. J., Herrick, J. L., Stilwell-Morecraft, K. S., Louie, J. L., Schroeder, C. M., Ottenbacher, J. G., and Schoolfield, M. W. (2002). Localization of arm representation in the corona radiata and internal capsule in the non-human primate. *Brain* 125, 176–198.

Mulder, T., De Vries, S., and Zijlstra, S. (2005). Observation, imagination and execution of an effortful movement: more evidence for a central explanation of motor imagery. *Exp. Brain Res.* 163, 344–351.

Mulder, T., Zijlstra, S., Zijlstra, W., and Hochstenbach, J. (2004). The role of motor imagery in learning a totally novel movement. *Exp. Brain Res.* 154, 211–217.

Munzert, J., Lorey, B., and Zentgraf, K. (2009). Cognitive motor processes: the role of motor imagery in the study of motor representations. *Brain Res. Rev.* 60, 306–326.

Murphy, S., Nordin, S. M., and Cumming, J. (2008). "Imagery in sport, exercise and dance," in *Advances in Sport Psychology*, ed. T. Horn (Champaign, IL: Human Kinetics), 306–315.

Naito, E., Kochiyama, T., Kidata, R., Nakamura, S., Matsumura, M., Yonekura, Y., and Sadato, N. (2002). Internally simulated movement sensations during motor imagery activate cortical motor areas and the cerebellum. *J. Neurosci.* 22, 3683–3691.

Nico, D., Daprati, E., Rigal, F., Parsons, L., and Sirigu, A. (2004). Left and right hand recognition in upper limb amputees. *Brain* 127, 120–132.

Olsson, C. J. (2012). Complex motor representations may not be preserved after complete spinal cord injury. *Exp. Neurol.* 236, 46–49.

Pascual-Leone, A., Peris, M., Tormos, J. M., Pascual, A. P., and Catala, M. D. (1996). Reorganization of human cortical motor output maps following traumatic forearm amputation. *Neuroreport* 7, 2068–2070.

Patuzzo, S., Fiaschi, A., and Manganotti, P. (2003). Modulation of motor cortex excitability in the left hemisphere during action observation: a single- and paired-pulse transcranial magnetic stimulation study of self- and non-self-action observation. *Neuropsychologia* 41, 1272–1278.

Personnier, P., Paizis, C., Ballay, Y., and Papaxanthis, C. (2008). Mentally represented motor actions in normal aging II. The influence of the gravito-inertial context on the duration of overt and covert arm movements. *Behav. Brain Res.* 186, 273–283.

Porro, C. A., Cettolo, V., Francescato, M. P., and Baraldi, P. (2000). Ipsilateral involvement of primary motor cortex during motor imagery. *Eur. J. Neurosci.* 12, 3059–3063.

Porro, C. A., Francescato, M. P., Cettolo, V., Diamond, M. E., Baraldi, P., Zuiani, C., Bazzocchi, M., and Di Prampero, P. E. (1996). Primary motor and sensory cortex activation during motor performance and motor imagery: a functional magnetic resonance imaging study. *J. Neurosci.* 16, 7688–7698.

Quartarone, A., Bagnato, S., Rizzo, V., Morgante, F., Sant'Angelo, A., Crupi, D., Romano, M., Messina, C., Berardelli, A., and Girlanda, P. (2005). Corticospinal excitability during motor imagery of a simple tonic finger movement in patients with writer's cramp. *Mov. Disord.* 20, 1488–1495.

Raffin, E., Giraux, P., and Reilly, K. T. (2012a). The moving phantom: motor execution or motor imagery? *Cortex* 48, 746–757.

Raffin, E., Mattout, J., Reilly, K. T., and Giraux, P. (2012b). Disentangling motor execution from motor imagery with the phantom limb. *Brain* 135, 582–595.

Ramachandran, V. S., Brang, D., and Mcgeoch, P. D. (2010). Dynamic reorganization of referred sensations by movements of phantom limbs. *Neuroreport* 21, 727–730.

Ranganathan, V. K., Siemionow, V., Liu, J. Z., Sahgal, V., and Yue, G. H. (2004). From mental power to muscle power-gaining strength by using the mind. *Neuropsychologia* 42, 944–956.

Reilly, K. T., Mercier, C., Schieber, M. H., and Sirigu, A. (2006). Persistent hand motor commands in the amputees' brain. *Brain* 129, 2211–2223.

Reis, J., Swayne, O. B., Vandermeeren, Y., Camus, M., Dimyan, M. A., Harris-Love, M., Perez, M. A., Ragert, P., Rothwell, J. C., and Cohen, L. G. (2008). Contribution of transcranial magnetic stimulation to the understanding of cortical mechanisms involved in motor control. *J. Physiol. (Lond.)* 586, 325–351.

Roosink, M., and Zijdewind, I. (2010). Corticospinal excitability during observation and imagery of simple and complex hand tasks: implications for motor rehabilitation. *Behav. Brain Res.* 213, 35–41.

Ross, J. S., Tkach, J., Ruggieri, P. M., Lieber, M., and Lapresto, E. (2003). The mind's eye: functional MR imaging evaluation of golf motor imagery. *AJNR Am. J. Neuroradiol.* 24, 1036–1044.

Rossini, P. M., Rossi, S., Pasqualetti, P., and Tecchio, F. (1999). Corticospinal excitability modulation to hand muscles during movement imagery. *Cereb. Cortex* 9, 161–167.

Rothwell, J. C. (1991). Physiological studies of electric and magnetic stimulation of the human brain. *Electroencephalogr. Clin. Neurophysiol. Suppl.* 43, 29–35.

Roux, F. E., Ibarrola, D., Lazorthes, Y., and Berry, I. (2001). Virtual movements activate primary sensorimotor areas in amputees: report of three cases. *Neurosurgery* 49, 736–741; discussion 741–732.

Roux, F. E., Lotterie, J. A., Cassol, E., Lazorthes, Y., Sol, J. C., and Berry, I. (2003). Cortical areas involved in virtual movement of phantom limbs: comparison with normal subjects. *Neurosurgery* 53, 1342–1352; discussion 1352–1343.

Roy, F. D., Zewdie, E. T., and Gorassini, M. A. (2011). Short-interval intracortical inhibition with incomplete spinal cord injury. *Clin. Neurophysiol.* 122, 1387–1395.

Sabate, M., Gonzalez, B., and Rodriguez, M. (2004). Brain lateralization of motor imagery: motor planning asymmetry as a cause of movement lateralization. *Neuropsychologia* 42, 1041–1049.

Sabate, M., Gonzalez, B., and Rodriguez, M. (2007). Adapting movement planning to motor impairments: the motor-scanning system. *Neuropsychologia* 45, 378–386.

Sabbah, P., De, S. S., Leveque, C., Gay, S., Pfefer, F., Nioche, C., Sarrazin, J. L., Barouti, H., Tadie, M., and Cordoliani, Y. S. (2002). Sensorimotor cortical activity in patients with complete spinal cord injury: a functional magnetic resonance imaging study. *J. Neurotrauma* 19, 53–60.

Samuel, M., Ceballos-Baumann, A. O., Boecker, H., and Brooks, D. J. (2001). Motor imagery in normal subjects and Parkinson's disease patients: an H215O PET study. *Neuroreport* 12, 821–828.

Saturno, E., Bonato, C., Miniussi, C., Lazzaro, V., and Callea, L. (2008). Motor cortex changes in spinal cord injury: a TMS study. *Neurol. Res.* 30, 1084–1085.

Schwoebel, J., Boronat, C. B., and Branch Coslett, H. (2002). The man who executed "imagined" movements: evidence for dissociable components of the body schema. *Brain Cogn.* 50, 1–16.

Sharma, N., Baron, J. C., and Rowe, J. B. (2009a). Motor imagery after stroke: relating outcome to motor network connectivity. *Ann. Neurol.* 66, 604–616.

Sharma, N., Simmons, L. H., Jones, P. S., Day, D. J., Carpenter, T. A., Pomeroy, V. M., Warburton, E. A., and Baron, J. C. (2009b). Motor imagery after subcortical stroke: a functional magnetic resonance imaging study. *Stroke* 40, 1315–1324.

Sharma, N., Jones, P. S., Carpenter, T. A., and Baron, J. C. (2008). Mapping the involvement of BA 4a and 4p during motor imagery. *Neuroimage* 41, 92–99.

Sharma, N., Pomeroy, V. M., and Baron, J. C. (2006). Motor imagery: a backdoor to the motor system after stroke? *Stroke* 37, 1941–1952.

Shaw, W. A. (1938). The distribution of muscular action potentials during imaging. *Psychol. Rec.* 2, 195–216.

Shick, J. (1970). Effects of mental practice on selected volleyball skills for college women. *Res. Q.* 41, 88–94.

Shukla, G. D., Sahu, S. C., Tripathi, R. P., and Gupta, D. K. (1982). Phantom limb: a phenomenological study. *Br. J. Psychiatry* 141, 54–58.

Sirigu, A., Cohen, L., Duhamel, J. R., Pillon, B., Dubois, B., Agid, Y., and Pierrot-Deseilligny, C. (1995). Congruent unilateral impairments for real and imagined hand movements. *Neuroreport* 6, 997–1001.

Sirigu, A., Duhamel, J. R., Cohen, L., Pillon, B., Dubois, B., and Agid, Y. (1996). The mental representation of hand movements after parietal cortex damage. *Science* 273, 1564–1568.

Slade, J. M., Landers, D. M., and Martin, P. E. (2002). Muscular activity during real and imagined movements: a test of inflow explanations. *J. Sport Exerc. Psychol.* 24, 151–167.

Solodkin, A., Hlustik, P., Chen, E. E., and Small, S. L. (2004). Fine modulation in network activation during motor execution and motor imagery. *Cereb. Cortex* 14, 1246–1255.

Stinear, C. M. (2010). "Corticospinal facilitation during motor imagery," in *The Neurophysiological Foundations of Mental and Motor Imagery*, eds A. Guillot and C. Collet (New York, NY: Oxford University Press), 47–61.

Stinear, C. M., and Byblow, W. D. (2003a). Motor imagery of phasic thumb abduction temporally and spatially modulates corticospinal excitability. *Clin. Neurophysiol.* 114, 909–914.

Stinear, C. M., and Byblow, W. D. (2003b). Role of intracortical inhibition in selective hand muscle activation. *J. Neurophysiol.* 89, 2014–2020.

Stinear, C. M., and Byblow, W. D. (2004). Modulation of corticospinal excitability and intracortical inhibition during motor imagery is task-dependent. *Exp. Brain Res.* 157, 351–358.

Stinear, C. M., Byblow, W. D., Steyvers, M., Levin, O., and Swinnen, S. P. (2006). Kinesthetic, but not visual, motor imagery modulates corticomotor excitability. *Exp. Brain Res.* 168, 157–164.

Stinear, C. M., Coxon, J. P., and Byblow, W. D. (2009). Primary motor cortex and movement prevention: where Stop meets Go. *Neurosci. Biobehav. Rev.* 33, 662–673.

Stinear, C. M., Fleming, M. K., Barber, P. A., and Byblow, W. D. (2007). Lateralization of motor imagery following stroke. *Clin. Neurophysiol.* 118, 1794–1801.

Subramanian, L., Hindle, J. V., Johnston, S., Roberts, M. V., Husain, M., Goebel, R., and Linden, D. (2011). Real-time functional magnetic resonance imaging neurofeedback for treatment of Parkinson's disease. *J. Neurosci.* 31, 16309–16317.

Suinn, R. M. (1980). "Body thinking: psychology for Olympic champions," in *Psychology in Sports: Methods and Applications*, ed. R. M. Suinn (Mineapolis: Burgess), 306–315.

Szameitat, A. J., Shen, S., Conforto, A., and Sterr, A. (2012). Cortical activation during executed, imagined, observed, and passive wrist movements in healthy volunteers and stroke patients. *Neuroimage* 62, 266–280.

Tamir, R., Dickstein, R., and Huberman, M. (2007). Integration of motor imagery and physical practice in group treatment applied to subjects with Parkinson's disease. *Neurorehabil. Neural Repair* 21, 68–75.

Taniguchi, S., Kimura, J., Yamada, T., Ichikawa, H., Hara, M., Fujisawa, R., Shimizu, H., and Tani, T. (2008). Effect of motion imagery to counter rest-induced suppression of F-wave as a measure of anterior horn cell excitability. *Clin. Neurophysiol.* 119, 1346–1352.

Thobois, S., Dominey, P., Fraix, V., Mertens, P., Guenot, M., Zimmer, L., Pollak, P., Benabid, A. L., and Broussolle, E. (2002). Effects of subthalamic nucleus stimulation on actual and imagined movement in Parkinson's disease: a PET study. *J. Neurol.* 249, 1689–1698.

Thobois, S., Dominey, P. F., Decety, J., Pollak, P. P., Gregoire, M. C., Le Bars, P. D., and Broussolle, E. (2000). Motor imagery in normal subjects and in asymmetrical Parkinson's disease: a PET study. *Neurology* 55, 996–1002.

Topka, H., Cohen, L. G., Cole, R. A., and Hallett, M. (1991). Reorganization of corticospinal pathways following spinal cord injury. *Neurology* 41, 1276–1283.

Tremblay, F., Leonard, G., and Tremblay, L. (2008). Corticomotor facilitation associated with observation and imagery of hand actions is impaired in Parkinson's disease. *Exp. Brain Res.* 185, 249–257.

Verbruggen, F., and Logan, G. D. (2009). Models of response inhibition in the stop-signal and stop-change paradigms. *Neurosci. Biobehav. Rev.* 33, 647–661.

Vromen, A., Verbunt, J. A., Rasquin, S., and Wade, D. T. (2011). Motor imagery in patients with a right hemisphere stroke and unilateral neglect. *Brain Inj.* 25, 387–393.

Watanabe, J., Sugiura, M., Sato, K., Sato, Y., Maeda, Y., Matsue, Y., Fukuda, H., and Kawashima, R. (2002). The human prefrontal and parietal association cortices are involved in NO-GO performances: an event-related fMRI study. *Neuroimage* 17, 1207–1216.

Wehner, T., Vogt, S., and Stadler, M. (1984). Task-specific EMG-characteristics during mental training. *Psychol. Res.* 46, 389–401.

Wolf, S. L., Blanton, S., Baer, H., Breshears, J., and Butler, A. J. (2002). Repetitive task practice: a critical review of constraint-induced movement therapy in stroke. *Neurologist* 8, 325–338.

Wu, A. J., Hermann, V., Ying, J., and Page, S. J. (2010). Chronometry of mentally versus physically practiced tasks in people with stroke. *Am. J. Occup. Ther.* 64, 929–934.

Yahagi, S., and Kasai, T. (1998). Facilitation of motor evoked potentials (MEPs) in first dorsal interosseous (FDI) muscle is dependent on different motor images. *Electroencephalogr. Clin. Neurophysiol.* 109, 409–417.

Yahagi, S., Shimura, K., and Kasai, T. (1996). An increase in cortical excitability with no change in spinal excitability during motor imagery. *Percept. Mot. Skills* 83, 288–290.

Yue, G., and Cole, K. J. (1992). Strength increases from the motor program: comparison of training with maximal voluntary and imagined muscle contractions. *J. Neurophysiol.* 67, 1114–1123.

Zatsiorsky, V. M., Li, Z. M., and Latash, M. L. (2000). Enslaving effects in multi-finger force production. *Exp. Brain Res.* 131, 187–195.

Zijdewind, J., Toering, S. T., Bessen, B., Van-Der-Laan, O., and Diercks, R. L. (2003). Effects of imagery motor training on torque production of ankle plantar flexor muscles. *Muscle Nerve* 28, 168–173.

Zimmermann-Schlatter, A., Schuster, C., Puhan, M. A., Siekierka, E., and Steurer, J. (2008). Efficacy of motor imagery in post-stroke rehabilitation: a systematic review. *J. Neuroeng. Rehabil.* 5, 8.

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

Received: 29 March 2012; accepted: 10 August 2012; published online: 05 September 2012.

Citation: Guillot A, Di Rienzo F, MacIntyre T, Moran A and Collet C (2012) Imagining is not doing but involves specific motor commands: a review of experimental data related to motor inhibition. Front. Hum. Neurosci. **6**:*247. doi: 10.3389/fnhum.2012.00247*

frontiers in
HUMAN NEUROSCIENCE

REVIEW ARTICLE
published: 28 November 2012
doi: 10.3389/fnhum.2012.00314

Mental imagery of speech: linking motor and perceptual systems through internal simulation and estimation

Xing Tian* and David Poeppel

Poeppel Lab, Department of Psychology, New York University, New York, NY, USA

Edited by:
Joel Pearson, University of New South Wales, Australia

Reviewed by:
Arthur M. Jacobs, Freie Universität Berlin, Berlin
Joel Pearson, University of New South Wales, Australia

****Correspondence:***
Xing Tian, Poeppel Lab, Department of Psychology, New York University, 6 Washington Pl, Suite 275, New York, NY 10003, USA.
e-mail: xing.tian@nyu.edu

The neural basis of mental imagery has been investigated by localizing the underlying neural networks, mostly in motor and perceptual systems, separately. However, how modality-specific representations are top-down induced and how the action and perception systems interact in the context of mental imagery is not well understood. Imagined speech production ("articulation imagery"), which induces the kinesthetic feeling of articulator movement and its auditory consequences, provides a new angle because of the concurrent involvement of motor and perceptual systems. On the basis of previous findings in mental imagery of speech, we argue for the following regarding the induction mechanisms of mental imagery and the interaction between motor and perceptual systems: (1) Two distinct top-down mechanisms, *memory retrieval* and *motor simulation*, exist to induce estimation in perceptual systems. (2) Motor simulation is sufficient to internally induce the representation of perceptual changes that would be caused by actual movement (perceptual associations); however, this simulation process only has modulatory effects on the perception of external stimuli, which critically depends on context and task demands. Considering the proposed simulation-estimation processes as common mechanisms for interaction between motor and perceptual systems, we outline how mental imagery (of speech) relates to perception and production, and how these hypothesized mechanisms might underpin certain neural disorders.

Keywords: internal forward model, efference copy, corollary discharge, sensory-motor integration, mirror neurons, auditory hallucination, stuttering, phantom limb

INTRODUCTION

Mental imagery can be characterized as a quasi-perceptual experience, induced in the absence of external stimulation. Neuroimaging studies have shown that common neural substrates mediate mental imagery and the corresponding perceptual processes, such as in visual (e.g., Kosslyn et al., 1999; O'Craven and Kanwisher, 2000), auditory (e.g., Zatorre et al., 1996; Kraemer et al., 2005), somatosensory (e.g., Yoo et al., 2003; Zhang et al., 2004), and olfactory domains (e.g., Bensafi et al., 2003; Djordjevic et al., 2005). The demonstration of activation in corresponding perceptual regions during mental imagery has provided strong evidence to support the claim that the perceptual experience during mental imagery is mediated by modality-specific neural representations (see the review by Kosslyn et al., 2001). However, the top-down "induction mechanism" for the neural activity mediating mental imagery is not well understood.

We focus here on the role of the motor system in the construction of perceptual experience in mental imagery. We propose a motor-based mechanism that is an alternative (additional) mechanism to Kosslyn's memory-attention-based account (Kosslyn, 1994, 2005; Kosslyn et al., 1994): planned action is simulated in motor systems to internally derive the representation of perceptual changes that would be caused by the actual action (perceptual associations). We suggest that the deployment of these two distinct mechanisms depends on task demands and contextual influence. Studies of mental imagery of speech are summarized to provide evidence for the proposed account—and for the coexistence of both mechanisms. We discuss the motor-to-sensory integration process and propose some working hypotheses regarding certain neural and neuropsychiatric disorders from the perspective of the proposed internal simulation and estimation mechanisms.

DIFFERENT ROUTES FOR INDUCING MENTAL IMAGES

MENTAL IMAGERY OF PERCEPTION AS MEMORY RETRIEVAL (DIRECT SIMULATION)

Mental imagery has been proposed to be essentially a memory retrieval process. That is, perceptual experience is simulated by reconstructing stored perceptual information in modality-specific cortices (Kosslyn, 1994, 2005; Kosslyn et al., 1994). In particular, the process, guided by attention, retrieves object and spatial properties stored in long-term memory to reactivate the topographically organized sensory cortices that represent the object features. Through top-down (re)construction of the neural representation that is similar to the result of bottom-up perceptual processes, the perceptual experience can be re-elicited without the presence of any physical stimuli during mental imagery. This attention-guided memory retrieval process has been demonstrated, for example, in the visual imagery of faces (Ishai et al., 2002).

Mental imagery is further hypothesized to be a predictive process (for future perceptual states), in which the dynamics of perceptual experience can be retrieved/calculated and reconstructed internally (Moulton and Kosslyn, 2009). That is, given an initial point, the series of future perceptual states can be internally simulated by following the regularity (temporal and causal constraints) stored in declarative memory (general knowledge). The mapping between internal simulation and the perception of external stimulation is thought not to be necessarily isomorphic (Goldman, 1989), as only the essential intermediate states are required to have a one-to-one mapping (Fisher, 2006). Because this proposed simulation process is executed entirely within perceptual domains on the basis of memory retrieval—without any representational transformation between motor and perceptual systems—we refer to this account as *direct simulation*.

MENTAL IMAGERY OF MOTOR ACTION AS ESTIMATION DERIVING FROM SIMULATION

Motor imagery is thought to be the process that internally simulates planned actions, by activating similar neural substrates that mediate motor intention and preparation (Jeannerod, 1995, 2001; Decety, 1996). Numerous studies have demonstrated both frontal and parietal activity during motor imagery (Decety et al., 1994; Lotze et al., 1999; Gerardin et al., 2000; Ehrsson et al., 2003; Hanakawa et al., 2003; Dechent et al., 2004; Meister et al., 2004; Nikulin et al., 2008). However, motor system activation does not necessarily link to the kinesthetic feeling generated during motor imagery. The residual neural activity, resulting from the absence of external somatosensory feedback, is thought to mediate the kinesthetic experience during motor imagery (Jeannerod, 1994, 1995). The implicit assumption of the "residual activity account" is that the internal motor simulation during imagery should be transformed into the same representational format as the one resulting from somatosensory feedback. That is, the somatosensory consequences of motor simulation should be estimated. This is consistent with the view that parietal rather than frontal motor regions mediate motor awareness (Desmurget and Sirigu, 2009). In support of the claim that parietal regions mediate somatosensory estimation, direct current stimulation over parietal cortex induces false belief of movement (Desmurget et al., 2009); parietal lesions also impaired the temporal precision of performing motor imagery tasks (Sirigu et al., 1996). Cumulatively, the results suggest that motor simulation in frontal cortex converges in parietal regions to form a kinesthetic representation.

The internal transformation between motor simulation and somatosensory estimation has been proposed in the context of internal forward models in the motor control literature [see the review by Wolpert and Ghahramani (2000)]. The core presupposition is that the neural system can predict the perceptual consequences by internal simulating a copy of a planned action command (the efference copy). Mental imagery has been linked to the concept of internal forward models by the argument that the subjective feeling in mental imagery is the result of the internal estimation of the perceptual consequences following the internal simulation of an action (Grush, 2004). Consistent with this hypothesis, we propose here that the kinesthetic feeling in motor imagery is the result of somatosensory estimation, derived from internal simulation that closely mimics the dynamics of a motor action. We refer to this account as *motor simulation and estimation*.

The *motor simulation and estimation* account differs from the *direct simulation* (memory retrieval) account in that it requires a transformation between motor and somatosensory systems. Our question here, though, extends beyond this: can a motor simulation deliver perceptual consequences that extend to other sensory domains (such as visual and auditory) as well? If so, internal simulation and estimation processes would serve as an additional path to induce modality-specific neural representations similar to the ones induced on the basis of memory retrieval. In the next section, we discuss this possibility in the framework of internal forward models and propose a *sequential simulation and estimation* account. We will use the interaction of motor, somatosensory, and auditory systems in speech production as an example to illustrate such internal cascaded processes, which can generalize to other sensory domains.

MENTAL IMAGERY OF SPEECH AS SEQUENTIAL ESTIMATION

Perception and production systems are functionally connected: perceptual systems analyze the sensory input generated by self actions; the motor system is also regulated by perceptual feedback to perform updates on actions in the future. For example, when people talk, they move their articulators, feel the movement, and hear the self-produced speech that can be used to detect and correct any pronunciation errors. The temporal sequence of physical articulation, proprioception of the articulators, and auditory perception of one's own vocalization makes it possible—on the basis of co-occurrence and associative learning during development—to create internal connections among the neural processes that mediate motor action, somatosensory feedback, and auditory perception. After establishing the connections, motor commands can cycle internally through somatosensory regions and "reach" auditory regions. That is, the estimation in the somatosensory system can serve as a link between motor and auditory systems. Theoretically, such a cascaded estimation architecture has been hypothesized by Hesslow (2002). Anatomically and functionally, the connections between parietal regions and auditory temporal regions have also been demonstrated (Schroeder et al., 2001; Foxe et al., 2002; Fu et al., 2003).

On the basis of recent neurophysiological (MEG) studies, we proposed that a process of auditory inference after somatosensory estimation occurs during *overt speech processing* [**Figure 1**; adapted from Tian and Poeppel (2010)]. Specifically, the estimation of auditory consequences relies on the somatosensory estimation that derives from the simulation of planned action. That is, the internal *auditory* prediction is the result of a coordinate transformation from the somatosensory to the auditory domain. This sequential estimation mechanism (motor plan → somatosensory estimation → auditory prediction/estimation) can derive detailed auditory predictions that are then compared with auditory feedback for self-monitoring and online control.

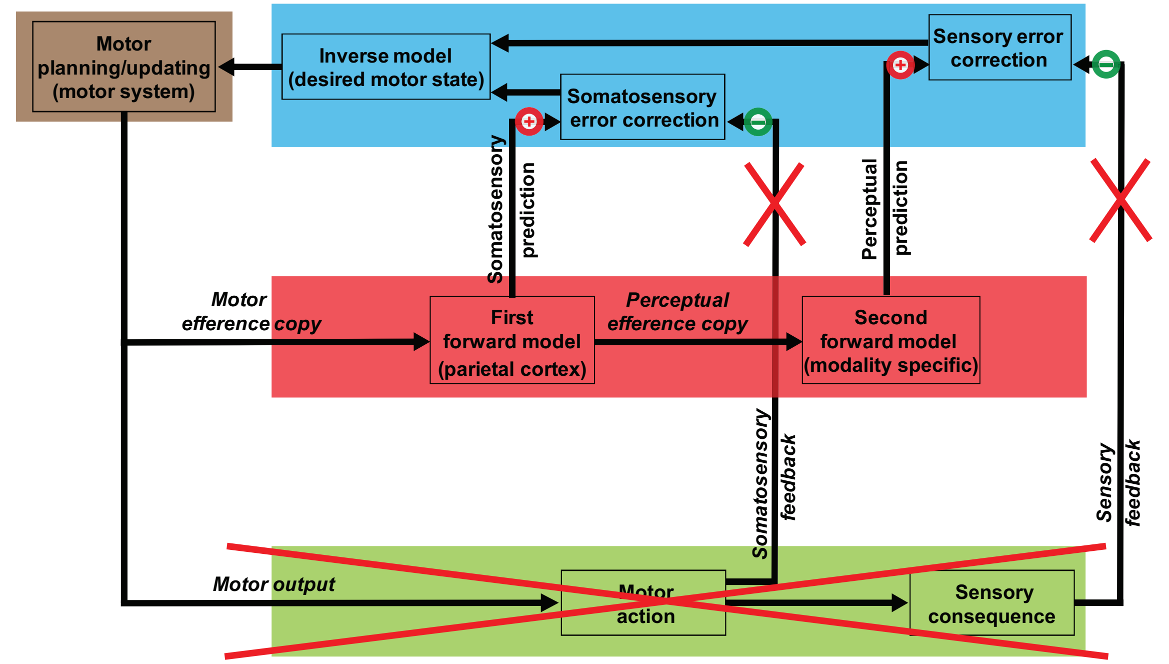


FIGURE 1 | Model of speech processing and its implication for mental imagery of speech. The internal simulation and estimation model proposed as a second route to generate mental images. The motor systems that mediate action preparation carry out the same functions in mental imagery of speech, but only perform motor simulation, in the sense that the planned motor commands are truncated along the path to primary motor cortex and are not executed (the red cross over external outputs). A copy of such planned motor commands (motor efference copy) is processed internally and is used to estimate the associated somatosensory consequences. A copy of the somatosensory estimation is further sent to modality-specific areas, and the associated perceptual consequences that would be produced by the overt action are estimated. The quasi-perceptual experience during mental imagery (the feeling of movement of the articulators and the feeling of auditory perception in the case of articulation imagery) is the result of residual activity from these internal estimation processes, because of the absence of cancellation from the external feedback (the red crosses over external somatosensory and perceptual feedback).

In the case of the *mental imagery of speech*, we propose that the quasi-perceptual experience of articulator movement and the subsequent auditory percept are induced by the same sequential estimation mechanism. However, the "cancellation" deriving from somatosensory and auditory feedback, which is generated by the overt outputs during production, is absent in the imagery case (**Figure 1**). Therefore, similar to the case of motor imagery (Jeannerod, 1994, 1995), the feeling of articulator movement is the result of residual somatosensory representation resulting from motor simulation; the subsequent auditory perceptual experience, we suggest, is the residual auditory representation from the second estimation stage.

On the basis of sequential estimation account, particular neural activity patterns for the two sequential estimates are predicted to occur in a temporal order. Specifically, an auditory pattern should follow a somatosensory one during mental imagery of speech. Applying a novel multivariate technique (Tian and Huber, 2008; Tian et al., 2011) to MEG data, we observed such a temporal order for somatosensory and auditory estimations during articulation imagery (Tian and Poeppel, 2010), manifested in the sequential activity patterns over modality-specific regions at different latencies (**Figure 2**). A left parietal response pattern was observed during *articulation imagery* at the same latency as when motor responses occurred in the articulation condition[1]. Following such a left parietal response pattern, a second pattern was identified at a latency of 150–170 ms after the parietal response. This second pattern was very similar to the response elicited by external auditory stimuli. Moreover, in a further experimental condition, *hearing imagery*, we also observed an auditory-like neural response pattern; however, its latency was faster than the same auditory pattern observed in *articulation imagery*. The existence of these two spatially highly similar auditory-like neural representations, with different latencies for *articulation* versus *hearing* imagery tasks, suggests that the same (or strikingly similar) neural representations can be generated either by internal estimation or by memory retrieval, based on contextual variation and task demands.

Note that the auditory estimation is presumably formed along the canonical auditory hierarchy, but the induction process will be in reversed order. That is to say, the abstract representation is (re-)constructed first in higher level associative areas and conveyed to a perceptual-sensory representation in lower areas. The observation of neural activity in the posterior superior temporal

---

[1] Because previous findings suggest that the time courses for completing execution and imagery are comparable (Decety and Michel, 1989; Decety et al., 1989; Sirigu et al., 1995, 1996), the observed neural responses over parietal regions presumably mediate somatosensory estimation.

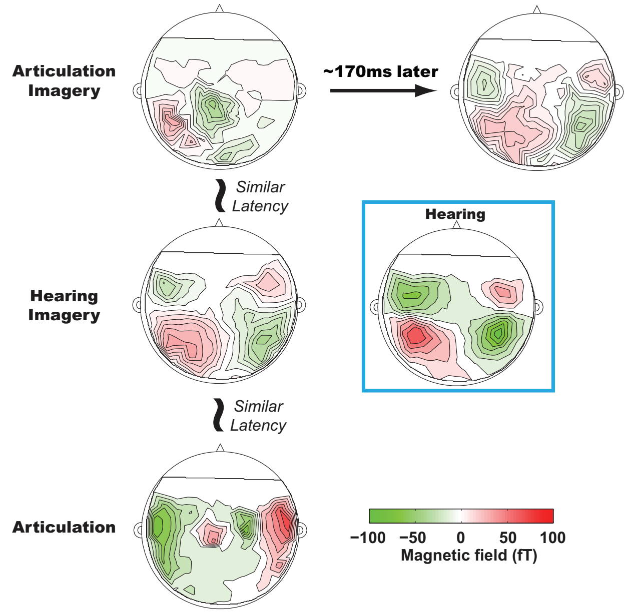


FIGURE 2 | Results from Tian and Poeppel (2010). The sequential estimation during articulation imagery revealed by MEG recordings. All plots are MEG topographies (response patterns) when participants actually speak (lower row), imagine hearing (middle row), and imagine speaking (top row). The activity patterns in the first column are temporally aligned with the onset of articulation movement. At a similar latency, bilateral frontal, bilateral temporal, and left lateralized parietal activity patterns are observed in articulation, hearing imagery, and articulation imagery conditions. In articulation imagery, about 150–170 ms later after the parietal activity, bilateral temporal activity is also observed. All the bilateral temporal activity patterns in *hearing imagery* and *articulation imagery* resemble the topography of the auditory response during actual hearing (highlighted in a blue box, response pattern when participants listen to the same auditory stimuli as in other conditions).

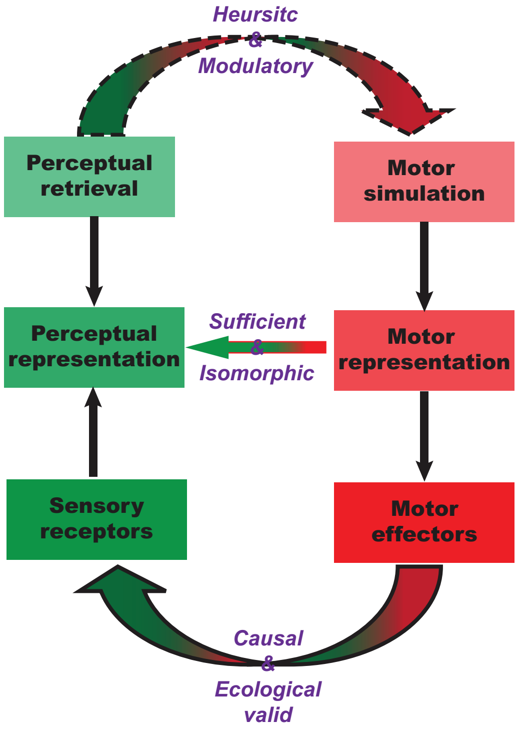


FIGURE 3 | Sufficiency and necessity between motor simulation and perceptual estimation. The characteristics of the proposed motor simulation and perceptual estimation processes, and the nature of motor involvement during perceptual tasks. The internal motor simulation can take a similar path as motor preparation to derive a corresponding motor representation that in turn derives associated perceptual representations in a one-to-one fashion. Such one-to-one mapping is the same as the one in the external connections between the similar motor action and perceptual consequences. In the other direction, when the perceptual representation is needed, different paths can be taken. It can rely on memory retrieval to directly recreate the perceptual representation. It can also take another less demanding path that relies on the motor simulation to derive the associated perceptual representation.

sulcus (pSTS) during silent speaking (Price et al., 2011) could be the result of an earlier reconstruction. Whereas the observations of similarity between responses to mental imagery and to external stimulation, such as in visual (e.g., Kosslyn et al., 1999) and auditory (e.g., **Figure 2**, Tian and Poeppel, 2010) domains, are the results of process continuation to lower perceptual-sensory regions. How much further back the reconstruction process might go seems to depend on the sensory modality and demands of the imagery tasks (Kosslyn and Thompson, 2003; Kraemer et al., 2005; Zatorre and Halpern, 2005).

INTERNAL SIMULATION-ESTIMATION AND RELATION TO SENSORY-MOTOR INTEGRATION

Mental imagery of speech exemplifies a top-down mechanism for sensory-motor integration. The proposal here is motor simulation and sequential estimation. In the first part of this section, we describe the nature of this sequential transformation between motor, somatosensory, and other perceptual systems. We postulate that there is a one-to-one transformation between motor simulation and somatosensory estimation, as well as isomorphic mapping between somatosensory estimation and subsequent perceptual estimation (**Figure 3**). The entire transformation process is carried out in a continuous manner, beginning with motor simulation, then somatosensory estimation, and ending with modality-specific perceptual estimation. In the second part of this section, we argue that the implementation of motor simulation depends on context and task demands and may only exert modulatory effects on perception.

MOTOR-TO-SENSORY MAPPING: ISOMORPHISM VIA ESTABLISHED CONNECTIONS

The central idea underpinning motor simulation and subsequent perceptual estimation is the conjectured *one-to-one mapping* or *isomorphism* between mental and physical processes. This isomorphism has been proposed for motor simulation (Jeannerod, 1994) and visual mental rotation (Shepard and Cooper, 1986): the intermediate stages of the internal process must have a one-to-one correspondence to intermediate stages of an actualized physical process. We extend this isomorphism to the associations between the motor simulation and perceptual estimation: the one-to-one mapping between the trajectory of motor simulation and perceptual estimation is a close analog to the causal relation between motor outputs and perceptual changes. That is, not only should the starting and ending points of an action simulation lead to the initiation and results of perceptual estimation, but intermediate points on this action simulation trajectory should result in a sequence of perceptual estimates, even though no external signals are physically presented. Notice that the analogy between internal simulation-estimation and external action-perception does not require the preservation of first-order isomorphism: only the one-to-one relation in the transformation of internal representation from motor to perceptual systems is required, as if the action was actually performed and the percept was actually induced.

The isomorphic transformation from motor to perceptual systems relies on the established internal associations between motor and perceptual representations, which are presumably formed following the causal and ecologically valid sequential occurrence of action-perception pairs, through the mechanisms of associative learning (Mahon and Caramazza, 2008). For example, the movement of articulators can induce somatosensory feedback and subsequent auditory perception of one's own speech. On the basis of the occurrence order (action first, then somatosensory activation, followed by auditory perception), an internal association can be established to link a particular movement trajectory of articulators with the specific somatosensory sensation, followed by a given auditory perception of speech. Note that we do *not* exclude the possible existence of a parallel estimation process that links motor simulation to somatosensory and auditory systems separately (Guenther et al., 2006; Price et al., 2011). Such an additional mechanisms which runs in parallel may mediate the early comparison between auditory estimation from an articulatory plan and intended auditory targets during speech production (Hickok, 2012). The redundancy of the compensation in somatosensory and auditory domains offers a hint for the co-existence of sequential and parallel estimation structures (Lametti et al., 2012). We suggest that the serial updating structure as one of the possible underlying estimation mechanisms naturally follows the biological sequences, providing advantages in learning and plasticity during development as well as online speech control.

Speech-induced suppression and enhancement caused by feedback perturbation provides strong evidence for the one-to-one mapping between motor simulation and estimation of perceptual consequences. When participants speak and listen to their own speech, the evoked auditory responses are smaller compared with the auditory responses to the same speech played back without spoken outputs (Numminen et al., 1999; Houde et al., 2002; Eliades and Wang, 2003, 2005; Ventura et al., 2009). However, when the auditory feedback is perturbed (manipulating, e.g., pitch or format frequencies), the auditory responses during speaking become larger compared with the ones during playback (Eliades and Wang, 2008; Tourville et al., 2008; Zheng et al., 2010; Behroozmand et al., 2011). The suppression caused by articulation demonstrates that an internal signal labels the onset of movement and down-regulates sensitivity to subsequent auditory perception (general suppression). However, the enhancement caused by feedback perturbation suggests that the internal signal during articulation is not a generic gain control mechanism for all auditory stimuli, but rather provides a precise perceptual prediction and only blocks the feedback that is identical to the prediction. In other words, there is a one-to-one mapping between motor simulation and auditory estimation, and the precise auditory consequence can be predicted based on particular motor trajectory.

The hypothesized intermediate neurocomputational step of somatosensory estimation that lies between motor simulation and auditory estimation has also been suggested by recent experiments. The sequential neural activity underlying somatosensory and auditory estimation has been observed during articulation imagery using MEG (Tian and Poeppel, 2010), as discussed above (**Figure 2**). Lesions over the left pars opercularis (pOp) in the inferior frontal gyrus (IFG) as well as adjacent to the left supramarginal gyrus (SMG) in parietal cortex correlate with the ability to imagine speech; this demonstrates the possible neural implementation underlying the proposed simulation and (somatosensory) estimation (Geva et al., 2011). Moreover, the causal role of somatosensory feedback in speech perception has also been demonstrated (Ito et al., 2009). There, participants were asked to listen to ambiguous stimuli (e.g., head-had vowel continuum) while their facial skin was manipulated with a robotic device. When the skin at the side of mouth was stretched upward (as in the case of pronouncing "head"), participants were biased toward hearing the ambiguous sound as "head." That is, the somatosensory status affected the auditory perception in a systematic way: there was a one-to-one representational mapping between somatosensory and auditory systems.

THE SIMULATION-ESTIMATION PROCESS IN PERCEPTION

The debates surrounding motor theories of perception and cognition [see the review by Scheerer (1984)] have heated up since the discovery of the putative "mirror neuron system" in monkeys (di Pellegrino et al., 1992; Gallese et al., 1996; see Rizzolatti and Craighero, 2004 for a review) and the observation of motor activity observed during numerous perceptual studies in humans (e.g., Rizzolatti et al., 1996; Iacoboni et al., 1999; Buccino et al., 2001; Wilson et al., 2004). Although these debates are beyond the scope of this review, the proposed mechanism of sequential estimation following motor simulation may provide insight to reconcile some of the observations, providing a top-down perspective.

We propose, building on arguments in the recent literature (Mahon and Caramazza, 2008; Hickok, 2009; Lotto et al., 2009; Rumiati et al., 2010), that the deployment of motor simulation in perceptual tasks is (1) strategy-dependent and (2) exerts modulatory effects on the formation of perceptual representations. That is, the selection of motor involvement in perceptual tasks depends on context and task demands. It is a top-down strategic step to provide modality-specific representations in advance (cf. Moulton and Kosslyn, 2009) and reduce perceptual variance by generating more precise estimation (Mahon and Caramazza, 2008; but also see Pulvermüller and Fadiga, 2010 for an opposite view from a embodied perspective).

The implementation of motor-to-sensory transformations is strategy-dependent

We describe two types of evidence. First, the recruitment/involvement of motor simulation is influenced by task demands. For example, motor imagery can be performed from a "first person" perspective that relies on kinesthetic feeling, in contrast with when a task is executed from a "third person" perspective in which the action-related visual changes are recreated (Jeannerod, 1994, 1995). Reaction times of hand rotation imagery showed an interaction between imagery perspectives and limb posture: when asked to imagine rotating their hands from first person perspective, participants responded faster when their hands were on the lap but slower when their hands are in the back; the reverse pattern was observed when imagining from third person perspective (Sirigu and Duhamel, 2001). Activation in the motor system was observed when participants were explicitly told to imagine rotating an object with their own hands, but was absent when they were told to imagine rotating the same object with a robotic motor (Kosslyn et al., 2001). Both behavioral and neuroimaging results highlight that the task demands influence the implementation of neural pathways that mediate either direct simulation (memory retrieval) or motor simulation-estimation (transformation between motor and perceptual systems).

Second, motor-to-sensory transformations are influenced by context and the properties of stimuli. For example, neural responses in frontal motor regions have been observed during observation of meaningful actions, contrasted with occipital activity for meaningless actions (Decety et al., 1997). Relatedly, when participants mentally rotated their hands, premotor, primary motor, and posterior parietal cortices were activated. However, frontal motor areas were silent when they mentally rotated objects (Kosslyn et al., 1998). These results suggest that contextual influence and task demands can determine the implementation of motor simulation in a top-down, voluntary, strategic way.

In the context of action observation, understanding/comprehension and imitation could be the result of heuristic engagement of motor simulation. That is, humans can deploy a top-down mechanism that transfers perceptual goals into the motor domain and initiates motor simulation to derive perceptual consequences (**Figure 3**). The strategic and heuristic initiation of motor involvement can be considered as a top-down mental imagery process (possibly exclusive to humans) (cf. Iacoboni et al., 1999; Papeo et al., 2009), wherein the motor action is internally simulated and perceptual consequences estimated thereafter (cf. Tkach et al., 2007).

Modulatory function of motor simulation on perception

The major evidence supporting a modulatory role of motor simulation in perception (rather than a primary causal role) comes from lesion studies. For example, lesions in the frontal lobe only caused deficits in action production, whereas lesions in the parietal lobe caused deficits both during production and perception of movement (Heilman et al., 1982). A deficit in gesture recognition has also been linked to inferior parietal cortex lesions but not lesions in the frontal lobe (Buxbaum et al., 2005). Action comprehension also relies on a network that includes inferior parietal cortex but not IFG (Saygin et al., 2004). Although patients with IFG lesions demonstrated deficits in action comprehension in the same study, the static stimuli (pictures of pantomimed actions or objects) could require participants to implement the strategy of motor simulation to form the dynamic display of action and to derive the perceptual consequences so that they can fulfill the action-object association task. Such lesion results indicate that a damaged motor system (and the deficits in motor simulation) dissociates from action-perception and comprehension. The abstract meaning of motor action is probably "stored" in parietal regions, and the motor simulation mediated by frontal regions is one of many paths to access the stored representation (in line with our proposed *simulation* over frontal cortex and *estimation* over parietal cortex). Therefore, motor simulation to estimate perceptual consequences is only modulatory and not necessary for perceptual tasks.

Analogous to the advantage of multisensory integration in minimizing perceptual variance (Ernst and Banks, 2002; Alais and Burr, 2004; van Wassenhove et al., 2005; von Kriegstein and Giraud, 2006; Morgan et al., 2008; Poeppel et al., 2008; Fetsch et al., 2009), the modulatory effects of motor simulation convey benefits by providing additional, more detailed information to enrich the perceptual representation using internal sequential estimation mechanism (cf. Mahon and Caramazza, 2008). Human observers can adopt motor strategies to provide more precise perceptual representations and deal with perceptual ambiguity, for example in the case of speech perception. That is, the motor simulation and estimation can provide improved priors to reduce perceptual variance.

In summary, various perceptual tasks can use the motor system to derive perceptual consequences, by implementing the same top-down motor simulation and perceptual estimation mechanism, as in mental imagery of speech. We hypothesize that this motor simulation is modulatory and only serves as one of many possible corridors to induce perceptual representations. Such strategies of sensory-to-motor and motor-to-sensory transformation would be implemented depending on task demands and contextual influence.

IMPLICATIONS FOR THE NEURAL CORRELATES OF SOME DISORDERS

In this section we argue that the internal processes of motor simulation and estimation, revealed originally for the mental imagery of speech, can shed light on possible neural correlates of certain disorders, including auditory hallucinations, stuttering, and phantom limb syndrome. We outline some working hypotheses regarding these disorders, complementing other existing hypotheses. It is suggested that the proposed idea for mental imagery generation, motor simulation, and sequential perceptual estimation, points to the practical value of mental imagery research for understanding the internal mechanisms of such neural disorders.

AUDITORY HALLUCINATIONS: INTACT ESTIMATION VERSUS BROKEN MONITORING

Internal simulation and sequential estimation has been proposed to be a way to distinguish between the perceptual changes caused by self-generated actions and exogenous external events (Blakemore and Frith, 2003; Jeannerod and Pacherie, 2004; Tsakiris and Haggard, 2005). The perceptual consequences of intended movement can be predicted, and the processing of external sensory feedback can be dampened by the internal prediction, such as in the case of speech production (e.g., Houde et al., 2002; Eliades and Wang, 2003, 2005) and somatosensory perception in tickling (e.g., Blakemore et al., 1998). This suggests that the action-induced perceptual signals are identified as self-generated and cancelled by the virtually identical representation generated by internal perceptual prediction. However for patients suffering from auditory hallucinations, deficits of these hypothesized dampening mechanisms for self-induced perceptual changes have been observed in both somatosensory (e.g., Blakemore et al., 2000) and auditory (e.g., Ford et al., 2007; Heinks-Maldonado et al., 2007) domains. These results suggest that patients with auditory hallucinations cannot separate self-induced from external-induced perceptual signals.

Critically, deficits of distinguishing self-induced from externally induced perceptual changes are not enough to account for auditory hallucinations, because the positive symptoms typically occur in the absence of any external stimuli. There must exist an internal mechanism to induce the auditory representations that are then misattributed to an external source/voice. In fact, we face a similar situation during mental imagery: the neural representations mediating perception and mental imagery are very similar, but there is no mechanism in the perceptual system to distinguish them. A *source monitoring function* is required to keep track of the origins of the perceptual neural representation. Therefore, we hypothesize that a higher order function monitors and distinguishes internally versus externally induced neural representations. Such a monitoring operation is functionally independent from the perceptual estimation process that internally reconstructs the perceptual representation. Under this hypothesis, auditory hallucinations are caused by incorrect operation of the monitoring function, resulting in incorrectly labeling the self-induced auditory representation during the intact internal perceptual estimation processes.

Computationally, the independence of the monitoring function versus internal simulation and estimation is demonstrated by the nuanced differences between *corollary discharge* and the *efference copy* [see the review by Crapse and Sommer (2008)]. The efference copy is a duplicate of the planned motor command and provides the dynamics of an action trajectory that can be used to estimate the perceptual consequences (von Holst and Mittelstaedt, 1950, 1973). Corollary discharge is a more general motor related mechanism that can be available at all levels of a motor process. The corollary discharge does not necessarily contain the same representational information as an efference copy; rather, it serves as a generic signal to inform sensory-perceptual systems of the potential occurrence of perceptual changes caused by one's own actions (Sperry, 1950). In the case of speech articulation, these two functions originate at the same stage of motor simulation, but their functional roles are still separate. The efference copy is used to estimate the detailed perceptual consequences, whereas the corollary discharge labels the internally and externally induced perceptual consequences.

Empirically, the finding that auditory hallucination patients can generate inner speech (e.g., Shergill et al., 2003) demonstrates the relatively intact motor-to-sensory transformation function. The neural responses in IFG and superior temporal gyrus/sulcus (STG/STS) were observed during auditory hallucinations, hinting at the derivation of auditory perceptual consequences from motor simulation during the positive symptom (e.g., McGuire et al., 1993; Shergill et al., 2003). Moreover, the left lateralization during covert speech versus right lateralization during auditory hallucinations offers tantalizing hints about the independence between self-monitoring and the sequential simulation-estimation (Sommer et al., 2008).

We summarize the hypothetical mechanistic account for auditory hallucinations (of this type) as follows: when patients prepare to articulate speech covertly or subvocally (either consciously or unconsciously), the internal motor simulation leads to perceptual estimation (intact efference copy). But the source monitoring process malfunctions (broken corollary discharge). Therefore, the internal prediction of a perceptual consequence, which has the same neural representation as an external perception, is erroneously interpreted as the result of external sources, resulting in an auditory hallucination.

STUTTERING: NOISY ESTIMATION AND CORRECTION PROCESSES

The comparison between internal estimation and external feedback provides information to fine-tune motor control. However, if the internal estimate from motor simulation malfunctions and generates imprecise perceptual predictions, an inaccurate or incorrect feedback control signal would be conveyed. Stuttering could be an example of such erroneous correction. We suggest, along the lines of similar theories (Max et al., 2004; Hickok et al., 2011), that one of the neural mechanisms causing stuttering is a deficit in the motor-to-sensory transformation. That is, the noisy perceptual estimation is mismatched to the external feedback. Such a discrepancy would signal an incorrect error message, and the feedback control system would interpret such an apparent error as the requirement to correct motor action. Hence, unnecessary attempts would be performed to modify the correct articulation, resulting in repetitive/prolonged sound or silent pauses/blocks.

The noise in the estimation process can come both from the somatosensory and auditory domains (since there is sequential estimation). Stutterers showed speed and latency deficits when required to sequentially update articulator movement (Caruso et al., 1988). Smaller magnitude compensation with longer latency adjustment to the perturbation on the jaws was also observed in stutterers (Caruso et al., 1987). In the auditory domain, smaller magnitude compensation to the perturbation of F1 formant in auditory feedback is observed (Cai et al., 2012). The inaccurate compensation to external perturbation in both somatosensory and auditory domains (with intact somatosensory and auditory processes) demonstrates that inaccurate prediction in both domains could be causal for stuttering.

Interestingly, dramatically altering auditory feedback (e.g., by delaying feedback onset or shifting frequency) can enhance speech fluency in people who stutter (Martin and Haroldson, 1979; Stuart et al., 1997, 2008). The improvement could be because the magnitude of error signals is scaled down when the distance between feedback and prediction is beyond some threshold, so that fewer correction attempts are made.

PHANTOM LIMBS: MISMATCH BETWEEN INTERNAL ESTIMATION AND EXTERNAL FEEDBACK

The mismatch between internal prediction and external feedback could also be caused by an acute change of conditions leading to the absence of feedback. One such example is the phantom limb phenomenon, where amputees feel control over a lost limb (phantom limb) accompanied with chronic and sometimes acute pain. We hypothesize that the apparent awareness and control of a lost limb occurs as follows: the missing somatosensory feedback is "replaced" by the results of internal estimation (cf. Frith et al., 2000; Fotopoulou et al., 2008). Such a hypothesis is similar to the mislabeling of the internal estimation as an external perception (due to the malfunction of source monitoring) in auditory hallucinations.

The causes of pain in phantom limbs are more intriguing. The most significant physical changes are loss of proprioception, or somatosensory afference, after lost limbs. Because motor control as well as motor simulation of the lost limb are still in some sense valid (e.g., Raffin et al., 2012), we hypothesize that a mismatch between the intact internal estimation and absent external somatosensory feedback can cause the pain associated with phantom limbs. In fact, consistent with our hypothesis, limb pain can be induced in normal participants by mismatching visual and proprioceptive feedback (McCabe et al., 2005) and spinal cord injured patients report that neuropathic pain increases while they imagine moving their ankles (Gustin et al., 2008).

This mismatch hypothesis may represent an intermediate step between cortical reorganization and pain induction. Lost limbs cause reorganization in both motor (Maihöfner et al., 2007) and somatosensory (Maihöfner et al., 2003) cortices, and pain reduction has been demonstrated to correlate with more granular organization in the same areas (MacIver et al., 2008). Motor imagery can lead to cortical reorganization that correlates with pain reduction in phantom limbs (Moseley, 2006). Seeing the movement of the opposite functioning arm in a mirror can reduce the pain associated with the phantom limb (Ramachandran et al., 1995). Such behavioral and psychological training can provide more precise topographic maps in both motor and somatosensory cortices and hence reduce the inaccurate motor firing caused by the "take over" effect (e.g., cortex of lip movement expand to the cortex mediated a lost hand), as well as erroneous somatosensory estimation. The internal estimation hypothesis offers a new perspective on pain induction. However, there is neither a clear pain center (Mazzola et al., 2012) nor a mechanistic pain induction account (Flor, 2002). Further research is needed to understand how the proposed mismatch hypothesis could underpin pain induction.

CONCLUSION

In this perspective, we argued that mental imagery is an internal predictive process. Using mental imagery of speech as an example, we demonstrated a variety of principles underlying how the mechanism of motor simulation and sequential perceptual estimation in mental imagery works. We conclude that the simulation-estimation mechanism provides a novel conceptual and practical perspective that allows for new types of research on predictive functions and sensory-motor integration, as well as stimulating some new insights into several neural disorders. Typically, mental imagery has been studied in cognitive psychology and cognitive neuroscience, while the concepts of internal forward models (and sensory-motor integration) are the focus of motor control research from an engineering perspective. Our atypical pairing of internal models as an additional source for mental imagery yields, in our view, some provocative new angles on mental imagery in both basic research and applied contexts.

ACKNOWLEDGMENTS

This study was supported by MURI ARO #54228-LS-MUR and NIH 2R01DC 05660.

REFERENCES

Alais, D., and Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Curr. Biol.* 14, 257–262.

Behroozmand, R., Liu, H., and Larson, C. R. (2011). Time-dependent neural processing of auditory feedback during voice pitch error detection. *J. Cogn. Neurosci.* 23, 1205–1217.

Bensafi, M., Porter, J., Pouliot, S., Mainland, J., Johnson, B., Zelano, C., et al. (2003). Olfactomotor activity during imagery mimics that during perception. *Nat. Neurosci.* 6, 1142–1144.

Blakemore, S. J., and Frith, C. (2003). Self-awareness and action. *Curr. Opin. Neurobiol.* 13, 219–224.

Blakemore, S. J., Smith, J., Steel, R., Johnstone, E., and Frith, C. (2000). The perception of self-produced sensory stimuli in patients with auditory hallucinations and passivity experiences: evidence for a breakdown in self-monitoring. *Psychol. Med.* 30, 1131–1139.

Blakemore, S. J., Wolpert, D. M., and Frith, C. D. (1998). Central cancellation of self-produced tickle sensation. *Nat. Neurosci.* 1, 635–640.

Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., et al. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. *Eur. J. Neurosci.* 13, 400–404.

Buxbaum, L. J., Kyle, K. M., and Menon, R. (2005). On beyond mirror neurons: internal representations subserving imitation and recognition of skilled object-related actions in humans. *Cogn. Brain Res.* 25, 226–239.

Cai, S., Beal, D. S., Ghosh, S. S., Tiede, M. K., Guenther, F. H., and Perkell, J. S. (2012). Weak responses to auditory feedback perturbation during articulation in persons who stutter: evidence for abnormal auditory-motor transformation. *PLoS ONE* 7:e41830. doi: 10.1371/journal.pone.0041830

Caruso, A. J., Abbs, J. H., and Gracco, V. L. (1988). Kinematic analysis of multiple movement coordination during speech in stutterers. *Brain* 111, 439.

Caruso, A. J., Gracco, V., and Abbs, J. H. (1987). "A speech motor control perspective on stuttering: preliminary observations," in *Speech Motor Dynamics in Stuttering,* eds H. F. M. Peters and W. Hulstijn (Wien, Austria: Springer-Verlag), 245–258.

Crapse, T. B., and Sommer, M. A. (2008). Corollary discharge across the animal kingdom. *Nat. Rev. Neurosci.* 9, 587–600.

Decety, J. (1996). The neurophysiological basis of motor imagery. *Behav. Brain Res.* 77, 45–52.

Decety, J., Grezes, J., Costes, N., Perani, D., Jeannerod, M., Procyk, E., et al. (1997). Brain activity during observation of actions. Influence of action content and subject's strategy. *Brain* 120, 1763.

Decety, J., Jeannerod, M., and Prablanc, C. (1989). The timing of mentally represented actions. *Behav. Brain Res.* 34, 35–42.

Decety, J., and Michel, F. (1989). Comparative analysis of actual and mental movement times in two graphic tasks. *Brain Cogn.* 11, 87–97.

Decety, J., Perani, D., Jeannerod, M., Bettinardi, V., Tadary, B., Woods, R., et al. (1994). Mapping motor representations with positron emission tomography. *Nature* 371, 600–602.

Dechent, P., Merboldt, K. D., and Frahm, J. (2004). Is the human primary motor cortex involved in motor imagery? *Cogn. Brain Res.* 19, 138–144.

Desmurget, M., Reilly, K. T., Richard, N., Szathmari, A., Mottolese, C., and Sirigu, A. (2009). Movement intention after parietal cortex stimulation in humans. *Science* 324, 811.

Desmurget, M., and Sirigu, A. (2009). A parietal-premotor network for movement intention and motor awareness. *Trends Cogn. Sci.* 13, 411–419.

di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., and Rizzolatti, G. (1992). Understanding motor events: a neurophysiological study. *Exp. Brain Res.* 91, 176–180.

Djordjevic, J., Zatorre, R., Petrides, M., Boyle, J., and Jones-Gotman, M. (2005). Functional neuroimaging of odor imagery. *Neuroimage* 24, 791–801.

Ehrsson, H. H., Geyer, S., and Naito, E. (2003). Imagery of voluntary movement of fingers, toes, and tongue activates corresponding body-part-specific motor representations. *J. Neurophysiol.* 90, 3304–3316.

Eliades, S. J., and Wang, X. (2003). Sensory-motor interaction in the primate auditory cortex during self-initiated vocalizations. *J. Neurophysiol.* 89, 2194.

Eliades, S. J., and Wang, X. (2005). Dynamics of auditory–vocal interaction in monkey auditory cortex. *Cereb. Cortex* 15, 1510.

Eliades, S. J., and Wang, X. (2008). Neural substrates of vocalization feedback monitoring in primate auditory cortex. *Nature* 453, 1102–1106.

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

Fetsch, C. R., Turner, A. H., DeAngelis, G. C., and Angelaki, D. E. (2009). Dynamic reweighting of visual and vestibular cues during self-motion perception. *J. Neurosci.* 29, 15601.

Fisher, J. C. (2006). Does simulation theory really involve simulation? *Philos. Psychol.* 19, 417–432.

Flor, H. (2002). Phantom-limb pain: characteristics, causes, and treatment. *Lancet Neurol.* 1, 182–189.

Ford, J. M., Gray, M., Faustman, W. O., Roach, B. J., and Mathalon, D. H. (2007). Dissecting corollary discharge dysfunction in schizophrenia. *Psychophysiology* 44, 522–529.

Fotopoulou, A., Tsakiris, M., Haggard, P., Vagopoulou, A., Rudd, A., and Kopelman, M. (2008). The role of motor intention in motor awareness: an experimental study on anosognosia for hemiplegia. *Brain* 131, 3432–3442.

Foxe, J. J., Wylie, G. R., Martinez, A., Schroeder, C. E., Javitt, D. C., Guilfoyle, D., et al. (2002). Auditory-somatosensory multisensory processing in auditory association cortex: an fMRI study. *J. Neurophysiol.* 88, 540.

Frith, C. D., Blakemore, S. J., and Wolpert, D. M. (2000). Abnormalities in the awareness and control of action. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 355, 1771–1788.

Fu, K. M. G., Johnston, T. A., Shah, A. S., Arnold, L., Smiley, J., Hackett, T. A., et al. (2003). Auditory cortical neurons respond to somatosensory stimulation. *J. Neurosci.* 23, 7510–7515.

Gallese, V., Fadiga, L., Fogassi, L., and Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain* 119, 593.

Gerardin, E., Sirigu, A., Lehericy, S., Poline, J.-B., Gaymard, B., Marsault, C., et al. (2000). Partially overlapping neural networks for real and imagined hand movements. *Cereb. Cortex* 10, 1093–1104.

Geva, S., Jones, P. S., Crinion, J. T., Price, C. J., Baron, J. C., and Warburton, E. A. (2011). The neural correlates of inner speech defined by voxel-based lesion–symptom mapping. *Brain* 134, 3071–3082.

Goldman, A. I. (1989). Interpretation psychologized. *Mind Lang.* 4, 161–185.

Grush, R. (2004). The emulation theory of representation: motor control, imagery, and perception. *Behav. Brain Sci.* 27, 377–396.

Guenther, F. H., Ghosh, S. S., and Tourville, J. A. (2006). Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain Lang.* 96, 280–301.

Gustin, S. M., Wrigley, P. J., Gandevia, S. C., Middleton, J. W., Henderson, L. A., and Siddall, P. J. (2008). Movement imagery increases pain in people with neuropathic pain following complete thoracic spinal cord injury. *Pain* 137, 237–244.

Hanakawa, T., Immisch, I., Toma, K., Dimyan, M. A., Van Gelderen, P., and Hallett, M. (2003). Functional properties of brain areas associated with motor execution and imagery. *J. Neurophysiol.* 89, 989–1002.

Heilman, K. M., Rothi, L. J., and Valenstein, E. (1982). Two forms of ideomotor apraxia. *Neurology* 32, 342–342.

Heinks-Maldonado, T. H., Mathalon, D. H., Houde, J. F., Gray, M., Faustman, W. O., and Ford, J. M. (2007). Relationship of imprecise corollary discharge in schizophrenia to auditory hallucinations. *Arch. Gen. Psychiatry* 64, 286.

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

Hickok, G. (2009). Eight problems for the mirror neuron theory of action understanding in monkeys and humans. *J. Cogn. Neurosci.* 21, 1229–1243.

Hickok, G. (2012). Computational neuroanatomy of speech production. *Nat. Rev. Neurosci.* 13, 135–145.

Hickok, G., Houde, J., and Rong, F. (2011). Sensorimotor integration in speech processing: computational basis and neural organization. *Neuron* 69, 407–422.

Houde, J. F., Nagarajan, S. S., Sekihara, K., and Merzenich, M. M. (2002). Modulation of the auditory cortex during speech: an MEG study. *J. Cogn. Neurosci.* 14, 1125–1138.

Iacoboni, M., Woods, R. P., Brass, M., Bekkering, H., Mazziotta, J. C., and Rizzolatti, G. (1999). Cortical mechanisms of human imitation. *Science* 286, 2526.

Ishai, A., Haxby, J. V., and Ungerleider, L. G. (2002). Visual imagery of famous faces: effects of memory and attention revealed by fMRI. *Neuroimage* 17, 1729–1741.

Ito, T., Tiede, M., and Ostry, D. J. (2009). Somatosensory function in speech perception. *Proc. Natl. Acad. Sci. U.S.A.* 106, 1245.

Jeannerod, M. (1994). The representing brain: neural correlates of motor intention and imagery. *Behav. Brain Sci.* 17, 187–202.

Jeannerod, M. (1995). Mental imagery in the motor context. *Neuropsychologia* 33, 1419–1432.

Jeannerod, M. (2001). Neural simulation of action: a unifying mechanism for motor cognition. *Neuroimage* 14, S103–S109.

Jeannerod, M., and Pacherie, E. (2004). Agency, simulation and self-identification. *Mind Lang.* 19, 113–146.

Kosslyn, S. M. (1994). *Image and Brain: The Resolution of the Imagery Debate*. Cambridge, MA: MIT Press.

Kosslyn, S. M. (2005). Mental images and the brain. *Cogn. Neuropsychol.* 22, 333–347.

Kosslyn, S. M., Alpert, N. M., Thompson, W. L., Chabris, C. F., Rauch, S. L., and Anderson, A. K. (1994). Identifying objects seen from different viewpoints A PET investigation. *Brain* 117, 1055.

Kosslyn, S. M., Digirolamo, G. J., Thompson, W. L., and Alpert, N. M. (1998). Mental rotation of objects versus hands: neural mechanisms revealed by positron emission tomography. *Psychophysiology* 35, 151–161.

Kosslyn, S. M., Ganis, G., and Thompson, W. L. (2001). Neural foundations of imagery. *Nat. Rev. Neurosci.* 2, 635–642.

Kosslyn, S. M., Pascual-Leone, A., Felician, O., Camposano, S., Keenan, J., Ganis, G., et al. (1999). The role of area 17 in visual imagery: convergent evidence from PET and rTMS. *Science* 284, 167.

Kosslyn, S. M., and Thompson, W. L. (2003). When is early visual cortex activated during visual mental imagery? *Psychol. Bull.* 129, 723–746.

Kosslyn, S. M., Thompson, W. L., Wraga, M., and Alpert, N. M. (2001). Imagining rotation by endogenous versus exogenous forces: distinct neural mechanisms. *Neuroreport* 12, 2519.

Kraemer, D. J. M., Macrae, C. N., Green, A. E., and Kelley, W. M. (2005). Musical imagery: sound of silence activates auditory cortex. *Nature* 434, 158.

Lametti, D. R., Nasir, S. M., and Ostry, D. J. (2012). Sensory preference in speech production revealed by simultaneous alteration of auditory and somatosensory feedback. *J. Neurosci.* 32, 9351–9358.

Lotto, A. J., Hickok, G. S., and Holt, L. L. (2009). Reflections on mirror neurons and speech perception. *Trends Cogn. Sci.* 13, 110–114.

Lotze, M., Montoya, P., Erb, M., Hulsmann, E., Flor, H., Klose, U., et al. (1999). Activation of cortical and cerebellar motor areas during executed and imagined hand movements: an fMRI study. *J. Cogn. Neurosci.* 11, 491–501.

MacIver, K., Lloyd, D., Kelly, S., Roberts, N., and Nurmikko, T. (2008). Phantom limb pain, cortical reorganization and the therapeutic effect of mental imagery. *Brain* 131, 2181.

Mahon, B. Z., and Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *J. Physiol. Paris* 102, 59–70.

Maihöfner, C., Baron, R., DeCol, R., Binder, A., Birklein, F., Deuschl, G., et al. (2007). The motor system shows adaptive changes in complex regional pain syndrome. *Brain* 130, 2671.

Maihöfner, C., Handwerker, H. O., Neundörfer, B., and Birklein, F. (2003). Patterns of cortical reorganization in complex regional pain syndrome. *Neurology* 61, 1707–1715.

Martin, R., and Haroldson, S. K. (1979). Effects of five experimental treatments on stuttering. *J. Speech Hear. Res.* 22, 132.

Max, L., Guenther, F. H., Gracco, V. L., Ghosh, S. S., and Wallace, M. E. (2004). Unstable or insufficiently activated internal models and feedback-biased motor control as sources of dysfluency: a theoretical model of stuttering. *Contemp. Issues Commun. Sci. Disord.* 31, 105–122.

Mazzola, L., Isnard, J., Peyron, R., and Mauguière, F. (2012). Stimulation of the human cortex and the experience of pain: Wilder Penfield's observations revisited. *Brain* 135, 631–640.

McCabe, C., Haigh, R., Halligan, P., and Blake, D. (2005). Simulating sensory–motor incongruence in healthy volunteers: implications for a cortical model of pain. *Rheumatology* 44, 509.

McGuire, P., Murray, R., and Shah, G. (1993). Increased blood flow in Broca's area during auditory hallucinations in schizophrenia. *Lancet* 342, 703–706.

Meister, I. G., Krings, T., Foltys, H., Boroojerdi, B., Müller, M., Töpper, R., et al. (2004). Playing piano in the mind: an fMRI study on music imagery and performance in pianists. *Cogn. Brain Res.* 19, 219–228.

Morgan, M. L., DeAngelis, G. C., and Angelaki, D. E. (2008). Multisensory integration in macaque visual cortex depends on cue reliability. *Neuron* 59, 662–673.

Moseley, G. L. (2006). Graded motor imagery for pathologic pain. *Neurology* 67, 2129–2134.

Moulton, S. T., and Kosslyn, S. M. (2009). Imagining predictions: mental imagery as mental emulation. *Philos. Trans. R. Soc. B Biol. Sci.* 364, 1273–1280.

Nikulin, V. V., Hohlefeld, F. U., Jacobs, A. M., and Curio, G. (2008). Quasi-movements: a novel motor-cognitive phenomenon. *Neuropsychologia* 46, 727–742.

Numminen, J., Salmelin, R., and Hari, R. (1999). Subject's own speech reduces reactivity of the human auditory cortex. *Neurosci. Lett.* 265, 119–122.

O'Craven, K. M., and Kanwisher, N. (2000). Mental imagery of faces and places activates corresponding stimulus-specific brain regions. *J. Cogn. Neurosci.* 12, 1013–1023.

Papeo, L., Vallesi, A., Isaja, A., and Rumiati, R. I. (2009). Effects of TMS on different stages of motor and non-motor verb processing in the primary motor cortex. *PLoS ONE* 4:e4508. doi: 10.1371/journal.pone.0004508

Poeppel, D., Idsardi, W. J., and van Wassenhove, V. (2008). Speech perception at the interface of neurobiology and linguistics. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 363, 1071.

Price, C. J., Crinion, J. T., and MacSweeney, M. (2011). A generative model of speech production in Broca's and Wernicke's areas. *Front. Psychology* 2:237. doi: 10.3389/fpsyg.2011.00237

Pulvermüller, F., and Fadiga, L. (2010). Active perception: sensorimotor circuits as a cortical basis for language. *Nat. Rev. Neurosci.* 11, 351–360.

Raffin, E., Mattout, J., Reilly, K. T., and Giraux, P. (2012). Disentangling motor execution from motor imagery with the phantom limb. *Brain* 135, 582–595.

Ramachandran, V. S., Rogers-Ramachandran, D., and Cobb, S. (1995). Touching the phantom limb. *Nature* 377, 489.

Rizzolatti, G., and Craighero, L. (2004). The mirror-neuron system. *Annu. Rev. Neurosci.* 27, 169–192.

Rizzolatti, G., Fadiga, L., Matelli, M., Bettinardi, V., Paulesu, E., Perani, D., et al. (1996). Localization of grasp representations in humans by PET: 1. Observation versus execution. *Exp. Brain Res.* 111, 246–252.

Rumiati, R. I., Papeo, L., and Corradi-Dell'Acqua, C. (2010). Higher-level motor processes. *Ann. N.Y. Acad. Sci.* 1191, 219–241.

Saygin, A. P., Wilson, S. M., Dronkers, N. F., and Bates, E. (2004). Action comprehension in aphasia: linguistic and non-linguistic deficits and their lesion correlates. *Neuropsychologia* 42, 1788–1804.

Scheerer, E. (1984). "Motor theories of cognitive structure: a historical review," in *Cognition and Motor Processes*, eds W. Prinz and A. Sanders (Berlin: Springer-Verlag), 77–98.

Schroeder, C. E., Lindsley, R. W., Specht, C., Marcovici, A., Smiley, J. F., and Javitt, D. C. (2001). Somatosensory input to auditory association cortex in the macaque monkey. *J. Neurophysiol.* 85, 1322.

Shepard, R. N., and Cooper, L. A. (1986). *Mental Images and Their Transformations*. Cambridge, MA: The MIT Press.

Shergill, S. S., Brammer, M. J., Fukuda, R., Williams, S. C. R., Murray, R. M., and McGuire, P. K. (2003). Engagement of brain areas implicated in processing inner speech in people with auditory hallucinations. *Br. J. Psychiatry* 182, 525–531.

Sirigu, A., Cohen, L., Duhamel, J. R., Pillon, B., Dubois, B., Agid, Y., et al. (1995). Congruent unilateral impairments for real and imagined hand movements. *Neuroreport* 6, 997–1001.

Sirigu, A., and Duhamel, J. (2001). Motor and visual imagery as two complementary but neurally dissociable mental processes. *J. Cogn. Neurosci.* 13, 910–919.

Sirigu, A., Duhamel, J., Cohen, L., Pillon, B., Dubois, B., and Agid, Y. (1996). The mental representation of hand movements after parietal cortex damage. *Science* 273, 1564–1568.

Sommer, I. E. C., Diederen, K. M. J., Blom, J. D., Willems, A., Kushan, L., Slotema, K., et al. (2008). Auditory verbal hallucinations predominantly activate the right inferior frontal area. *Brain* 131, 3169.

Sperry, R. (1950). Neural basis of the spontaneous optokinetic response produced by visual inversion. *J. Comp. Physiol. Psychol.* 43, 482.

Stuart, A., Frazier, C. L., Kalinowski, J., and Vos, P. W. (2008). The effect of frequency altered feedback on stuttering duration and type. *J. Speech Lang. Hear. Res.* 51, 889.

Stuart, A., Kalinowski, J., and Rastatter, M. P. (1997). Effect of monaural and binaural altered auditory feedback on stuttering frequency. *J. Acoust. Soc. Am.* 101, 3806.

Tian, X., and Huber, D. E. (2008). Measures of spatial similarity and response magnitude in MEG and scalp EEG. *Brain Topogr.* 20, 131–141.

Tian, X., and Poeppel, D. (2010). Mental imagery of speech and movement implicates the dynamics of internal forward models. *Front. Psychology* 1:166. doi: 10.3389/fpsyg.2010.00166

Tian, X., Poeppel, D., and Huber, D. E. (2011). TopoToolbox: using sensor topography to calculate psychologically meaningful measures from event-related EEG/MEG. *Comput. Intell. Neurosci.* 2011, 8.

Tkach, D., Reimer, J., and Hatsopoulos, N. G. (2007). Congruent activity during action and action observation in motor cortex. *J. Neurosci.* 27, 13241.

Tourville, J. A., Reilly, K. J., and Guenther, F. H. (2008). Neural mechanisms underlying auditory feedback control of speech. *Neuroimage* 39, 1429–1443.

Tsakiris, M., and Haggard, P. (2005). Experimenting with the acting self. *Cogn. Neuropsychol.* 22, 387–407.

van Wassenhove, V., Grant, K. W., and Poeppel, D. (2005). Visual speech speeds up the neural processing of auditory speech. *Proc. Natl. Acad. Sci. U.S.A.* 102, 1181.

Ventura, M., Nagarajan, S., and Houde, J. (2009). Speech target modulates speaking induced suppression in auditory cortex. *BMC Neurosci.* 10:58. doi: 10.1186/1471-2202-10-58

von Holst, E., and Mittelstaedt, H. (1950). Daz Reafferenzprinzip. Wechselwirkungen zwischen Zentralnerven-system und Peripherie. *Naturwissenschaften* 37, 467–476.

von Holst, E., and Mittelstaedt, H. (1973). *The Reafference Principle (R. Martin, Trans.). The Behavioral Physiology of Animals and Man: The Collected Papers of Erich von Holst.* (Coral Gables, FL: University of Miami Press), 139–173.

von Kriegstein, K., and Giraud, A. L. (2006). Implicit multisensory associations influence voice recognition. *PLoS Biol.* 4:e326. doi: 10.1371/journal.pbio.0040326

Wilson, S. M., Saygin, A. P., Sereno, M. I., and Iacoboni, M. (2004). Listening to speech activates motor areas involved in speech production. *Nat. Neurosci.* 7, 701–702.

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

Yoo, S. S., Freeman, D. K., McCarthy, J. J. 3rd., and Jolesz, F. A. (2003). Neural substrates of tactile imagery: a functional MRI study. *Neuroreport* 14, 581.

Zatorre, R. J., and Halpern, A. R. (2005). Mental concerts: musical imagery and auditory cortex. *Neuron* 47, 9–12.

Zatorre, R. J., Halpern, A. R., Perry, D. W., Meyer, E., and Evans, A. C. (1996). Hearing in the mind's ear: a PET investigation of musical imagery and perception. *J. Cogn. Neurosci.* 8, 29–46.

Zhang, M., Weisser, V. D., Stilla, R., Prather, S., and Sathian, K. (2004). Multisensory cortical processing of object shape and its relation to mental imagery. *Cogn. Affect. Behav. Neurosci.* 4, 251–259.

Zheng, Z. Z., Munhall, K. G., and Johnsrude, I. S. (2010). Functional overlap between regions involved in speech perception and in monitoring one's own voice during speech production. *J. Cogn. Neurosci.* 22, 1770–1781.

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: 20 April 2012; accepted: 06 November 2012; published online: 28 November 2012.

Citation: Tian X and Poeppel D (2012) Mental imagery of speech: linking motor and perceptual systems through internal simulation and estimation. Front. Hum. Neurosci. ***6****:314. doi: 10.3389/fnhum.2012.00314*

frontiers in
HUMAN NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 01 November 2012
doi: 10.3389/fnhum.2012.00299

Effect of biomechanical constraints in the hand laterality judgment task: where does it come from?

Gilles Vannuscorps[1,2]*, **Agnesa Pillon**[1,2] **and Michael Andres**[2,3]

[1] *Institute of Psychological Sciences and Institute of Neuroscience, Université catholique de Louvain, Louvain-la-Neuve, Belgium*
[2] *Fonds National de la Recherche Scientifique, Brussels, Belgium*
[3] *Faculty of Psychology, Department of Experimental Psychology, Ghent University, Ghent, Belgium*

Edited by:
Joel Pearson, The University of New South Wales, Australia

Reviewed by:
Bettina E. Bläsing, Bielefeld University, Germany
Joel Pearson, The University of New South Wales, Australia

****Correspondence:***
Gilles Vannuscorps, Institut de Recherche en Sciences Psychologiques, Université catholique de Louvain, Place du Cardinal Mercier 10, Louvain-la-Neuve 1348, Belgium.
e-mail: gilles.vannuscorps@uclouvain.be

Several studies have reported that, when subjects have to judge the laterality of rotated hand drawings, their judgment is automatically influenced by the biomechanical constraints of the upper limbs. The prominent account for this effect is that, in order to perform the task, subjects mentally rotate their upper limbs toward the position of the displayed stimulus in a way that is consistent with the biomechanical constraints underlying the actual movement. However, the effect of such biomechanical constraints was also found in the responses of motor-impaired individuals performing the hand laterality judgment (HLJ) task, which seems at odds with the "motor imagery" account for this effect. In this study, we further explored the source of the biomechanical constraint effect by assessing the ability of an individual (DC) with a congenital absence of upper limbs to judge the laterality of rotated hand or foot drawings. We found that DC was as accurate and fast as control participants in judging the laterality of both hand and foot drawings, without any disadvantage for hands when compared to feet. Furthermore, DC's response latencies (RLs) for hand drawings were influenced by the biomechanical constraints of hand movements in the same way as control participants' RLs. These results suggest that the effect of biomechanical constraints in the HLJ task is not strictly dependent on "motor imagery" and can arise from the visual processing of body parts being sensitive to such constraints.

Keywords: motor imagery, motor simulation, biomechanical knowledge, aplasia, body part perception

INTRODUCTION

In the past 20 years, an increased interest has been devoted to the study of motor imagery. Most research has been achieved by the means of the hand laterality judgment (HLJ) task (Cooper and Shepard, 1975; Sekiyama, 1982; Parsons, 1987, 1994). In this task, participants are asked to decide whether drawings of different hand postures rotated with different angles from the upright view depict a left or a right limb (see **Figure 1**). If a hand drawing were processed like any other visual object, response latencies (RLs) should increase as a function of the angular disparity between the hand drawing and its upright view, as evidenced for other 2D or 3D stimuli (Cooper and Shepard, 1975). Instead, it has been shown that the impact of the angular disparity on the speed of HLJs was strongly modulated by the biomechanical limits that constrain the movement of the hand toward the displayed position (Sekiyama, 1982; Parsons, 1987). This observation was taken as evidence that, to solve the task, participants internally simulate a movement of their own hand and, moreover, that mental imagery of human movements—i.e., motor imagery—relies on the same representations and processes as those involved in action planning and/or control (Parsons, 1987, 1994; Jeannerod and Frak, 1999; Kosslyn et al., 2001; Nico et al., 2004; Wraga et al., 2005; de Lange et al., 2006, 2008; Fiorio et al., 2006; Helmich et al., 2007; Munzert et al., 2009). Furthermore, neuroimaging studies showed that HLJs induced increased activity in a parieto-frontal network known for its contribution to the planning and execution of hand movements (Kosslyn et al., 1998; Parsons et al., 1998; de Lange et al., 2006). On the basis of this behavioral and neural evidence, the HLJ task was considered as a privileged tool to read out the unconscious and normally covert process of motor planning (Jeannerod, 1994; Fiorio et al., 2006; de Lange et al., 2008; Munzert et al., 2009). In this paper, we report evidence that calls for a re-examination of the prominent "motor imagery" account of the effect of biomechanical constraints in the HLJ task.

The effect of biomechanical constraints on participants' judgments in the HLJ task is characterized by three features unveiled by the analysis of the RL pattern (Parsons, 1987; Funk and Brugger, 2008; Pelgrims et al., 2009). First, the RLs for the hand drawings depicted in the various rotation angles (from 0° to 360° in a clockwise direction) from the upright display (0°) are modulated by both the specific posture of the depicted hand (e.g., view of the hand side from the thumb vs. view of the palm from the wrist) and its laterality (right vs. left hand). This three-way interaction reflects the different biomechanical constraints that limit the amplitude of the rotation angle of each hand in a given posture. Thus, for instance, the recognition of a side view takes more time for right than left hands if the drawing is rotated at 150° clockwise given that such an angle corresponds to the outcome of a possible left hand movement but it is almost impossible to

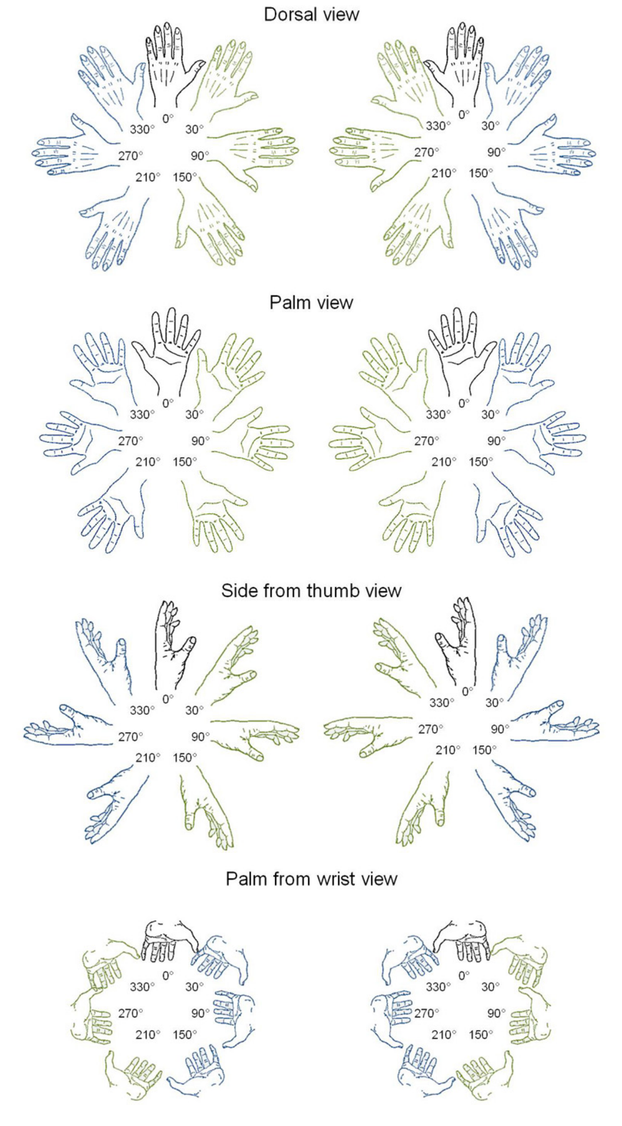


FIGURE 1 | Line drawings of left and right hands in four postures and seven rotation angles. In blue: lateral orientations; in green: medial orientations.

achieve with the right hand in this position. Likewise, the palm of a right hand viewed from the wrist, with the thumb pointing down toward the right side of the screen (i.e., between 210° and 270° clockwise) is recognized more slowly than a left hand displayed in the same orientation, in line with the fact that it is much more difficult to reach this orientation with the right than with the left hand in this posture. The second feature indexing the effect of biomechanical constraints on participants' responses is the overall chronometric advantage for judging the laterality of hands oriented in medial positions (stimuli rotated toward the mid-sagittal plane) when compared to lateral positions (stimuli rotated away of the mid-sagittal plane). This effect, called the "Medial Over Lateral Advantage" (MOLA) effect, reflects the impact of the biomechanical constraints of hand movements that make it easier to move one's hand toward medial than lateral directions. The third feature is the significant correlation between the RL to a given stimulus and its degree of awkwardness that is, how difficult participants rated it to actually place their own hand in the displayed position (Parsons, 1987).

There is, however, evidence showing that the effect of biomechanical constraints in the HLJ task can be observed even in the condition of impaired motor planning or execution processes. After transient disruption of motor-related areas with transcranial magnetic stimulation (TMS), the performance in the HLJ task was either normal (Sauner et al., 2006) or characterized by a small RL increase that nevertheless did not hamper the effect of biomechanical constraints (Ganis et al., 2000; Pelgrims et al., 2010). Furthermore, studies of patients suffering from motor disorders that prevent normal execution of hand movements such as congenital hemiparesis (Steenbergen et al., 2007), Parkinson's disease (Helmich et al., 2007), dystonia (Fiorio et al., 2006), conversion paralysis (de Lange et al., 2008), amputation of an upper limb (Nico et al., 2004), or chronic arm pain (Schwoebel et al., 2001) showed that these conditions delayed HLJs but, again, without affecting the effect of biomechanical constraints. Finally, a study using the HLJ task with two individuals suffering from bilateral upper limb aplasia reported an effect of biomechanical constraints on RLs in at least one individual (Funk and Brugger, 2008).

Evidence for an effect of biomechanical constraints in the HLJ task despite impaired motor planning or execution raises the possibility that the effect does not arise from motor but instead from visual processes that ensure the perception of human body parts. Several authors indeed suggested that, when representing the human body in whole or in parts, the visual system encodes information about the biomechanical constraints of body part movements, information that in turn constrains whole body or body part perception (Marr and Vaina, 1982; Kourtzi and Shiffrar, 1999). Evidence advanced in support of this view includes findings from the apparent motion paradigm. When an object is displayed sequentially in two different positions, it elicits in participants the perception of an apparent motion along the shortest pathway (Kolers and Pomerantz, 1971). However, the apparent motion induced by two hand postures presented sequentially can follow a longer pathway when the shortest one is not biomechanically possible (Shiffrar and Freyd, 1990, 1993). According to some authors (Shiffrar and Freyd, 1993; Chatterjee et al., 1996), this effect indicates that implicit perceptual knowledge of how the body moves impacts how body parts are perceived (but see Stevens et al., 2000, for an alternative account in terms of motor simulation).

However, before considering the perceptual hypothesis as an alternative account for the effect of biomechanical constraints in the HLJ task, a number of ambiguities that are present in the aforementioned studies of motor-impaired or aplasic individuals must be addressed. First, one cannot rule out that the motor-impaired patients in whom the effect of biomechanical constraints was found actually suffered from a deficit affecting processes that take place *after* the stage where biomechanical constraints influence motor planning and/or execution. Indeed, motor execution was not totally abolished in these patients. At first sight, the effect found in a bilateral aplasic individual (Funk and Brugger, 2008) should not present this ambiguity since, in such condition, none of the processes involved in the planning and execution of hand movements is functional. Nevertheless, this individual (AZ) reported vivid phantom sensations of her missing body parts that include phantom movements corresponding to those of normal upper limbs (Funk et al., 2005). Thus, one cannot rule out the existence of limb representations in AZ, which could explain her ability to carry out motor imagery for congenitally absent limbs (Brugger et al., 2000; Funk and Brugger, 2008). Furthermore, and this is also true for the studies with motor-impaired patients, the issue of whether mere visual familiarity with the various hand positions could have contributed to the observed pattern of RLs in the HLJ task was not addressed. What looks like an effect of awkwardness might in fact be an artifact of the differential occurrences of the various hand positions—awkward hand positions are also likely to be less often seen than easy ones. Within this visual familiarity account, the RLs in the HLJ task would mainly depend on how often a given hand position has been seen in everyday settings, with the more frequent hand positions being recognized faster as a right or a left hand than the less frequent ones.

In this study, we sought further evidence for the presence of the effect of biomechanical constraints in the HLJ task in the context of a motor disability to plan and execute hand movements. We presented the HLJ task to DC, a man born without upper limbs, as well as to 7 normally limbed control participants, and analyzed their pattern of RLs vis-a-vis the main features indexing the effect of biomechanical constraints. Because of the congenital disability of DC, who has also never experienced any phantom limb sensation, the present study overcomes the difficulties raised by previous studies with motor-impaired patients or aplasic individuals. In the case of DC, no hand motor planning or execution ability of any kind could be invoked to explain, if any, an effect of biomechanical constraints in the HLJ task. Furthermore, we examined the effect of biomechanical constraints on the RLs of DC and control participants by taking into account the potential effect of the rated visual familiarity of hand positions.

MATERIALS AND METHODS

PARTICIPANTS

DC is a 51 year-old man with a Master's Degree in Psychology. He presents a congenital bilateral upper limb aplasia (right side: two fingers attached to a foreshortened humerus; left side: completely aplasic) due to *in utero* thalidomide exposure. He had no experience of prosthesis or phantom limb sensations. His performance was compared with that of 7 right-handed, normally limbed control participants matched in gender, age (mean age = 53.5), and educational level. All participants had a normal or corrected to normal vision and no history of psychiatric or neurological disorder. The study was approved by the biomedical ethic committee of the *Cliniques universitaires Saint-Luc* (Brussels) and all participants gave written informed consent prior to the study.

TASKS AND STIMULI

Participants were presented with drawings of a hand outlined in black on a white background and asked to decide as fast as possible whether the drawing corresponded to a right or a left hand.

Stimuli were left or right hands, presented according to four different postures (dorsal view, palm view, side from thumb view, and palm from wrist view) and at 7 different rotation angles (upright 0° and 30°, 90°, 150°, 210°, 270°, and 330° in a clockwise direction, from Parsons, 1987; see **Figure 1**). The total number of different stimuli was thus of 56 hands (2 laterality × 4 postures × 7 angles). Ratings of motor awkwardness for each hand drawing were extracted from a previous study (Parsons, 1987), in which judges were asked to position their own hand at the orientation of each stimulus and, afterwards, to estimate the awkwardness of the reached position on a 5-point scale (1 = easy to place the appropriate limb into the orientation of the stimulus and 5 = difficult to place the appropriate limb into the orientation of the stimulus). Ratings of visual familiarity were collected in 25 students of the *Université catholique de Louvain* (8 males) who were asked to rate how often they saw a hand in each posture and angle of rotation in everyday life (1 = very unfamiliar and 5 = very familiar). DC's personal visual familiarity was collected separately following the same procedure; it was strongly correlated to the students' ratings [$r_{(56)} = 0.53$; $p < 0.001$]. As a control, we also tested foot laterality judgments using stimuli that were found to evoke implicit activation of the biomechanical constraints of foot movements in seminal studies of mental imagery (Parsons, 1987). Fifty-six drawings of left and right feet were presented at the same angle of rotation as hand drawings, according to four similar postures (sole view, top view, view from inside, and sole from heel view).

Participants seated in front of a computer screen located at a distance of about 60 cm; their feet were lying at rest on the ground and the hands of control participants were placed palms down on their knees without visual feedback.

During the experiment, participants performed 5 blocks of 56 trials with hand drawings and then 5 blocks of 56 trials with foot drawings. In each block, all postures and rotation angles were mixed in a different pseudo-randomized order. The first blocks of hand and foot laterality judgments included a familiarization with the four postures and 10 practice trials. Within each block, each trial started with the presentation of a central cross for 200 ms followed by a hand or foot drawing displayed until a response was recorded. Trials were separated by a blank screen of random duration between 500 ms to 1000 ms.

The experiment was controlled with the E-Prime software (Psychological Software, 2002, Pittsburgh, PA). Stimuli were presented on a 15.4 inch laptop screen set at 1024 × 768 pixels and subtended 5° of visual angle. During the testing, participants were asked to produce a verbal response ("right" or "left"). The RLs corresponded to the post-stimulus onset latency of the subject's vocalization, whose amplitude was electrically compared to a trip level voltage using a voice key controlled by E-prime. Malfunctioning of the voice key and response accuracy were monitored on-line by the experimenter.

RESULTS

Voice key failures (0% and 0.6% of the data in DC and controls, respectively), trials with RLs deviating more than 2 standard deviations from the mean RL within each participant (5.14% and 4.43% of the data in DC and controls, respectively), and trials with errors were discarded from RL analyses.

GENERAL ANALYSIS OF DC'S AND CONTROLS' PERFORMANCE

First of all, we performed Crawford and Howell's (1998) modified t-tests to test whether DC's performance (accuracy and speed) in hand and foot laterality judgments was impaired in comparison to the control group's performance. In hand judgments, no significant difference was observed between the performance of DC (correct responses: 97%; mean RL: 1237 ms) and control participants [mean % correct responses ± SD: 91% ± 4%, modified $t_{(6)} = 1.56$, $p > 0.1$; mean RL ± SD: 1277 ms ± 345 ms, modified $t_{(6)} = -0.11$, $p > 0.9$]. The analysis of foot judgments showed a similar pattern, with a non-significant difference in accuracy or RLs between DC (correct responses: 97%; mean RL: 1512 ms) and controls [mean % of correct responses ± SD: 90% ± 6%, modified $t_{(6)} = 1.13$, $p > 0.3$; mean RL ± SD: 1446 ms ± 386 ms, modified $t_{(6)} = 16$, $p > 0.8$]. Second, Crawford and Garthwaite's (2005) Revised Standardized Difference Test (RSDT) was applied to test whether the difference between hand and foot laterality judgments in DC deviated from the difference observed between hand and foot judgments in the control group. The results showed that DC's difference in performance between hand and foot drawings was not significantly different from that found in the control participants, either in accuracy [RSDT: $t_{(6)} = 0.31$, $p > 0.7$] or speed [RSDT: $t_{(6)} = 0.4$, $p > 0.7$]. Third, independent samples t-tests were performed in order to compare the RLs across posture, laterality, and angle of rotation for foot and hand judgments in DC and controls, respectively. The results indicated that both DC [$t_{(67.76)} = -2.99$, $p < 0.01$] and control participants [$t_{(84.1)} = -4.29$, $p < 0.01$] were significantly faster for hand than for foot judgments.

EFFECT OF BIOMECHANICAL CONSTRAINTS ON DC's AND CONTROLS' RLs

Having shown that the performance of DC in HLJs was within the normal range, we looked for the presence of the three behavioral features that were classically reported as evidence for an effect of biomechanical constraints in the HLJ task.

First, we investigated the presence of a three-way interaction between LATERALITY, ANGLE, and POSTURE in the chronometric data gathered for all participants. To do so, the RLs of control participants were entered in a repeated measure analysis of variance (ANOVA) with SUBJECT as the random factor and LATERALITY (left vs. right), POSTURE (dorsal view vs. palm view vs. palm from the wrist view vs. side from the thumb view), and ANGLE with respect to the upright view (0°–330°, in a clockwise direction) as within-subject factors. The data of control participants required a log transformation to satisfy the ANOVA's homoscedasticity and normality assumptions. In order to explore the effects of biomechanical constraints in DC's judgments, we performed an ANOVA with ITEM as the random factor and LATERALITY (left vs. right), POSTURE (dorsal view vs. palm view vs. palm from the wrist view vs. side from the thumb view), and ANGLE with respect to the upright view (0°–330°, in a clockwise direction) as between-item factors. For the analysis of DC's performance, the data were inverse transformed to fulfill the criteria of homoscedasticity and normality. The results replicated the significant three-way interaction between LATERALITY, ANGLE, and POSTURE, not only in control participants [$F_{(18,\ 108)} = 3.77$, $p < 0.001$] but also in DC [$F_{(18,\ 202)} = 2.69$, $p < 0.001$]. This three-way interaction showed that the effect of the angle of rotation on RLs was not symmetric for left and right hands, with the angle associated to the maximal increase varying as a function of hand posture.

In order to further exemplify the effect of biomechanical constraints, we decomposed the three-way interaction as a function of hand posture. In control participants, the log transformed data were analyzed separately for each hand posture using repeated measure ANOVAs with SUBJECT as a random factor and LATERALITY and ANGLE as within-subject factors. The inverse transformed data of DC were entered in similar ANOVAs with ITEM as a random factor and LATERALITY and ANGLE as between-item factors. In control participants, a significant LATERALITY by ANGLE interaction was found for the palm from the wrist view [$F_{(6,\ 36)} = 4.32$, $p < 0.01$] and for the palm view [$F_{(6,\ 36)} = 7.68$, $p < 0.001$]. A near significant interaction effect was observed for the side from the thumb view [$F_{(6,\ 36)} = 2.15$, $p = 0.07$] but, for the dorsal view, no significant interaction was found [$F_{(6,\ 36)} = 1.38$, $p > 0.2$]. The performance of DC also revealed a significant LATERALITY by ANGLE interaction for the palm from wrist view [$F_{(6,\ 51)} = 3.98$, $p < 0.01$] as well as for the side from the thumb view [$F_{(6,\ 56)} = 2.5$, $p < 0.05$], but not for the palm [$F_{(6,\ 41)} = 1.49$, $p > 0.2$] or dorsal view [$F_{(6,\ 54)} < 1$]. In **Figure 2**, we represented the pattern of RLs of control participants (upper panel) and DC (lower panel) for the two postures that showed a significant LATERALITY by ANGLE interaction in DC after decomposition of the three-way interaction. The Figure clearly illustrates that the angle of rotation had a different effect on RLs depending on hand laterality. In both controls (**Figure 2A**) and DC (**Figure 2C**), the side view of a left hand led to maximal RLs at angle 210°, whereas the slowest responses for the side view of a right hand were observed at angle 150°. In contrast, the palm from the wrist views showing the longest RLs ranged from angles 90° to 150° for the left hand and from angles 210° to 270° for the right hand (**Figures 2B,D**). The comparison of **Figures 2B** and **2D** showed that maximal RLs for the palm from the wrist views were not observed exactly at the same angles in controls and DC, with a RL curve skewed to angle 150° for DC. Except for this slight difference, the angles associated to maximal RLs in DC and controls reflect the most difficult positions to reach while adopting the displayed posture with one's left or right hand (see the awkwardness estimates provided by Parsons, 1987).

DC did not show the expected pattern of asymmetric RL curves for the left and right hands viewed from the dorsal or from the palm view. However, the absence of LATERALITY by ANGLE interaction for these two postures was already pointed out in previous experiments with normally limbed participants (Parsons, 1987; see for discussion ter Horst et al., 2010). Likewise, in our control participants, the effect of angular disparity was not modulated by hand laterality for dorsal views and a look at

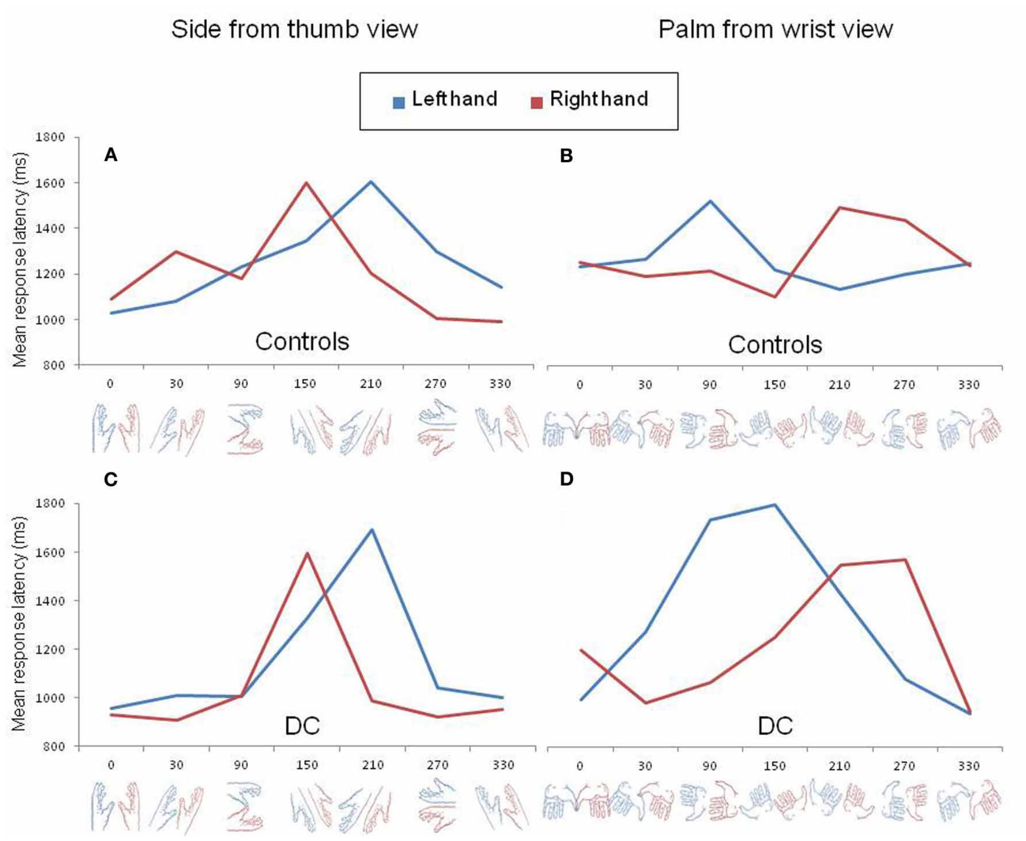


FIGURE 2 | Mean response latency for left (blue) and right (red) hand drawings rotated clockwise by steps of 60° viewed from the side (A,C) and from the wrist (B,D). The pattern of response latency in control participants and DC are shown in the upper (**A,B**) and lower (**C,D**) panels, respectively.

the individual data for the palm views revealed that the expected interaction was observed in only 3 out of 7 participants.

Second, in order to test for the presence of the MOLA effect, we recoded the trials according to the "medial" or "lateral" orientation of the displayed hand posture with respect to the body mid-sagittal plane (cf. **Figure 1**: medial orientations are displayed in green and lateral orientations in blue). In that way, the analyses were performed on the mean RLs calculated for each 24 medial and each 24 lateral hand displays (2 hands × 3 angles × 4 postures). **Figure 3** shows the mean RLs associated to medial and lateral orientations in DC and every control participant. Independent samples *t*-tests revealed that both control participants [medial: 1191 ms ± 134 ms; lateral: 1345 ms ± 139 ms; $t_{(46)} = 3.89, p < 0.001$] and DC [medial: 1205 ms ± 274 ms; lateral 1378 ms ± 335; $t_{(46)} = 1.95$, $p = 0.057$] responded faster to hand drawings in medial than in lateral orientation, which mimicked the effect of biomechanical constraints on actual hand movements. Furthermore, RSDT indicated that the MOLA effect did not significantly differ in size between DC and controls [RSDT: $t_{(6)} = 0.13, p > 0.9$].

Third, we calculated the correlation between the mean RL obtained for each of the 56 items and the estimates of motor awkwardness collected independently for the same items (Parsons, 1987). The correlation plots are displayed in **Figures 4A,B**. The analyses showed that the RLs of both controls [$r_{(56)} = 0.55$, $p < 0.001$] and DC [$r_{(56)} = 0.52$, $p < 0.001$] were significantly correlated to the motor awkwardness of the hand drawing. A final set of analyses was performed in order to partial out the

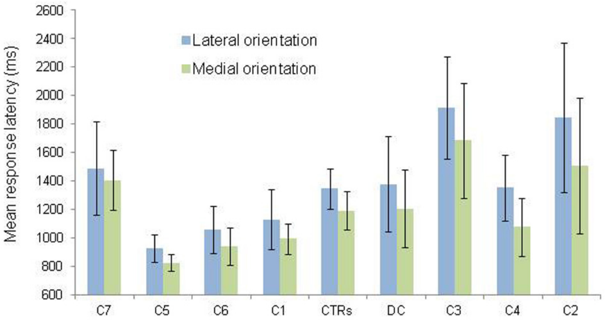


FIGURE 3 | Mean response latency and standard deviation (vertical bars) are shown as a function of hand orientation (lateral vs. medial). The data for the lateral and medial orientation were obtained by pooling the data from all postures and rotation angles corresponding to the lateral (blue) and medial (green) positions depicted in **Figure 1**. Individual data (C = Control; CTRs = mean of controls) are aligned in ascending order on the X axis as a function of the size of the medial over lateral advantage.

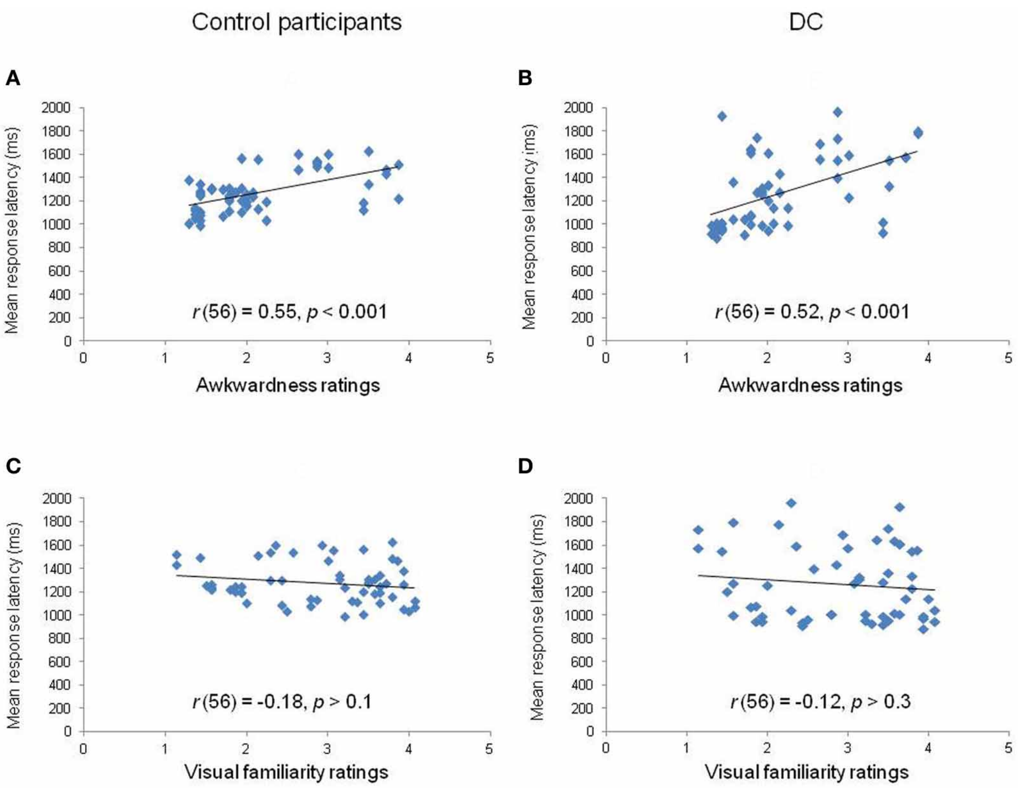


FIGURE 4 | Correlation plots between the mean response latency of controls (left panels) and DC (right panels) for each hand stimulus in relation to rated awkwardness (A,B) and visual familiarity (C,D).

respective influence of motor awkwardness and visual familiarity on RLs (see **Figures 4C,D**). We conducted stepwise regression and partial correlation analyses between RLs and, respectively, ratings of motor awkwardness and visual familiarity. The regression analysis performed in control participants [$F_{(1,\ 54)} = 23.94$, $p < 0.001$, $R^2 = 0.31$] showed that the motor awkwardness associated to a given hand drawing was the best predictor of the observed RLs ($\beta = 0.55$, $t = 4.89$, $p < 0.001$). The contribution of visual familiarity was not significant ($\beta = -0.04$, $t = -0.35$, $p > 0.7$). Likewise, the regression of DC's RLs [$F_{(1,\ 54)} = 19.87$, $p < 0.001$, $R^2 = 0.27$] revealed a significant effect of motor awkwardness ($\beta = 0.52$, $t = 4.46$, $p < 0.001$) in the context of a non-significant contribution of visual familiarity ($\beta = 0.01$, $t = 0.11$, $p > 0.9$). Similar results were observed when DC's own ratings of visual familiarity were taken into account in the regression equation [$F_{(1,\ 54)} = 19.87$, $p < 0.001$, $R^2 = 0.27$]. Furthermore, partial correlations showed that the correlation between RLs and awkwardness estimates remained significant in controls [$r_{(53)} = 0.54$, $p < 0.001$] and DC [$r_{(53)} = 0.51$, $p < 0.001$] after controlling for the part of variance explained by the effect of visual familiarity.

SUMMARY OF RESULTS

The goal of this study was to test whether the effect of biomechanical constraints in the HLJ task is strictly dependent on hand motor planning and execution abilities. To do so, we compared the performance of DC, a bilateral aplasic individual, with the performance of control participants in the HLJ task. First, we looked at the overall performance and found that DC was as fast and accurate as the control participants. Second, we found that (1) DC's RLs were influenced by the laterality, the angle, and the posture of hand stimuli, in a way that mirrors the biomechanical limits imposed by each posture on left and right hand movements; (2) DC's RLs showed a chronometric advantage of medial over lateral hand orientations, reflecting the difference of movement amplitude allowed by these two orientations; and (3) DC's RLs were strongly predicted by the motor awkwardness of the stimuli but not by their visual familiarity. To sum up, DC showed the three behavioral features that were classically reported as evidence for an effect of biomechanical constraints in the HLJ task. These effects were qualitatively and quantitatively comparable to the effects observed in control participants and they cannot be explained by the differential occurrences of the various hand positions in everyday life.

DISCUSSION

The effect of biomechanical constraints in the HLJ task is commonly assumed to reflect a process of motor simulation anchored in the same processes and representations as those involved in actual action planning and execution (Parsons, 1987, 1994; Jeannerod and Frak, 1999; Kosslyn et al., 2001; Nico et al., 2004;

Wraga et al., 2005; de Lange et al., 2006, 2008; Fiorio et al., 2006; Helmich et al., 2007; Munzert et al., 2009). In this paper, we found that a person born without upper limbs was as accurate and fast as a group of control participants in performing the HLJ task and that his RLs in this task were significantly influenced by the biomechanical constraints of upper limb movements, just like the RLs of normally limbed participants. These findings show that the effect of biomechanical constraints in the HLJ task is not strictly dependent on representations and processes involved in the planning and execution of hand movements. Given his total lack of motor experience with upper limb movements, and also of phantom limb experience, DC is not endowed with such motor representations and processes and could therefore not rely on them to perform the task.

Two kinds of accounts for the effect of biomechanical constraints in DC can be dismissed. First, this effect is not an artifact of the visual familiarity of the various hand positions. We found that DC's RLs were better predicted by the degree of motor awkwardness associated to each hand drawing than by their visual familiarity. Furthermore, the correlation between his RLs and motor awkwardness estimates remained significant even after removing the influence of visual familiarity on the data, which indicated that the influence of motor awkwardness and visual familiarity does not fully overlap in HLJs. Second, our data allow us to rule out that the effect is due to DC performing the HLJ task by mentally rotating the representation of his feet. Seminal studies in fact showed that the motor awkwardness estimates associated to certain foot positions correlate with those gathered for the homologue hand positions (see Tables 3 and 4 in Parsons, 1987). However, DC was 275 ms faster for hand than foot judgments. This advantage of hand over foot responses even reached an average of 583 ms for the side views that showed the typical effect of biomechanical constraints[1].

Our findings thus provide strong evidence for the presence of an effect of biomechanical constraints in the HLJ task in a condition that totally prevents the planning and execution of hand movements. Uncovering the source of this effect in such condition was beyond the scope of this study and we have no direct evidence that speaks to this issue. Nevertheless, the finding that knowledge of biomechanical constraints was implicitly and automatically recruited in DC's HLJs is consistent with the view that such knowledge is an intrinsic component of body part visual perception processes. These perceptual processes would provide us with a representation of the human body that takes into account information about the range of movement allowed by the different body parts (Marr and Vaina, 1982; Shiffrar and Freyd, 1990, 1993; Kourtzi and Shiffrar, 1999). Such information might have an adaptive value for humans because it facilitates the anticipation of the outcome of movements performed by others (Kourtzi and Shiffrar, 1999). The type of body representation we propose to explain the results of DC should not be confused with the "body schema" because this representation refers specifically to one's own body (Corradi-Dell'Acqua and Tessari, 2010; de Vignemont, 2010). It is also different from the "body structural description" (i.e., a visuospatial representation of body parts) and the "body image" (i.e., a conceptual representation of the body) because none of these representations include knowledge of the biomechanical constraints of the body (Sirigu et al., 1991; Schwoebel and Coslett, 2005).

The role of visual processes in the effect of biomechanical constraints in the perception of body parts was already emphasized by Brugger and colleagues in order to explain the influence of such constraints on HLJs (Brugger et al., 2000; Funk and Brugger, 2008) and apparent motion perception (Funk et al., 2005) in an bilateral aplasic individual, AZ, who experienced phantom limb movement sensations. It should be noted, however, that this proposal deviates from the idea that biomechanical knowledge is an intrinsic component of visual perception. In Brugger and colleagues' proposal, visual experience is assumed to activate pre-existing limb representations common to both action observation and execution, thereby allowing AZ to engage in a process of motor imagery in the HLJ task. Our finding that biomechanical constraints also affect the performance of a bilateral aplasic individual *without* phantom limb sensations in the HLJ task makes it unnecessary to assume that the role of visual experience is mediated by processes involved in action planning or execution.

In their study, Funk and Brugger (2008) also presented the HLJ task to a bilateral aplasic individual, CL, who did *not* experience phantom limb sensations, like DC, but, contrary to the results we reported here, they found no evidence for an effect of biomechanical constraints in CL's response pattern. These discrepant results in the HLJ task deserve some methodological considerations. While performing the task, CL was influenced by the rotation angle but not by the biomechanical constraints, as evidenced by an absence of interaction between hand laterality and rotation angle. However, CL was tested with a short version of the HLJ task including only two postures (palm vs. back) and four rotation angles (from 0° to 270° by steps of 90°). Previous research in healthy participants showed that minor modifications in the stimulus set, such as the reduction of the number of rotation axes, can suppress the effect of biomechanical constraints (ter Horst et al., 2010). Furthermore, laterality judgment for hands viewed from the back and the palm are less regularly associated with an effect of the biomechanical constraints, even in normally limbed participants (e.g., Parsons, 1987; this study). It is therefore unclear whether the effect of biomechanical constraints in CL's performance was not observed because he had no motor experience or because the stimulus set allowed him to base his responses on strategic processes that do not make use of biomechanical knowledge. This observation underlines the need to assess the influence of

---

[1] Several authors (e.g., Tomasino et al., 2003; Steenbergen et al., 2007; Daprati et al., 2010; ter Horst et al., 2010) argued that, when motor simulation is not possible (e.g., aplasia or severe motor deficit), alternative strategies may be used to perform the HLJ task. These strategies imply, for instance, mental rotation of the visual stimulus (Tomasino et al., 2003), visual analysis based on non-motor features of the hands such as the number of visible fingers or the position of the thumb and/or pinkie (Daprati et al., 2010), or third-person visual imagery of the action "as if someone else performed the action" (Steenbergen et al., 2007). It is important to note, however, that all these visually based strategies predict no effect of biomechanical constraints on response latencies. Thus, any account for DC's pattern in the HLJ task in terms of such alternative strategies would be inappropriate.

biomechanical knowledge on HLJs with an extended set of hand drawings.

Finally, our findings call for re-examining the motor imagery account for the effect of biomechanical constraints in normally limbed individuals performing HLJs. Obviously, the effect observed in DC and controls could be driven by distinct kinds of representations and processes, i.e., visual vs. motor processes, both being sensitive to the biomechanics of the human body. Nevertheless, we find it unlikely that an implicit and automatic access to biomechanics, such as the one revealed in DC's laterality judgments, would be totally suspended when normal subjects perform the task. Our findings thus encourage further empirical studies to entertain the hypothesis that visual in addition to motor processes, or even visual processes alone, contribute to the effect of biomechanical constraints in the HLJ task. In the meanwhile, the HLJ task should not be considered anymore as an unambiguous window on the covert stages of motor control and motor planning in normal or motor-impaired individuals.

ACKNOWLEDGMENTS

We are indebted to DC and the control participants for their motivation during the experiments carried out in this study. We are also very grateful to Lawrence Parsons who kindly provided us with the hand and foot drawings used in this study. Agnesa Pillon is a senior researcher, Gilles Vannuscorps is a research fellow, and Michael Andres was a post-doctoral researcher at the Fonds National de la Recherche Scientifique (FNRS). This research was supported by the FNRS grant no. 1.5.192.10F to Agnesa Pillon.

REFERENCES

Brugger, P., Kollias, S. S., Müri, R. M., Crelier, G., Hepp-Reymond, M. C., and Regard, M. (2000). Beyond re-membering: phantom sensations of congenitally absent limbs. *Proc. Natl. Acad. Sci. U.S.A.* 97, 6167–6172.

Chatterjee, S. H., Freyd, J. J., and Shiffrar, M. (1996). Configural processing in the perception of apparent biological motion. *J. Exp. Psychol. Hum. Percept. Perform.* 22, 916–929.

Cooper, L. A., and Shepard, R. N. (1975). Mental transformation in the identification of left and right hands. *J. Exp. Psychol. Hum. Percept. Perform.* 1, 48–56.

Corradi-Dell'Acqua, C., and Tessari, A. (2010). Is the body in the eye of the beholder? Visual processing of bodies in individuals with anomalous anatomical sensory and motor features. *Neuropychologia* 48, 689–702.

Crawford, J. R., and Garthwaite, P. H. (2005). Testing for suspected impairments and dissociations in single-case studies in neuropsychology: evaluation of alternatives using monte carlo simulations and revised tests for dissociations. *Neuropsychology* 19, 318–331.

Crawford, J. R., and Howell, D. C. (1998). Comparing an individual's test score against norms derived from small samples. *Clin. Neuropsychol.* 12, 482–486.

Daprati, E., Nico, D., Duval, S., and Lacquaniti, F. (2010). Different motor imagery modes following brain damage. *Cortex* 46, 1016–1030.

de Lange, F. P., Helmich, R. C., and Toni, I. (2006). Posture influences motor imagery: an fMRI study. *Neuroimage* 33, 609–617.

de Lange, F. P., Roelofs, K., and Toni, I. (2008). Motor imagery: a window into the mechanisms and alterations of the motor system. *Cortex* 44, 494–506.

de Vignemont, F. (2010). Body schema and body image: pros and cons. *Neuropsychologia* 48, 669–680.

Fiorio, M., Tinazzi, M., and Aglioti, S. M. (2006). Selective impairment of hand mental rotation in patients with focal hand dystonia. *Brain* 129, 47–54.

Funk, M., and Brugger, P. (2008). Mental rotation of congenitally absent hands. *J. Int. Neuropsychol. Soc.* 14, 81–89.

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

Ganis, G., Keenan, J. P., Kosslyn, S. M., and Pascual-Leone, A. (2000). Transcranial magnetic stimulation of primary motor cortex affects mental rotation. *Cereb. Cortex* 10, 175–180.

Helmich, R. C., de Lange, F. P., Bloem, B. R., and Toni, I. (2007). Cerebral compensation during motor imagery in Parkinson's disease. *Neuropsychologia* 45, 2201–2215.

Jeannerod, M. (1994). The representing brain: neural correlates of motor intention and imagery. *Behav. Brain Sci.* 17, 187–245.

Jeannerod, M., and Frak, V. (1999). Mental imaging of motor activity in humans. *Curr. Opin. Neurobiol.* 9, 735–739.

Kolers, P. A., and Pomerantz, J. R. (1971). Figural change in apparent motion. *J. Exp. Psychol.* 87, 99–108.

Kosslyn, S. M., DiGirolamo, G. J., Thompson, W. L., and Alpert, N. M. (1998). Mental rotation of objects versus hands: neural mechanisms revealed by positron emission tomography. *Psychophysiology* 35, 151–161.

Kosslyn, S. M., Ganis, G., and Thompson, W. L. (2001). Neural foundations of imagery. *Nat. Rev. Neurosci.* 2, 635–642.

Kourtzi, Z., and Shiffrar, M. (1999). Dynamic representations of human body movement. *Perception* 28, 49–62.

Marr, D., and Vaina, L. (1982). Representation and recognition of the movements of shapes. *Proc. R. Soc. Lond. B Biol. Sci.* 214, 501–524.

Munzert, J., Lorey, B., and Zentgraf, K. (2009). Cognitive motor processes: the role of motor imagery in the study of motor representations. *Brain Res. Rev.* 60, 306–326.

Nico, D., Daprati, E., Rigal, F., Parsons, L., and Sirigu, A. (2004). Left and right hand recognition in upper limb amputees. *Brain* 127, 120–132.

Parsons, L. M. (1987). Imagined spatial transformations of one's hands and feet. *Cogn. Psychol.* 19, 178–241.

Parsons, L. M. (1994). Temporal and kinematic properties of motor behavior reflected in mentally simulated action. *J. Exp. Psychol. Hum. Percept. Perform.* 20, 709–730.

Parsons, L. M., Gabrieli, J. D., Phelps, E. A., and Gazzaniga, M. S. (1998). Cerebrally lateralized mental representations of hand shape and movement. *J. Neurosci.* 18, 6539–6548.

Pelgrims, B., Andres, M., and Olivier, E. (2009). Double dissociation between motor and visual imagery in the posterior parietal cortex. *Cereb. Cortex* 19, 2298–2307.

Pelgrims, B., Michaux, N., Olivier, E., and Andres, M. (2010). Contribution of the primary motor cortex to motor imagery: a subthreshold TMS study. *Hum. Brain Mapp.* 32, 1802–1810.

Sauner, D., Bestmann, S., Siebner, H. R., and Rothwell, J. C. (2006). No evidence for a substantial involvement of primary motor hand area in handedness judgements: a transcranial magnetic stimulation study. *Eur. J. Neurosci.* 23, 2215–2224.

Schwoebel, J., and Coslett, H. B. (2005). Evidence for multiple, distinct representations of the human body. *J. Cogn. Neurosci.* 17, 543–553.

Schwoebel, J., Friedman, R., Duda, N., and Coslett, H. B. (2001). Pain and the body schema: evidence for peripheral effects on mental representations of movement. *Brain* 124, 2098–2104.

Sekiyama, K. (1982). Kinesthetic aspects of mental representations in the identification of left and right hands. *Percept. Psychophys.* 32, 89–95.

Shiffrar, M., and Freyd, J. J. (1990). Apparent motion of the human body. *Psychol. Sci.* 1, 257–264.

Shiffrar, M., and Freyd, J. J. (1993). Timing and apparent motion path choice with human body photographs. *Psychol. Sci.* 4, 379–384.

Sirigu, A., Grafman, J., Bressler, K., and Sunderland, T. (1991). Multiple representations contribute to body knowledge processing. Evidence from a case of autotopagnosia. *Brain* 114, 629–642.

Steenbergen, B., van Nimwegen, M., and Crajé, C. (2007). Solving a mental rotation task in congenital hemiparesis: motor imagery versus visual imagery. *Neuropsychologia* 45, 3324–3328.

Stevens, J. A., Fonlupt, P., Shiffrar, M. A., and Decety, J. (2000). New aspects of motion perception: selective neural encoding of apparent human movements. *Neuroreport* 11, 109–115.

ter Horst, A. C., van Lier, R., and Steenbergen, B. (2010). Mental rotation task of hands: differential influence number of rotational axes. *Exp. Brain Res.* 203, 347–354.

Tomasino, B., Rumiati, R. I., and Umilta, C. A. (2003). Selective deficit of motor imagery as tapped by a left-right decision of visually presented hands. *Brain Cogn.* 53, 376–380.

Wraga, M., Shephard, J. M., Church, J. A., Inati, S., and Kosslyn, S. M. (2005). Imagined rotations of self versus objects: an fMRI study. *Neuropsychologia* 43, 1351–1361.

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

Received: 04 May 2012; accepted: 14 October 2012; published online: 01 November 2012.

Citation: Vannuscorps G, Pillon A and Andres M (2012) Effect of biomechanical constraints in the hand laterality judgment task: where does it come from? Front. Hum. Neurosci. ***6****:299. doi: 10.3389/fnhum.2012.00299*

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 10 August 2012
doi: 10.3389/fpsyg.2012.00284

Understanding immersivity: image generation and transformation processes in 3D immersive environments

Maria Kozhevnikov[1,2]* **and Rupali P. Dhond**[2]

[1] Psychology Department, National University of Singapore, Singapore
[2] Department of Radiology, Martinos Center for Biomedical Imaging, Harvard Medical School, Charlestown, MA, USA

Edited by:
Joel Pearson, The University of New South Wales, Australia

Reviewed by:
Natasha Sigala, University of Sussex, UK
Adriane E. Seiffert, Vanderbilt University, USA

***Correspondence:**
Maria Kozhevnikov, Department of Radiology, Martinos Center for Biomedical Imaging, Harvard Medical School, 149 Thirteenth Street, Suite 2301, Charlestown, MA 02129, USA.
e-mail: mkozhevn@nmr.mgh.harvard.edu

Most research on three-dimensional (3D) visual-spatial processing has been conducted using traditional non-immersive 2D displays. Here we investigated how individuals generate and transform mental images within 3D immersive (3DI) virtual environments, in which the viewers perceive themselves as being surrounded by a 3D world. In Experiment 1, we compared participants' performance on the Shepard and Metzler (1971) mental rotation (MR) task across the following three types of visual presentation environments; traditional 2D non-immersive (2DNI), 3D non-immersive (3DNI – anaglyphic glasses), and 3DI (head mounted display with position and head orientation tracking). In Experiment 2, we examined how the use of different backgrounds affected MR processes within the 3DI environment. In Experiment 3, we compared electroencephalogram data recorded while participants were mentally rotating visual-spatial images presented in 3DI vs. 2DNI environments. Overall, the findings of the three experiments suggest that visual-spatial processing is different in immersive and non-immersive environments, and that immersive environments may require different image encoding and transformation strategies than the two other non-immersive environments. Specifically, in a non-immersive environment, participants may utilize a scene-based frame of reference and allocentric encoding whereas immersive environments may encourage the use of a viewer-centered frame of reference and egocentric encoding. These findings also suggest that MR performed in laboratory conditions using a traditional 2D computer screen may not reflect spatial processing as it would occur in the real world.

Keywords: mental rotation, immersivity, three-dimensional immersive virtual environments

INTRODUCTION

Our ability to generate and transform three-dimensional (3D) visual-spatial images is important not only for our every-day activities (locomotion, navigation) but also for a variety of professional activities, such as architecture, air traffic control, and telerobotics. Difficulties of studying visual-spatial cognition within real world environments, where controlling the experimental stimuli and recording participants' behavior is often impossible, have led researchers to increasingly employ 3D immersive (3DI) virtual environments (Chance et al., 1998; Klatzky et al., 1998; Loomis et al., 1999; Tarr and Warren, 2002; Macuga et al., 2007; Kozhevnikov and Garcia, 2011). Specifically, 3DI technology allows one to create a complex immersive environment of high ecological validity, in which participants are presented with and manipulate a variety of 3D stimuli under controlled conditions.

An immersive virtual environment involves computer simulation of a 3D space and a human computer-interaction within that space (Cockayne and Darken, 2004). There are two major characteristics of 3DI environments that distinguish them from non-immersive 2D non-immersive (2DNI) and 3D non-immersive (3DNI) environments. First, 3DI involves *egocentric* navigation (the user is surrounded by the environment) rather than *exocentric* navigation where the user is outside the environment, looking in. Second, unlike non-immersive environments where a scene is fixed on a 2D computer screen, 3DI involves image updating achieved by position and head orientation tracking. Although little is known about cognitive processes and neural dynamics underlying image encoding and transformation in 3DI environments, researchers have speculated that immersivity would differentially affect selection of a spatial frame of reference (i.e., spatial coordinate system) during object encoding processes (Kozhevnikov and Garcia, 2011).

Two different spatial frames of reference, *environmental* and *viewer-centered*, can be used for encoding and transforming visual-spatial images. An environmental frame may involve the "permanent environment" which is bound by standard orthogonal planes, i.e., the floor, walls, ceiling, and perceived direction of gravity or the local "scene-based" spatial environment where the target object's components are encoded allocentrically in relation to another object, i.e., table-top, blackboard, computer screen, etc. In contrast to environmental frames of reference, the viewer-centered frame is egocentric, that is, it defines object configurations and orientations relative to the viewer's gaze and it includes an embedded retinal coordinate system. In the case of imagined spatial transformations such as mental rotation (MR), the prevailing hypothesis is that individuals rely more upon an environmental, scene-based, rather than a viewer-centered frame of reference (Corballis et al., 1976, 1978; Rock, 1986; Hinton and Parsons, 1988; Palmer, 1989; Pani and Dupree, 1994). For example, Corballis et al. tested normal-mirror discriminations of rotated alphanumeric characters when participants' heads or bodies were either aligned with the gravitational vertical or misaligned by up to 60°. The results showed that the participants made their judgments by rotating the characters to the gravitational vertical (*Y* axis) rather than using a viewer-centered (head-centered or retina-centered) reference frame. Furthermore, Hinton and Parsons (1988) reported that while mentally rotating two shapes positioned on a table into congruence, participants often rotated one shape until it had the same relationship to the table-top (and room) as the other shape (thus achieving scene-based alignment), even though this produced quite different retinal images. Thus, it appears that the orientation of the viewer is defined relative to the scene, rather than the orientation of the scene being defined relative to the viewer. This lends support for theories suggesting that the representation of spatial relationships is established primarily in terms of scene-based reference systems.

Additional evidence for primacy of scene-based reference frames comes from experiments (e.g., Parsons, 1987, 1995) comparing the speed of MR of classical Shepard and Metzler's (1971) 3D forms around different axes (see **Figure 1A**). MR around different axes places different demands on the transformation processes, and results in different brain activity (Gauthier et al., 2002). Rotation in the picture plane preserves the feasibility of all the features of a shape, but perturbs the top-bottom relations between features. Rotation in depth around the vertical axis alters side-to-side relationships between features and the visibility of features, some coming into view and others becoming occluded. Rotation in depth around a horizontal axis is the most demanding rotation; it alters top-bottom relations between features and feature visibility. Interestingly, it has been consistently found that participants mentally rotate shapes in the depth plane just as fast as or even faster than in the picture plane (Shepard and Metzler, 1971; Parsons, 1987, 1995). If participants were in fact rotating viewer-centered 2D retina-based visual representations, the depth rotation would take longer than rotation in the picture plane since rotation in depth would have to carry out additional foreshortening and hidden line removal operations, not required during picture plane rotation.

Shepard and Metzler (1971) were the first to interpret similar slopes for rotation in depth and in the picture plane to indicate that latency was a function of the angle of rotation in three dimensions, not two, as in a retinal projection (for additional discussion see Pinker, 1988). In order to investigate this further, Parsons (1987) conducted an extensive experimental study examining the rates of imagined rotation not only around three principal axes of the observer's reference frame, but also around diagonal axes lying within one of the principal planes (frontal, midsagittal, or horizontal) and around "skew" axes not lying in any of the principal planes. The findings indicated that the rotation around different axes, including rotation in depth around a horizontal axis perpendicular to the line of sight (*Z* axis, see **Figure 1B**) were as fast as or even faster than rotations in the picture plane (rotations around the axis defined by the line of sight, *X*-axis in this study). Parsons concluded that this equal ease of rotating images around different axes support scene-based encoding, during which the observers rely largely on representations containing more "structural" information (e.g., information about spatial relations among the elements of the object and their orientations with respect to the scene in which the objects lie) rather than on retina-based 2D representations of visual-spatial images.

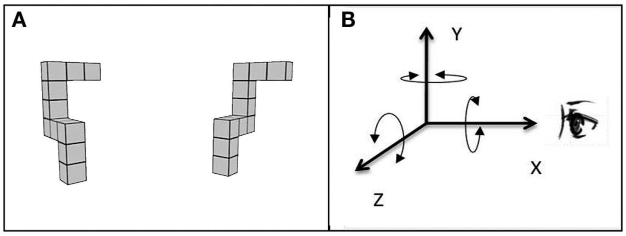


FIGURE 1 | Example of the MR test trial stimulus with (A) showing the subject's view of the stimulus and (B) showing the three principle axes of rotation, *X*, *Y*, and *Z*, used in the current study. Note that this coordinate frame differs from the one normally used in computer graphics in which the positive *Z*-direction is perpendicular to the plane of the display, pointing toward the viewer.

One limitation of previous studies on MR is that they have been conducted using traditional non-immersive environments, where the stimuli were presented on a 2D computer screen or another flat surface (e.g., a table-top), which defines a fixed local frame of reference. This limited and fixed field of view (FOV) may encourage the use of a more structural scene-based encoding, during which the parts of the 3D image are encoded in relation to the sides of the computer screen or another salient object in the environment. However, because 3DI environments enclose an individual within the scene and allow images to be updated with respect to the observer's head orientation, egocentric, viewer-centered encoding may predominate.

The primary goal of the current research was to examine how individuals process visual-spatial information (specifically encode and rotate 3D images) and what spatial frames of reference they rely upon in 3DI virtual environments vs. conventional non-immersive displays. In our first experiment, in order to control the effect of "three-dimensionality" vs. "immersivity," we compared participants' performance on the Shepard and Metzler (1971) MR task across the following three types of environments; traditional 2DNI, 3DNI (anaglyphic glasses), and 3DI [head mounted display (HMD) with position and head orientation tracking]. In the second experiment, we compared how participants encode and transform visual-spatial images in different 3DI environments with different backgrounds where shapes were embedded in a realistic scene vs. in a rectangular frame. Furthermore, if the neurocognitive correlates of visual-spatial imagery are affected by immersivity of visual presentation environment, this should be evidenced in the underlying temporal dynamic and/or spatial distribution of the electroencephalogram (EEG) response. Thus, in the third experiment, EEG was recorded while participants performed the MR task in 3DI and 2DNI environments.

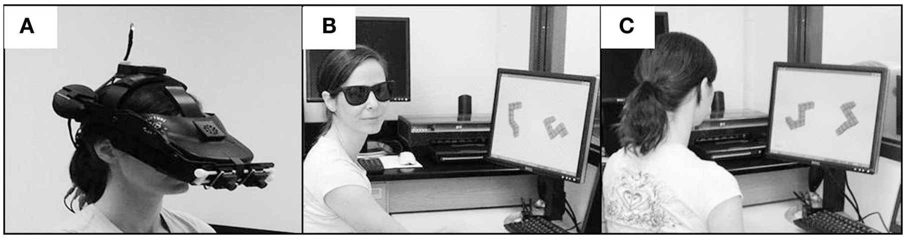


FIGURE 2 | Three different viewing environments (A) 3DI, which includes HMD with position tracking, (B) 3DNI with anaglyphic glasses to present a stereo picture of three-dimensional spatial forms, and (C) 2D monocular viewing environment.

EXPERIMENT 1

MATERIALS AND METHODS

Participants

Fourteen volunteers (eight males and eight females, average age = 21.5) participated in the study for monetary compensation. The study was approved by George Mason University (Fairfax, VA, USA) as well as by The Partners Human Research Committee (PHRC, MA, USA) and informed consent was obtained from all subjects. Participants were asked about their ability to perceive stereoscopic images prior to the start of the experiment, and only those who did report difficulty with stereopsis were included.

Materials and design

Each participant completed the MR task – a computerized adaptation of Shepard and Metzler's (1971) task – in three different viewing environments: 3DI, 3DNI, and 2DNI. For each trial, participants viewed two spatial figures, one of which was rotated relative to the position of the other (**Figure 1A**). Participants were to imagine rotating one figure to determine whether or not it matched the other figure and to indicate whether they thought the figures were the same or different by pressing a left (same) or right (different) button on a remote control device. Participants were asked to respond as quickly and as accurately as possible. Twelve rotation angles were used: 20, 30, 40, 60, 80, 90, 100, 120, 140, 150, 160, and 180°. The figures were rotated around three spatial axes: line of sight (X), vertical (Y), and horizontal (Z) corresponding to rotations parallel with the frontal (YZ), horizontal (XZ), and midsagittal (XY) anatomical planes, respectively (**Figure 1B**). The test included: 12 trial groups for the 12 rotation angles, 3 trial pairs for the 3 axes, and each pair had 1 trial with matching figures and 1 trial with different figures; thus, there were 72 ($12 \times 3 \times 2$) trials in total.

In the 3DI virtual environment, the shapes were presented to the participant through an nVisor SX60 (by Nvis, Inc.) HMD (**Figure 2A**). The HMD has a 44° horizontal by 34° vertical FOV with a display resolution of 1280×1024 and under 15% geometric distortion. During the experiment, participants sat on a chair in the center of the room, wearing the HMD to view "virtual" Shepard and Metzler images in front of them. Sensors on the HMD enabled real-time simulation in which any movement of the subject's head immediately caused a corresponding change to the image rendered in the HMD. The participant's head position was tracked by four cameras located in each corner of the experimental room and sensible to an infrared light mounted on the top of the HMD. The rotation of user's head was captured by a digital compass mounted on the back of the HMD.

In the 3DNI environment, the shapes were presented to the participant on a computer screen. Stereoscopic depth was provided by means of anaglyphic glasses (**Figure 2B**). In the 2DNI environment, the shapes were presented for on a standard computer screen (**Figure 2C**).

The retinal image size of the stimuli was kept constant across all the environments (computed as ratio of image size over the participant's distance to the screen). The Vizard Virtual Reality Toolkit v. 3.0 (WorldViz, 2007) was used to create the scenes and to record the dependent variables (latency and accuracy).

Before beginning the MR trials, participants listened to verbal instructions while viewing example trials in each environment. Eight practice trials were given to ensure participants' comprehension of the instructions and that they were using a MR strategy (as opposed to a verbal or analytical strategy). If a response to a practice trial was incorrect, the participants were asked to explain how they solved the task in order to ensure the use of a rotation strategy (i.e., rather than verbal or analytical strategy). In 3DI, to familiarize the participants with immersive virtual reality, there was also an exploratory phase prior to the practice trials in which the participants were given general instructions about virtual reality and the use of the remote control device (about 7–10 min). During the practice and test phases, the participants remained seated in the chair, but were allowed to move and rotate their head to view 3D Shepard and Metzler shapes. The participants were also given similar time to familiarize themselves with the shapes in the 3DNI and 2DNI environment, and were also allowed to move and rotate their head to view Shepard and Metzler shapes.

RESULTS

Descriptive statistics for performance in the three environments are given in **Table 1**. Outlier response times (RTs; i.e., RTs ± 2.5 SD from a participant's mean) were deleted (a total of 2.59% of all trials). All simple main effects were examined using the Bonferroni correction procedure. Two participants that performed below chance level were not included in the analysis, thus the final analysis was performed on 12 participants only.

Response accuracy (proportion correct) and RT for correct responses were assessed as a function of the rotation axis (X, Y, and Z) and environment (3DI, 3DNI, and 2DNI). Data were analyzed using a 3 (axis) × 3 (environment) repeated measures

ANOVA with a General Linear Model (GLM). The effect of environment was marginally significant [$F(2{,}22) = 2.9, p = 0.040$] and as pairwise comparison showed, the accuracy in 3DNI and 3DI environments was slightly less than in 2DNI ($p = 0.08$). There was a significant main effect of axis [$F(2{,}22) = 19.83, p < 0.001$] where Y axis rotations were more accurate than X and Z axis rotations ($ps < 0.01$). The interaction was not significant ($F < 1$). Overall, the accuracy level was relatively high for all the environments and all axes, with the proportion correct ranging from 0.84 to 0.97. Given the high rate of accuracy, indicating that ceiling performance was reached for some rotations, we focused our remaining analyses on the RTs.

With respect to RT, there was a significant effect of axis [$F(2{,}22) = 15.40, p < 0.001$] with Y axis rotations being the fastest ($ps < 0.05$), see **Figure 3**. There was no significant effect of environment ($F < 1$), however, there was a significant interaction between axis and test environment [$F(4{,}44) = 6.45$, $p < 0.001$]. Analysis of simple main effects revealed that, RT for rotation around the Y axis was significantly faster than either around X (all $ps < 0.05$) or Z (all $ps < 0.05$) for 3DNI and 2DNI environments, consistently with previous studies (Shepard and Metzler, 1971; Parsons, 1987, 1995). However, rotations around X and Z axes were similar ($p = 0.98$ and 0.79 for 2DNI and 3DNI respectively). Interestingly the opposite occurred for MR in the 3DI environment. In 3DI, rotation around Z was significantly longer than X ($p = 0.001$) or Y ($p = 0.01$), while rotations around X and Y were similar ($p = 0.97$).

Table 1 | Descriptive statistics for three versions of the MR test in 2DNI, 3DNI, and 3DI.

Test	Proportion correct	SD	RT (s)	SD
2D	0.90	0.07	5.33	1.02
3D non-immersive	0.86	0.10	5.47	1.64
3D immersive	0.87	0.09	5.42	1.46

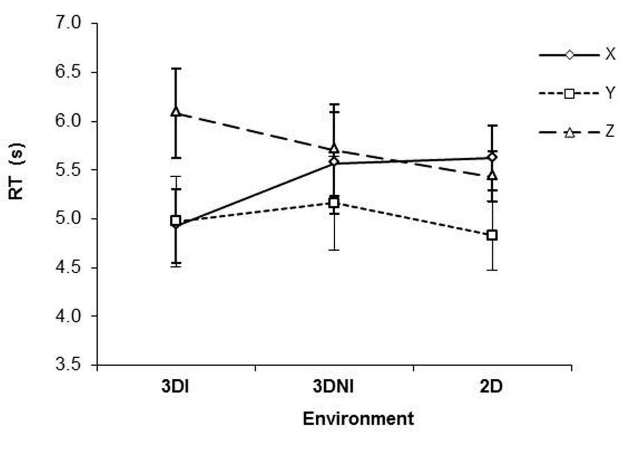


FIGURE 3 | Response time as a function of axis of rotation and viewing environment (2DNI, 3DNI, and 3DI). Error bars represent standard error means.

Thus, our central finding is that in 3DI, the RT of rotation differed between X and Z axes (Z was slower) and that rotation around the Y axis was faster than Z but not faster than X rotations. In contrast, RT patterns for 2DNI and 3DNI environments were similar to those found in previous MR studies (i.e., Y rotations are faster than X and Z and X and Z are similar).

Rate of rotation as a function of axis and environment. RT as a function of rotation angles (i.e., orientation differences between two Shepard and Metzler shapes) around X, Y, and Z axes for 3DI, 3DNI, and 2DNI environments respectively are shown in **Figure 4**.

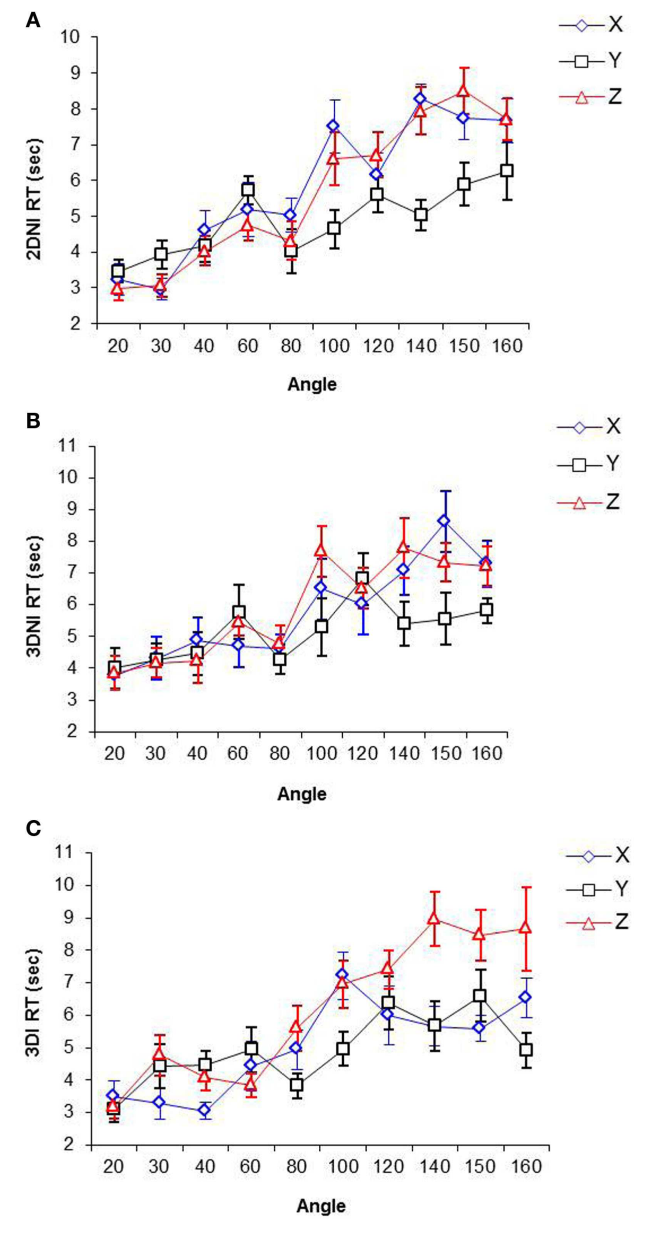


FIGURE 4 | Response time as a function of angle and axis of rotation in (A) 2DNI, (B) 3DNI, and (C) 3DI environments.

The range of rotation angles was from 20 to 160; 180° was omitted due to participant's reports that for this particular angle, they did not rotate shapes mentally, but only scanned two images for mirror-reversed symmetry.

The slopes of the best-fit linear RT-Rotation Angle functions for each axis and each environment (representing *rates of rotation* around different axes in different environments) were computed and are presented in **Table 2**.

A repeated measures ANOVA of slopes of best-fit linear regression equations of RT on Rotation Angle show a significant effect of axis on the slope [$F(2,22) = 51.34$, $p < 0.001$] and a significant interaction between environment and axis of rotation [$F(4,44) = 3.38$, $p < 0.05$], while the effect of environment is not significant. For 3DI, the rate of rotation around Z was more than 1.5 times slower than around X ($p < 0.05$). In both 3DNI and 2DNI, the rate of rotation around X and Z did not differ. Across the environments, the rate of rotation around X was significantly faster for 3DI than for 2DNI ($p < 0.05$), whereas the rate of rotation around Z was significantly slower for 3DI than for either 3DNI or 2DNI ($ps < 0.05$). There were no significant differences in the rate of rotation around the Y axis across environment, and the rate of rotation around Y seems to be the one of the fastest rotations. This is consistent with the findings of previous investigators (Rock and Leaman, 1963; Attneave and Olson, 1967; Parsons, 1987; Corballis, 1988) who argued that rotation around Y, a "gravitational vertical" axis, is the most common of all rotations in our ecology, so that the fast rate of rotation around it may result from our extraexperimental familiarity.

In summary, the results of Experiment 1 show that the rate of MR about the horizontal axis (Z axis) in 3DI (and only 3DI) was significantly slower than the rate of rotation about the line of sight (X-axis). This finding suggests that in the 3DI environment the participants were encoding and rotating 2D retina-based visual representations in relation to a *viewer-centered* frame of reference since only then would depth rotation take longer than rotation in the picture plane, due to the involvement of additional foreshortening and hidden line removal transformations. In contrast, in 2DNI and 3DNI environments, the rates of MR around the X and Z axes were not different, consistent with previous findings for MR using 2D traditional computer displays (Shepard and Metzler, 1971; Parsons, 1987). Thus, in non-immersive environments, participants seem to generate visual representations containing more allocentric information such as information about spatial relations among the elements of the object and their orientations with respect to the scene (i.e., the computer screen) in which the object is presented. The fact that there was equivalent performance in 2D and 3DNI environments suggests that depth information *per se*, which is provided in a 3DNI environment is insufficient to encourage the use of viewer-centered frame of reference.

Table 2 | Mean regression slopes of RT-Rotation angle function (s/°).

Environment	Axis of rotation		
	X	***Y***	***Z***
3DI	0.028	0.014	0.043
3DNI	0.029	0.013	0.031
2DNI	0.032	0.016	0.036

EXPERIMENT 2

One of possible limitations of Experiment 1 is that, in the 3DI environment, the Shepard and Metzler shapes were presented to participants on a non-realistic "empty" background lacking any points of reference (such as ceilings, walls, other objects), which would usually be present in a real scene. Thus, viewer-centered encoding observed in 3DI could be due not to the immersivity of the environment, but rather due to the lack of any other objects – except the observers themselves – in relation to which Shepard and Metzler shapes could have been encoded. In Experiment 2, the participants in a 3DI condition were presented with Shepard and Metzler forms embedded in a realistic scene (city). In addition, we added a second condition in which the participants viewed Shepard and Metzler forms embedded in a virtual rectangular-shaped frame within the 3DI environment. This was done to examine whether the fixed frame around objects in a 3DI environment induces scene-based encoding similar to a computer screen in the real world.

MATERIALS AND METHODS

Participants

Twenty-six volunteers (10 males and 16 females, average age = 20) recruited by advertisement participated in the study for monetary compensation. The study was approved by George Mason University (Fairfax, VA, USA) as well as by The Partners Human Research Committee (PHRC, MA, USA) and the informed consent was obtained from all subjects. Participants who reported difficulty with stereopsis were excluded from participation.

Materials and design

Each participant completed the MR task used in Experiment 1 in three different viewing environments: 3DIC (3DI City) environment, where Shepard and Metzler forms were embedded in a realistic scene of a city (**Figure 5A**), 3DF (3D Frame), where Shepard and Metzler forms where embedded in a rectangular frame (**Figure 5B**), and a 2DNI environment. The order of environments was counterbalanced. The experimental procedure was similar to that described in Experiment 1.

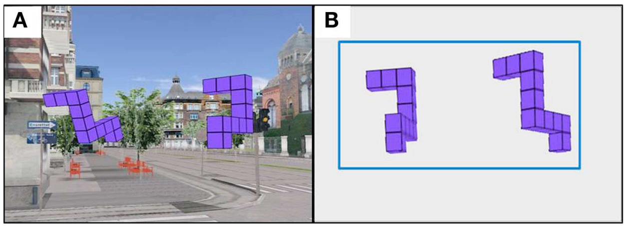


FIGURE 5 | Two different viewing environments (A) 3DIC (3DI City) environment, (B) 3DF (3D Frame) environment.

RESULTS

Descriptive statistics for proportion correct and RT correct for 2DNI, 3DF, and 3DIC environments are given in **Table 3**. Outlier RTs (i.e., RTs ± 2.5 SD from a participant's mean) were deleted (a total of 3.84% of all trials). One participant that performed below chance level was not included in the analysis.

The overall accuracy (proportion correct) ranged from 0.75 to 0.81. For accuracy, a 3 (axis) × 3 (environment) repeated measures ANOVA revealed a significant effect of axis [$F(2,48) = 5.38$, $p < 0.01$], where Z axis rotations were significantly less accurate than either Y or X (ps < 0.05). There was a significant effect of environment $F(2,48) = 13.59$, $p < 0.001$]. Participants were significantly less accurate in the 3DIC environment than in either the 2DNI or 3DIF environments (ps < 0.01), while 2DNI and 3DIF did not differ among themselves. Also, there was a significant interaction between axis and environment [$F(4,96) = 26.04$, $p < 0.001$]. Accuracy for rotation around the Y axis was higher than for rotations around X in 2DNI ($p < 0.01$), but there was no differences between the accuracy of X and Y rotation in 3DF and 3DIC.

Figure 6 presents RT against viewing environment for each axis. For RT, there was a significant effect of axis [$F(2,48) = 35.10$, $p < 0.001$], where Y axis rotations were significantly faster than Z ($p < 0.01$) or X ($p = 0.05$). There was no significant effect of environment [$F(2,48) = 3.85$, $p = 0.18$]. However, there was a significant interaction between axis and test environment [$F(4,96) = 6.18$, $p < 0.001$]. Examination of simple main effects revealed that rotation around the Y axis was faster than around X in 2DNI ($p < 0.001$) and in 3DIF ($p = 0.07$). In addition, rotation around Y was faster than around Z in both 2DNI and 3DF (all ps < 0.001). However, rotations around X and Z axes were similar ($p = 0.14$ and 0.24 for 2DNI and 3DF respectively).

In contrast, for 3DIC, rotation around Z was significantly longer than around X or Y (ps < 0.001), while rotations around X and Y were similar ($p = 0.78$). Thus, our central finding is that in 3DIC the RT of rotation differed between X and Z axes (Z was slower) and that rotation around the Y axis was faster than Z but not faster than around X. This pattern is similar to what we reported for the 3DI environment in Experiment 1. In contrast, RT patterns for 2DNI and 3DF are similar to those found reported for 2DNI environment in Experiment 1.

In conclusion, the findings of Experiment 2 suggest that individuals encode and rotate visual-spatial representations in relation to a viewer-centered frame only in 3DI environments in which the viewer is enclosed in the scene. Presumably, the participants used a viewer-centered reference frame because they perceived themselves to be "inside" the city scene and a scene-based reference frame when they perceived themselves to be "outside" the square-shaped frame that enclosed the stimuli. The square-shaped frame is similar to the sides of the computer screen in non-immersive environments, suggesting that whenever a person is observing a scene from the "outside" (e.g., a scene is defined by the frame or computer screen), it encourages the use of "scene-based" encoding.

Table 3 | Descriptive statistics for the MR test in 2DNI, 3DF, and 3DIC ($N = 25$).

Environment	Proportion correct	SD	RT correct (s)	SD
2DNI	0.81	0.09	5.88	1.76
3DF	0.76	0.09	6.19	1.87
3DIC	0.75	0.08	6.32	1.76

EXPERIMENT 3

Previous EEG research suggests that MR task involves four sequential cognitive stages which may also differentially modulate frontal and posterior brain (Desrocher et al., 1995). The first stage at ∼200–300 ms post-stimulus is independent of the object's angular disparity and involves early sensory processing and simple stimulus evaluation. Subsequently, at ∼300–400 ms a pre-rotation "set-up" stage involves evaluation of object orientation and rotation strategy selection. Third, is the act of MR at ∼400–800 ms post-stimulus which is followed by response selection and execution from ∼1000 ms onward.

Object encoding with respect to a specific frame of reference occurs prior to the actual process of MR. Thus, the selection of a frame of reference should begin in the earliest cognitive stages between ∼200 and 400 ms post-stimulus. Furthermore, the results of Experiments 1 and 2 demonstrate that selection of a frame of reference is determined by the viewing environment. Thus, we hypothesized that when performing a MR task in 2DNI vs. 3DI, brain response differences should be largest at early sensory and/or pre-rotation "set-up" stages occurring at ∼200–400 ms post-stimulus.

MATERIALS AND METHODS

Participants

Eight undergraduate psychology students (four males and four females) from the National University of Singapore (age between 19 and 25 years) participated in the study for monetary reimbursement. The study was approved by National University of Singapore committee and informed consent was obtained from all subjects.

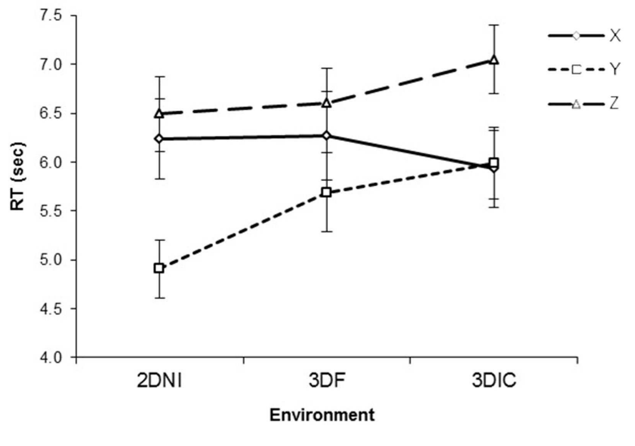


FIGURE 6 | Response time as a function of axis of rotation and viewing environment (2DNI, 3DIF, and 3DIC). Error bars represent standard error means.

Participants who reported difficulty with stereopsis were excluded from participation.

Design and analysis

Electroencephalogram was recorded while subjects completed the MR task in 2DNI and 3DI viewing environments. The order of environments was counterbalanced and, in general, the procedures were similar to the first two experiments except as follows. EEG was recorded using a 256-channel HydroCel Geodesic Sensor Net (Electrical Geodesics, Inc.). Signals were amplified using the EGI NetAmps 300 amplifier. The signal was sampled at 250 Hz and bandpass filtered online at 1.0−100 Hz. For the 3DI condition, the HMD was placed directly on top of the sensor net (**Figure 7**).

In order to assure an adequate number of trials for averaging, participants were administered 2 blocks of 72 randomly ordered trials in the 3DI condition (overall 144 trials), and then another 2 blocks of 72 trials in the 2DNI condition.

Data preprocessing and analysis were performed using a combination of EEGLAB (Delorme and Makeig, 2004), Neuromag software (Elekta, Stockholm) and MNE Software (http://www.martinos.org/mne/). Raw EEG signals were first low-pass filtered at 40 Hz to eliminate 50 Hz electrical noise generated by the 3DI headset. Eye blink artifacts were then removed using the Signal Space Projection method provided within the open source MNE toolbox. The resultant cleaned raw data was used to create ERP averages from −200 ms pre-stimulus to 800 ms post-stimulus. Baseline was defined as −150 to 0 ms.

RESULTS

All participants demonstrated clear centroparietal responses while performing the MR task in the 2DNI and also the 3DI environment (**Figure 8**). In the 2DNI environment, parietal ERPs were highly similar for shapes rotated around the *X* and *Z* axes but more negative for *Y* axis rotations from ∼250 ms onward (**Figure 8A**). In the 3DI environment *X*, *Z*, and *Y* rotation demonstrated increasing negativities at ∼350 ms (**Figure 8B**). When comparing rotations between 3DI and 2DNI environments, rotations in 3DI showed a trend for greater negativity prior to ∼250 ms but, larger positivity at latencies >350 ms (**Figures 8C–E**). This effect was significant for *Z*-rotations which were more negative at ∼270–300 ms post-stimulus for MR task in the 3DI environment.

FIGURE 7 | Electroencephalogram recording in 3DI virtual reality environment.

The results of Experiment 3 demonstrate that 2DNI and 3DI environments do evoke differential parietal ERP response. The early latency of these ERP differences further supports our hypothesis that 3DI and 2DNI environments differ at the level of shape encoding with respect to selection of a frame of reference. Our finding of significantly larger ERP differences for 3DI vs. 2DNI environment during *Z* axis rotations is consistent with our behavioral results from Experiments 1 and 2. Specifically, the selection of frame of reference (viewer-centered vs. scene-based) primarily affects MR in horizontal depth (*Z* axis rotation).

DISCUSSION

The results of this study suggest that cognitive processing in a 3DI environment differs from that occurring in 2D and 3DNI environments. Furthermore, only immersive environments seem to encourage individuals to use egocentric spatial reference frames in visual encoding and transformation tasks.

In Experiment 1 relative depth information provided by the 3DNI environment was insufficient to encourage the use of a viewer-centered frame of reference. In 2DNI and 3DNI environments parity judgments were fastest when objects were rotated around the *Y* axis, while RTs for *X* and *Z* axis rotations were similar. This suggests that the objects' components were encoded in terms of "vertical" and "horizontal" relations with regard to the sides of the computer screen. In particular, rotation in horizontal depth plane might be relatively easy because it does not alter the orientation of the "sides" of the object with respect to the "left" and "right" sides of the computer screen (see also Parsons, 1987, for a similar discussion). This also suggests that cognition in the 2DNI and 3DNI environments is scene-based and might be atypical of human interactions with large-scale, real environments.

In contrast, our results for a 3DI environment were unique and demonstrated that viewers employed an egocentric frame of reference during MR. Specifically, the rate of rotation in horizontal depth (around *Z*) was significantly slower than that in the picture plane (around *X*). At the same time, the rate of rotation in the picture plane (around *X*) was faster in immersive environments compared to non-immersive environments, which is expected for rotation in a plane where no object components are occluded. Furthermore, when comparing between environments, rotation in the picture plane (around *X*) was the fastest in 3DI while rate of rotation in horizontal depth (around *Z*) was the slowest in 3DI. This suggests that subjects were in fact rotating a depth 2D retina-based object representations in 3DI environments and experiencing difficulties with foreshortening and occlusion.

Furthermore, the results of Experiment 2 demonstrated that the simple use of 3DI technology is not sufficient to promote the use of viewer-based frames of reference. When presentation is in 3DI and target objects are embedded within a "fixed frame," the observer relies on scene-based encoding. This is similar to the effect of viewing in 2DNI where the fixed borders are defined by the sides of a computer screen. In this case, scene-based encoding is most efficient because the observer is free to move around without need

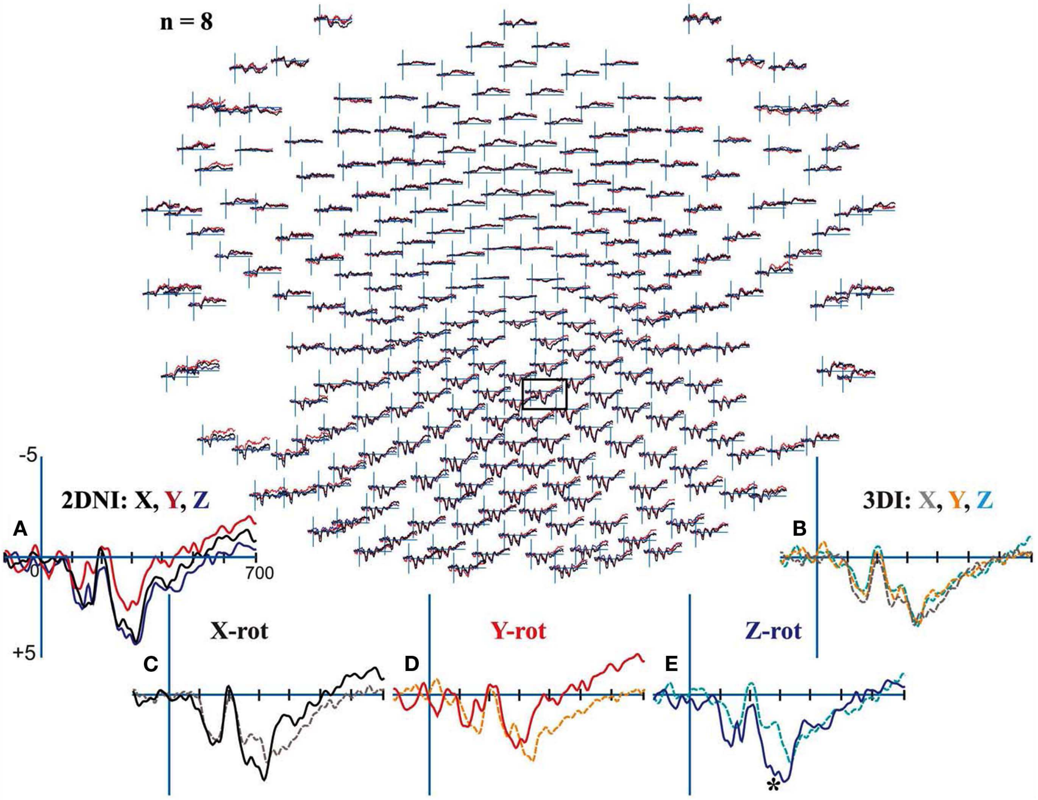


FIGURE 8 | Electroencephalograms for Mental Rotation under 2DNI and 3DI conditions: Evoked potential (μV) peaks were largest in centroparietal electrodes regardless of visual environment. (A) 2DNI ERPs were most negative for *Y* axis rotations from ∼250 ms onward. **(B)** 3DI ERP's demonstrated increasing negativities for *Z* and *Y* rotation with respect to *X* rotations at ∼350 ms. **(C–E)** When comparing rotations between 2DNI and 3DI environments 3DI rotations demonstrated a trend for greater negativity prior to 350 ms but slightly larger positivity at longer latencies >350 ms. The largest differences were between *Z*-rotations which were more positive for 3DI at ∼270–300 ms, (∗) with paired *t*-test $p < 0.03$.

to mentally update the position of every object within the scene. In contrast, within 3DI environments (without frame embedding) as well as in the large-scale real world, the positions of objects relative to an external frame may be constantly changing thus making viewer-centered encoding more efficient.

Finally, most previous EEG/ERP experiments on mental imagery have been performed only in 2DNI environments. The results of Experiment 3 demonstrate that 2DNI and 3DI environments do evoke differential parietal ERP responses, and that ERPs were more negative at ∼270–300 ms post-stimulus for MR in the 3DI vs. 2DNI environment. One interpretation is that this early modulation of ERP activity marks viewer-centered vs. scene-based orienting in preparation for subsequent MR from 400 ms onward. However, this early modulation may in addition involve other processes such as spatial attention and simple shape evaluation. Thus, future ERP research should evaluate the contribution of these component processes which may affect early stages of MR task performance in different visual environments. Also as Experiment 3 was a preliminary study, much work remains to be done to map the neural markers of brain processing differences between 2DNI vs. 3DI environments. This may include the use of structural MRI for anatomical localization of ERPs as well as the use of more ecologically valid task designs.

Importantly, our study is the first attempt to understand immersivity from a cognitive neuroscience perspective. Currently, there is no clear understanding of what "immersivity" means in cognitive terms. Most commonly used terms and definitions are merely descriptive, such as "perceiving oneself to be enveloped by, included in, and interacting with an environment" (Witmer and Singer, 1998, p. 227), and often confounded with such terms as "immersion" and "presence" describing the "extent to which the human operator loses his or her awareness of being present at the site and instead feels present in the artificial environment" (Durlach and Mavor, 1995, p. 22). Similarly, one problem in research aimed at understanding cognitive processing differences between virtual and real environments, is that they do not clearly define "immersivity." For example, Perani et al. (2001) reported that the right inferior parietal cortex (IPC) was activated when subjects observed real hand actions but not when observing hand movements in a 3D virtual environment or on a 2D display. Perani et al. suggested that within virtual environments, the right parietal system may not provide sufficient information for the computation of the spatial reference frame as it is used in the real world. However, what Perani et al. called an immersive environment was the condition in which participants laid in a PET scanner with their heads position fixed, and black curtains were arranged so that the participants were able to focus only on what was shown behind a single rectangular opening in front of them. We should note that this type of environment does not meet what we consider to be the essential requirement of "immersivity," namely viewing the scene from the "inside." We believe that an immersive 3D environment, in which a viewer is surrounded by the environment and no clear borders are present is necessary to encourage the use of an egocentric (viewer-centered) frame of reference.

Our findings have implications for future studies on spatial transformations of mental images and the design of testing environments. They show that the results of the previous experiments on MR, performed in laboratory conditions using a traditional 2D computer screen, might be limited in that they may not reflect the MR patterns that would be measured in a natural, 3D environment. In addition to its theoretical implications, this research could be of considerable interest from an applied perspective; specifically for the design of training and learning environments. Although 3D environments might be more attractive to the user, the results of the current research show that there will probably be no significant differences between encoding and spatial transformation of images under 2DNI and 3DNI conditions. On the other hand, a 3DI environment can provide a unique and possibly more realistic learning environment. In particular, a 3DI environment should provide advantage to those tasks that benefit from encoding from an egocentric frame of reference (e.g., navigation, wayfinding, laparoscopic surgery, and telerobotics). In general, using desktop graphics to train users for real world egocentric spatial tasks might not be effective, and may actually be counterproductive due to the differences in encoding and transformation processes in immersive vs. non-immersive environments. In fact, the findings of this research explain the results of previous studies that show no transfer from training in 2D non-environments to immersive virtual environments. For instance, Pausch et al. (1997) reported that immersive prior practice with conventional 2D displays in visual search tasks impaired performance in immersive virtual environments. The researchers suggested that using desktop graphics to train users for real world search tasks may not be efficient. The current study explains this finding by pointing out that the encoding of spatial relations and cognitive strategies applied to perform visual/spatial transformations in these two types of environments are different. We suggest that 3DI environments with a variety of simulated 3D stimuli will provide the most efficient environment for training egocentric visual-spatial skills that will generalize and transfer to real world tasks.

ACKNOWLEDGMENTS

This research was supported by the National University of Singapore (Singapore) and US Office of Naval Research (N00014-10-1-0638, Kozhevnikov) and the National Institutes of Health NCCAM (K01AT004481, Dhond). We thank Olesya Blazhenkova and David Schloerb for fruitful discussion. We thank Jodie Royan for her help in conducting Experiment 1, and Debbie Morton in conducting Experiment 3. We also thank (in alphabetical order): Andrey Gorbunov and Remi Tache for programming 3D virtual reality testing environments; Sheraz Khan and Emily Ruzich for assisting with EEG data analysis; as well as Klaus Gramann and Florian Friemel for EEG data analysis advice.

REFERENCES

Attneave, R., and Olson, R. K. (1967). Discriminability of stimuli varying in physical and retinal orientation. *J. Exp. Psychol.* 74, 149–157.

Chance, S., Gaunet, F., Beall, A., and Loomis, J. (1998). Locomotion mode affects the updating of objects encountered during travel: the contribution of vestibular and proprioceptive inputs to path integration. *Presence (Camb.)* 7, 168–178.

Cockayne, W., and Darken, R. (2004). "The application of human ability requirements to virtual environment interface design and evaluation," in *The Handbook of Task Analysis for Human-Computer Interaction*, eds D. Diaper and N. Stanton (Mahwah: Lawrence Erlbaum Associates), 401–421.

Corballis, M. C. (1988). Recognition of disoriented shapes. *Psychol. Rev.* 95, 115–123.

Corballis, M. C., Anuza, T., and Blake, L. (1978). Tachistoscopic perception under head tilt. *Percept. Psychophys.* 24, 274–284.

Corballis, M. C., Zbrodoff, J., and Roldan, C. E. (1976). What's up in mental rotation? *Percept. Psychophys.* 19, 525–530.

Delorme, A., and Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics. *J. Neurosci. Methods* 134, 9–21.

Desrocher, M. E., Smith, M. L., and Taylor, M. J. (1995). Stimulus and sex differences in performance of mental rotation: evidence from event-related potentials. *Brain Cogn.* 28, 14–38.

Durlach, N. I., and Mavor, A. S. (1995). *Virtual Reality: Scientific and Technological Challenges*. Washington, DC: National Academy Press.

Gauthier, I., Hayward, W. G., Tarr, M. J., Anderson, A. W., Skudlarski, P., and Gore, J. C. (2002). BOLD activity during mental rotation and viewpoint-dependent object recognition. *Neuron* 34, 161–171.

Hinton, G. E., and Parsons, L. M. (1988). Scene-based and viewer-centered representations for comparing shapes. *Cognition* 30, 1–35.

Klatzky, R., Loomis, J., Beall, A., Chance, S., and Golledge, R. (1998). Spatial updating of self-position and orientation during real, imagined, and virtual locomotion. *Psychol. Sci.* 9, 293–298.

Kozhevnikov, M., and Garcia, A. (2011). "Visual-spatial learning and training in collaborative design in virtual environments," in *Collaborative Design in Virtual Environments* (Intelligent Systems, Control and Automation: Science and Engineering) Vol. 48, eds X. Wang and J. Jen-Hung (Dordrecht: Springer), 17–26.

Loomis, J., Blascovich, J., and Beall, A. (1999). Immersive virtual environment technology as a basic research tool in psychology. *Behav. Res. Methods Instrum. Comput.* 31, 557–564.

Macuga, K. L., Beall, A. C., Kelly, J. W., Smith, R. S., and Loomis, J. M. (2007). Changing lanes: inertial cues and explicit path information facilitate steering performance when visual feedback is removed. *Exp. Brain Res.* 178, 141–150.

Palmer, S. E. (1989). "Reference frames in the perception of shape and orientation," in *Object Perception: Structure and Process*, eds B. Shepp and S. Ballesteros (Hillsdale, NJ: Lawrence Erlbaum Associates), 121–163.

Pani, J. R., and Dupree, D. (1994). Spatial reference systems in the comprehension of rotational motion. *Perception* 23, 929–946.

Parsons, L. M. (1987). Visual discrimination of abstract mirror-reflected three-dimensional objects at many orientations. *Percept. Psychophys.* 42, 49–59.

Parsons, L. M. (1995). Inability to reason about an object's orientation using an axis and angle of rotation. *J. Exp. Psychol. Hum. Percept.* 21, 1259–1277.

Pausch, R., Proffitt, D., and Williams, G. (1997). "Quantifying immersion in virtual reality," in *Proceedings of ACM SIGGRAPH '97 Conference on Computer Graphics*, Los Angeles.

Perani, D., Fazio, F., Borghese, N. A., Tettamanti, M., Ferrari, S., Decety, J., and Gilardi, M. C. (2001). Different brain correlates for watching real and virtual hand actions. *Neuroimage* 14, 749–758.

Pinker, S. (1988). "A computational theory of the mental imagery medium," in *Cognitive and Neuropsychological Approaches to Mental Imagery*, eds M. Denis, J. Engelkamp, and J. T. E. Richardson (Amsterdam: Martinus Nijhoff), 17–36.

Rock, I. (1986). "The description and analysis of object and event perception," in *Handbook of Perception and Human Performance*, Chap. 33, eds K. Boff, L. Kaufman, and J. Thomas (New York: John Wiley), 1–71.

Rock, I., and Leaman, R. (1963). An experimental analysis of visual symmetry. *Acta Psychol.* 21, 171–183.

Shepard, R. N., and Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science* 171, 701–703.

Tarr, M. J., and Warren, W. H. (2002). Virtual reality in behavioral neuroscience and beyond. *Nat. Neurosci.* 5(Suppl.), 1089–1092.

Witmer, B. G., and Singer, M. J. (1998). Measuring presence in virtual environments: a presence questionnaire. *Presence (Camb.)* 7, 225–240.

WorldViz. (2007). *Vizard VR Toolkit R3.* Available at: http://www.worldviz.com/products/vizard4/index.html

Conflict of Interest Statement: The research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. Furthermore, the content of this article is solely the responsibility of the authors and does not necessarily represent the official views of our funding agencies.

Received: 12 April 2012; accepted: 23 July 2012; published online: 10 August 2012.

Citation: Kozhevnikov M and Dhond RP (2012) Understanding immersivity: image generation and transformation processes in 3D immersive environments. Front. Psychology ***3****:284. doi: 10.3389/fpsyg.2012.00284*

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

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 28 August 2012
doi: 10.3389/fpsyg.2012.00315

Abacus in the brain: a longitudinal functional MRI study of a skilled abacus user with a right hemispheric lesion

Satoshi Tanaka[1]*, Keiko Seki[2], Takashi Hanakawa[3], Madoka Harada[4], Sho K. Sugawara[5,6], Norihiro Sadato[5,6], Katsumi Watanabe[7] and Manabu Honda[3]

[1] Center for Fostering Young and Innovative Researchers, Nagoya Institute of Technology, Nagoya, Japan
[2] Graduate School of Health Sciences, Kobe University, Kobe, Japan
[3] Department of Functional Brain Research, National Institute of Neuroscience, Kodaira, Japan
[4] Department of Rehabilitation Medicine, Eisei Hospital, Hachioji, Japan
[5] Division of Cerebral Integration, National Institute for Physiological Sciences, Okazaki, Japan
[6] School of Life Sciences, The Graduate University for Advanced Studies (SOKENDAI), Hayama, Japan
[7] Research Center for Advanced Technology and Science, The University of Tokyo, Komaba, Japan

Edited by:
Joel Pearson, The University of New South Wales, Australia

Reviewed by:
David Barner, University of California San Diego, USA
Feiyan Chen, Zhejiang University, China

***Correspondence:**
Satoshi Tanaka, Center for Fostering Young and Innovative Researchers, Nagoya Institute of Technology, Gokiso-cho Showa-Ku, Nagoya 466-8555, Japan.
e-mail: tanaka.satoshi@nitech.ac.jp

The abacus, a traditional physical calculation device, is still widely used in Asian countries. Previous behavioral work has shown that skilled abacus users perform rapid and precise mental arithmetic by manipulating a mental representation of an abacus, which is based on visual imagery. However, its neurophysiological basis remains unclear. Here, we report the case of a patient who was a good abacus user, but transiently lost her "mental abacus" and superior arithmetic performance after a stroke owing to a right hemispheric lesion including the dorsal premotor cortex (PMd) and inferior parietal lobule (IPL). Functional magnetic resonance imaging experiments were conducted 6 and 13 months after her stroke. In the mental calculation task, her brain activity was shifted from the language-related areas, including Broca's area and the left dorsolateral prefrontal and IPLs, to the visuospatial-related brain areas including the left superior parietal lobule (SPL), according to the recovery of her arithmetic abilities. In the digit memory task, activities in the bilateral SPL, and right visual association cortex were also observed after recovery. The shift of brain activities was consistent with her subjective report that she was able to shift the calculation strategy from linguistic to visuospatial as her mental abacus became stable again. In a behavioral experiment using an interference paradigm, a visual presentation of an abacus picture, but not a human face picture, interfered with the performance of her digit memory, confirming her use of the mental abacus after recovery. This is the first case report on the impairment of the mental abacus by a brain lesion and on recovery-related brain activity. We named this rare case "abacus-based acalculia." Together with previous neuroimaging studies, the present result suggests an important role for the PMd and parietal cortex in the superior arithmetic ability of abacus users.

Keywords: acalculia, arithmetic, calculation, expertise, imagery, memory, plasticity, stroke

INTRODUCTION

To perform complex calculations, most people rely on physical devices such as pencil and paper, mechanical calculators, and more recently digital computers. One such device is an abacus, which is still widely used in Asian countries. The abacus is a simple device of beads and rods, and numbers are represented by the spatial locations of beads (**Figure 1**). Skilled abacus users can calculate accurate answers to mathematical problems extremely rapidly. Interestingly, however, abacus users not only manipulate the tool skillfully in its physical form but also gain the ability to mentally calculate extraordinarily large numbers, often more than 10 digits at the expert level, with unusual speed and accuracy (Hatano et al., 1977). Psychological studies have shown that a non-linguistic strategy using visual imagery of the abacus (a "mental abacus") underlies this unusual calculation ability (Hatano et al., 1977, 1987; Hatano and Osawa, 1983; Stigler, 1984; Hatta et al., 1989; Hishitani, 1990; Hanakawa et al., 2004; Tanaka et al., 2008; Frank and Barner, 2012). These works have demonstrated examples of the role of mental imagery in mental arithmetic operations.

Several behavioral and neuroimaging studies have attempted to examine the neural correlates of the calculation strategy employed by abacus users (Hatta and Ikeda, 1988; Tanaka et al., 2002, 2008; Hanakawa et al., 2003; Chen et al., 2006; Wu et al., 2009; Hu et al., 2011; Ku et al., 2012). For example, recent neuroimaging studies have reported activation in the bilateral dorsal premotor cortex (PMd) and inferior and superior parietal lobule (IPL and SPL, respectively) during mental calculation and digit memory tasks in abacus users (Tanaka et al., 2002; Hanakawa et al., 2003; Chen et al., 2006; Wu et al., 2009; Ku et al., 2012). However, there have been no neuropsychological studies that report deficits in mental abacus ability after focal brain injury. Therefore, the causal relationship between mental abacus ability and region-specific brain structures remains unclear.

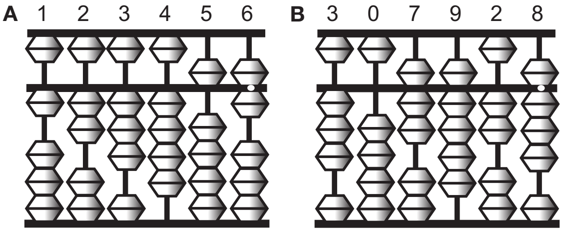


FIGURE 1 | Illustrations of an abacus. An abacus is a rectangular wooden calculator based on the decimal system. Each vertical rod has five sliding beads, one above and four below a middle horizontal bar. Numbers are represented by the configurations of the beads. A bead above the bar is equal to five when it is pushed down, and each of the four beads below is equal to one when pushed up. For example, the left figure **(A)** represents 123,456 and the right figure **(B)** represents 307,928.

Here, we report the case of a patient who was a well-experienced abacus user but had impaired mental arithmetic performance based on her mental abacus strategy due to a stroke. Her knowledge of basic arithmetic facts and her knowledge and operation of a physical abacus were intact. Only performance in mental calculation and digit memory tasks based on the mental abacus strategy was transiently impaired after the lesion. When we met her for the first time, she said "I lost my abacus in the brain."

The first purpose of the present study was to localize the lesion areas using high-resolution structural magnetic resonance imaging (MRI) with a 3T MRI scanner. We hypothesized that the lesion areas should include the PMd and/or parietal regions that were dominantly activated during the mental calculation and digit memory tasks in the previous functional MRI studies of abacus users (Tanaka et al., 2002; Hanakawa et al., 2003; Chen et al., 2006; Ku et al., 2012).

The second purpose of the present study was to examine the changes of brain activity with the recovery of mental abacus ability. Several neuroimaging studies have reported changes of brain activities with recovery from motor, attentional, or language deficits after stroke (Ward et al., 2003; Fridman et al., 2004; Corbetta et al., 2005; Price and Crinion, 2005; Heiss and Thiel, 2006). However, recovery-related changes in brain activity from deficits in arithmetic ability, especially in the non-linguistic aspects of arithmetic operation, remain totally unknown.

We hypothesized that the patient would change her strategy for mental calculation and digit memory from verbal to visuospatial with stroke recovery. Therefore, her brain activity during mental calculation would shift from language-related to visuospatial-related brain regions after recovery. As mentioned above, previous imaging studies have revealed dominant activation in the bilateral PMd, IPL, and SPL during mental calculation in abacus experts (Tanaka et al., 2002; Hanakawa et al., 2003). Neuroanatomical studies have shown that the PMd and parietal cortex have dense neuroanatomical connections (Wise et al., 1997; Luppino et al., 1999; Wise and Murray, 2000). Thus, the PMd, IPL, and SPL may work as a functional network during abacus-based mental calculation. Damage in one node may induce transient impairment of mental abacus ability. However, it is possible that the other intact nodes in the functional network could gain the ability to work without the damaged node, possibly because of functional reorganization within the remote intact nodes (Frost et al., 2003; Fridman et al., 2004; Dancause et al., 2006). Thus, we hypothesized that the intact PMd, IPL, and/or SPL would be active with the recovery of mental abacus ability.

In the present study, functional MRI experiments were conducted 6 and 13 months after her stroke and brain activity between the two sessions was compared in order to test this hypothesis. In addition, a behavioral experiment using dual-task interference paradigms was conducted to confirm her use of the mental imagery of an abacus on a digit memory task 13 months after her stroke.

MATERIALS AND METHODS

CASE REPORT

The patient was a 57-year old left handed female. She had worked as a professor in a national university before the stroke. She had a Ph.D. degree in medicine and had worked as a scientist in the field of neuropsychology for more than 25 years. She had published more than 20 international peer-reviewed papers. She had also engaged in rehabilitative medicine as a speech-language-hearing therapist for more than 25 years.

She started her abacus training at an abacus school when she was an elementary-school child, and had trained in physical and mental abacus operation for 3 years. We speculated that she was an excellent and skilled abacus user owing to the fact that she became a finalist at a domestic abacus competition in Japan in two successive years, although her training period was relatively shorter compared with the grand experts who participated in our previous functional MRI studies (Tanaka et al., 2002; Hanakawa et al., 2003). After she finished her abacus training, she kept using abacus-based mental calculation and mnemonic strategies in everyday activities for a long period and did not lose her ability. In fact, she reported that her forward digit span was around 12 before the stroke episode. This was far beyond the average score for her age group.

In July 2009, she suffered from a right hemispheric infarct in the territory of the anterior and middle cerebral arteries. When a therapist tested her digit span during a clinical neuropsychological evaluation in a hospital approximately 2 months after her stroke, she noticed that she was not able to use the mental abacus strategy for the digit span test. She was not able to generate vivid mental imagery of an abacus and the image of the abacus was very fragile. Detailed structural MRI scans were obtained in January 2010. Functional MRI scans were conducted at two different periods, the first in January 2010 and the second in August 2010.

NEUROPSYCHOLOGICAL EVALUATION

Neuropsychological evaluations were conducted approximately 1 month after stroke onset. Her score on Raven's Standard Progressive Matrices was in the average range (33/36). Similarly, her IQ measured by Kohs Block Design Test was also in the average range (108). The Standard Language Test for Aphasia (SLTA; Hasegawa et al., 1984), which has been widely used in Japan, did not detect any impairments of language. However, clinical observation detected mild impairments of her speech production: her prosody was impaired and speed of speech was slow with small volume. Clinical observation immediately after her stroke detected unilateral visual neglect. For motor function, the patient showed a severe paralysis in the left upper limb and mild paralysis in the left lower limb.

ARITHMETIC ABILITY

After her stroke onset, her arithmetic ability was not impaired according to the neuropsychological evaluation. She was able to perform four basic arithmetic operations without any problem. In fact, she was able to answer all arithmetic problems correctly in the SLTA. In addition, her long-term memory of digits was also intact because she correctly remembered the numbers of her bank accounts and airplane mileage accounts. However, she noticed that she was not able to generate visual imagery of a mental abacus, which had been easily generated before the stroke, when a neuropsychologist tested her maximum digit span 2 months after her stroke. Before the stroke, she used to use the mental abacus strategy especially when she calculated and memorized larger sequences of digits, because the visuospatial strategy, rather than a phonological strategy, was useful in coding a larger number of digits (Hatano et al., 1977; Hatano and Osawa, 1983). Due to the impairment of visual imagery after her stroke, she used the phonological strategy instead. She was able to perform four basic arithmetic operations correctly although she felt that her arithmetic ability had declined after her stroke.

Six months after her stroke, just before the first functional MRI session, we evaluated her knowledge of basic arithmetic facts, as well as her knowledge, and operation of a physical abacus. These aspects were all intact. However, she still felt that it was difficult to generate a vivid visual image of a mental abacus. She reported that she was not able to perform mental calculations and memorize digit sequences based on the mental abacus strategy because her mental abacus was fragile. However, 13 months after her stroke, she reported that her capacity for visual imagery of a mental abacus had recovered. At that time, she participated in the second functional MRI session.

Figure 2 shows her behavioral performance of maximum digit and alphabet span tasks. Forward digit span and forward and backward alphabet spans were all unchanged across the experimental period. In contrast, backward digit span improved over time after her stroke. Her backward digit span 13 months after her stroke was eight and almost equal to her forward digit span. It has been reported that abacus experts reproduce a series of digits in backward order almost as well as in the forward order, because both require experts to read off the digits from visuospatial mental representation of an abacus (Hatano and Osawa, 1983). Therefore, nearly identical maximum digit spans both backward and forward might be interpreted as evidence that she used her mental abacus 13 months after her stroke. In fact, she reported that she was able to use the mental abacus strategy for the backward digit span task 13 months after her stroke.

EXPERIMENTAL PROCEDURE

The patient gave written, informed consent before the experiments, which were approved by the local ethics committee of the National Institute for Neuroscience.

The patient participated in two functional MRI sessions of the mental calculation and digit memory tasks (Experiment 1). The first and second functional MRI sessions were conducted 6 months (January 2010) and 13 months (August 2010) after her stroke onset, respectively. The difference of the brain activities between the two sessions was compared. Structural MRI scans were obtained in January 2010. In addition, the patient participated in a behavioral experiment after the second functional MRI session in order to examine whether the patient would use abacus-based mental calculation and digit memory strategies in these tasks (Experiment 2).

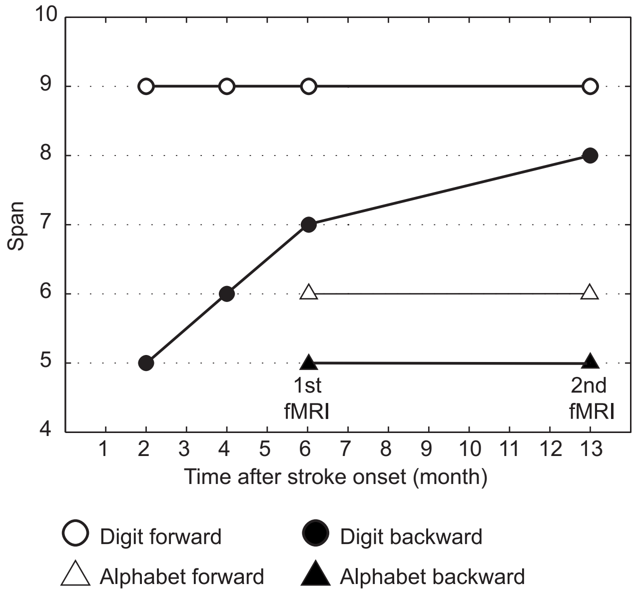


FIGURE 2 | Behavioral performance of maximum digit and alphabet span tasks. Maximum digit spans were higher compared with alphabet span, indicating that her superior performance was digit memory-specific. Her maximum forward digit span (white circle), as well as forward and backward alphabet spans (white and black triangles), was unchanged across the entire experimental period. However, the maximum backward digit span (black circle) was improved over time after her stroke.

EXPERIMENT 1

Behavioral task in functional MRI experiment

For the functional MRI experiment, the patient performed mental calculation and digit memory tasks that were used in our previous functional MRI studies of abacus experts (Tanaka et al., 2002; Hanakawa et al., 2003). Before the functional MRI experiment, she practiced these tasks outside the scanner to become familiar with the tasks. Presentation software (Neurobehavioral Systems Inc., Albany, CA, USA) was used for the visual stimulus presentation and to record her responses. Stimuli were presented on a screen using a liquid crystal display projector, and she viewed the screen though a mirror.

For the mental calculation task, white digit stimuli were presented for 1.5 s with inter-stimulus intervals of 2 s on the center of a screen (**Figure 3A**, Hanakawa et al., 2003). Digit stimuli were presented 10 times during each trial. The patient was asked to mentally add the presented series of digits without moving her fingers. After the presentation of these digit stimuli, a red digit stimulus

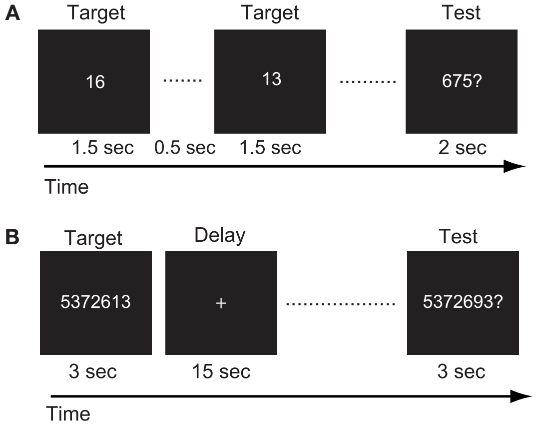


FIGURE 3 | (A) Schematic illustration of the mental calculation task. The patient was asked to add a series of numbers mentally that were visually presented on the computer screen. **(B)** Schematic illustration of the digit memory task. She was asked to retain the target sequence of digits during the delay period.

was presented for 3 s. She was asked to judge whether the addition answer in her mind and the test digit stimuli were the same or different, by pressing one of the response buttons with the right fingers. After each trial, there was an 18-s inter-trial interval (ITI) in which the patient simply watched the white fixation cross presented at the center of the screen (visual fixation condition). She performed additional tasks with single-digit and two-digit numbers. The experimental session consisted of five trials for each task in an alternate order.

For the digit memory task, a delayed match-to-sample task using a digit sequence as the stimulus was employed (**Figure 3B**, Tanaka et al., 2002). A target digit sequence was presented on the center of a screen for 3 s. The length of the digit sequence was a five digit number, which was two digits shorter than her digit span memory capacity measured before the first functional MRI session. After a 15-s delay period, during which only a fixation cross appeared on the screen, a test sequence of digits was presented for 3 s. She was asked to judge whether the target and test sequences were the same or different, by pressing one of the response buttons. Following these behavioral events, there was a 17-s visual fixation. The experimental session consisted of 10 trials.

The patient also participated in functional MRI experiments of verbal fluency and hand grip tasks 13 months after the stroke. These experiments were conducted to ascertain whether the region-specific brain activity during arithmetic tasks 13 months after the stroke would be task-specific or not. In the verbal fluency task, the subject was asked to generate in her mind as many words as possible from an indicated category (such as names of sports or fruits) during a 24-s trial. After each trial, there was a 24-s visual fixation condition. The task and fixation condition were alternately performed 10 times. In the hand grip task, the patient was asked to make the hand grip movement with her paretic hand every 2 s during a 24-s period. The hand grip task and visual fixation condition were alternately performed 10 times.

Imaging data acquisition and analysis

The functional MRI experiment was conducted using a 3.0-T MRI scanner (MAGNETOM Trio, Siemens, Erlangen, Germany). Functional images were acquired using a T2*-weighted echo planar imaging sequence (TR/TE/FA/FOV/voxel size/slice number = 3000 ms/30 ms/90˚/192 mm/3.0 mm × 3.0 mm × 3.0 mm/46 axial slices for the mental calculation task, and 2000 ms/40 ms/80˚/192 mm/3.0 mm × 3.0 mm × 4.0 mm/25 axial for the digit memory task). A total of 143 and 205 functional images on each mental calculation and digit memory task were collected during each session. The first three and five images of each task were discarded from data analysis to allow for the stabilization of the magnetization. Eighty-three images were obtained on each verbal fluency and hand grip task and the first three images were discarded. A high-resolution structural T1 image was acquired using a Magnetization Prepared Rapid Acquisition in Gradient Echo (MPRAGE) sequence.

SPM8 software (Wellcome Department of Cognitive Neurology, London, UK) was used for image processing and analysis. The T1 image was spatially normalized to fit a Montreal Neurological Institute (MNI) template (Evans et al., 1993). The damaged regions were masked to reduce the influence from non-brain or lesioned tissue (Brett et al., 2001). For functional images, the data were first realigned to the mean functional images in order to reduce the effect of head motion. These images were then normalized to the MNI template, with the same parameter obtained for T1 normalization. Then, the images were spatially smoothed using an isotropic Gaussian kernel of 6-mm full-width half maximum (FWHM).

Statistical analysis

Statistical analysis of the time course data at each voxel was conducted with a general linear model in order to identify voxels that showed task-specific and session-specific signal changes (Friston et al., 1994). The brain activities in the mental calculation and digit memory tasks were analyzed separately.

For the mental calculation task, one-digit and two-digit calculation tasks were separately modeled as regressors on each session with boxcar functions convolved with a hemodynamic response function. For the digit memory task, the presentations of the target and test sequences, and the delay period, were separately modeled on each session using three boxcar functions convolved with a hemodynamic response function. For the verbal fluency and hand grip task, the task period was modeled using three boxcar functions convolved with a hemodynamic response function. In all tasks, head-movement parameters were also included as regressors of no interest.

To test hypotheses about regionally specific task-effects or session-effects, the estimates for each model parameter were compared with the linear contrasts. The resulting set of voxel values constituted a statistical parametric map of the t statistic, SPM$\{t\}$. In all tasks, the statistical threshold was set at $p < 0.001$ at the voxel level. Control for multiple comparisons was achieved at the cluster level with Gaussian random field theory either in the whole brain (p corr < 0.05) or the small volume around the coordinates of the regions of interest (ROIs) based on the published papers (p svc < 0.05). On the basis of previous works on abacus experts (Tanaka et al., 2002; Hanakawa et al., 2003), spherical ROIs ($r = 8$ mm) were created at the peak voxel in the bilateral SPL (left $x = -18$, $y = -66$, $z = 60$; right $x = 14$, $y = -66$, $z = 64$ at MNI coordinate), left IPL ($x = -46$, $y = -40$, $z = 54$), left PMd ($x = -32$, $y = -6$, $z = 52$), and Broca's area ($x = -50$, $y = 10$, $z = 26$).

EXPERIMENT 2

Behavioral evaluation in mental abacus use

A behavioral experiment using interference paradigms was conducted to examine whether the patient would utilize the mental abacus strategy on a digit memory task 13 months after her stroke (**Figure 10A**). The behavioral paradigm was based on Hatta et al. (1989). She performed a delayed digit recall task. First, a target digit sequence was presented on the computer screen for 3 s. The length of the target digit sequence was eight, which was one-digit shorter than her maximum digit span memory capacity. After a 15-s retention interval, she was asked to recall and report the digit sequence orally. There were three experimental conditions which differed according to the types of visual distractors. Pictures of abacus figures, human faces, or gray rectangles were presented on the center of the screen during the retention interval. Each distractor stimulus was presented for 1 s with 0.5 s inter-stimulus intervals. She performed 15 trials for each distractor condition. We hypothesized that if she utilized a mental abacus for the digit memory task, the presentation of the pictures of abacus figures would interfere with task performance more than the presentation of the human faces and gray rectangles.

RESULTS

STRUCTURAL MRI

The T1-weighted MRI showed a right fronto-parietal lesion, involving the posterior parts of the inferior and superior frontal gyrus, anterior insula, anterior cingulate gyrus, pre and post central gyrus, and supramarginal gyrus (**Figure 4**). These lesioned areas included the right PMd and IPL, which were dominantly activated during the mental calculation and digit memory tasks in the previous functional MRI studies of abacus experts (Tanaka et al., 2002; Hanakawa et al., 2003; Chen et al., 2006; Ku et al., 2012). The lesion was not observed in the left hemisphere.

EXPERIMENT 1: FUNCTIONAL MRI EXPERIMENT

Mental calculation task

The patient responded correctly in all trials of the calculation tasks in both functional MRI sessions. **Figure 5A** shows brain activity associated with one- and two-digit mental calculation tasks relative to the visual fixation condition (see **Table A1** in Appendix online). In one-digit mental calculations, brain activity was generally lateralized to the left hemisphere both 6 and 13 months after her stroke. In contrast, brain activity in two-digit mental calculations was observed bilaterally both 6 and 13 months after her stroke. These brain regions include the middle frontal gyrus, pre- and postcentral gyrus, SPL, middle and superior occipital gyrus, inferior temporal gyrus, and cerebellum. This activity was not observed in the damaged regions of the right hemisphere. When brain activities during one- and two-digit mental calculation tasks were directly compared, significant brain activity in the left middle frontal gyrus was observed 6 months after her stroke (**Figure 5B**). In contrast, significant activity was observed in the bilateral SPL, right middle frontal gyrus, postcentral gyrus, and middle occipital gyrus 13 months after her stroke.

To investigate the time-specific brain activities, her whole brain activities between 6 and 13 months after her stroke were directly compared. A previous study has revealed that the region-specific brain activities in abacus users were more evident in the mental calculation task with a higher cognitive demand (Hanakawa et al., 2003). Therefore, the brain activities in the two-digit addition task between 6 and 13 months after her stroke were compared in the analysis.

The results are shown in **Figure 6**. The left hemispheric cortical activities including Broca's area (peak coordinate $x = -48$, $y = 8$, $z = 8$; $t = 4.73$, cluster size = 227 voxels, p corr < 0.05), the left dorsolateral prefrontal cortex (DLPFC, $x = -48$, $y = 38$, $z = 30$; $t = 4.81$, cluster size = 118 voxels, p corr < 0.05), and IPL ($x = -44$, $y = -50$, $z = 54$; $t = 4.38$, cluster size = 118 voxels, p corr < 0.05) were significantly greater at 6 months compared with 13 months after the stoke (**Figure 6A**). These brain regions were repeatedly activated in many language- related cognitive tasks (Paulesu et al., 1993; Fiez et al., 1996; Smith et al., 1998). In contrast, activity in the left SPL ($x = -20$, $y = -66$, $z = 66$; $t = 3.60$, cluster size = 10, p svc < 0.05) was significantly greater at 13 months compared with 6 months after her stroke

FIGURE 4 | T1-weighted structural MRI of the patient. The lesion was observed in the fronto-parietal cortex, including the posterior parts of the inferior and superior frontal gyrus, anterior insula, anterior cingulate gyrus, pre and post central gyrus, and supramarginal gyrus. No lesion was observed in the left hemisphere.

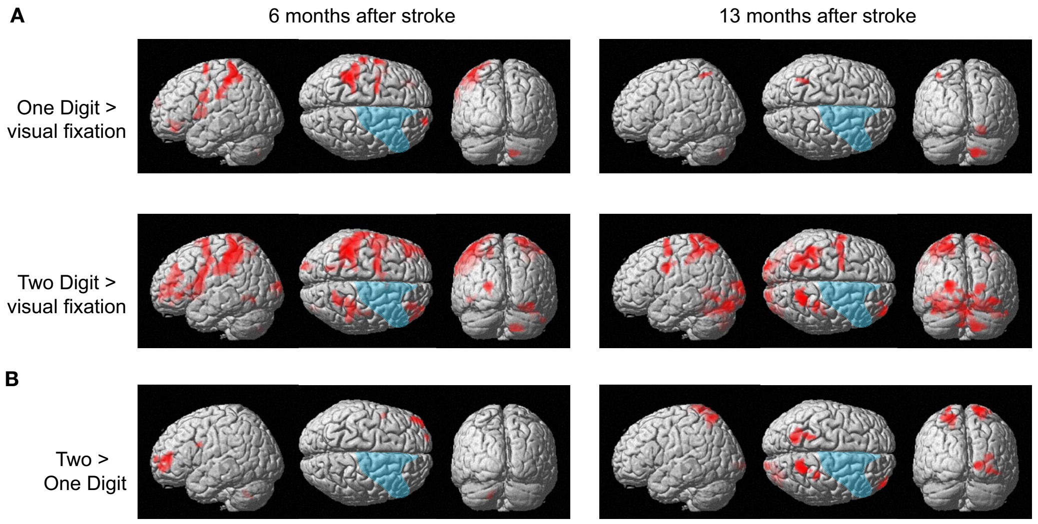


FIGURE 5 | (A) Active areas during the mental calculation task relative to those during the visual fixation condition 6 and 13 months after her stroke. **(B)** Active areas during the two-digit mental calculation task relative to those during the one-digit task 6 and 13 months after her stroke. The blue region indicates the lesion area. Brain activity is overlayed on the spatially normalized rendering images (voxel level, $p < 0.001$ uncorrected for multiple comparison; cluster level $p < 0.05$ FWE). The detailed coordinates and statistical values are listed in **Table A1** in Appendix online.

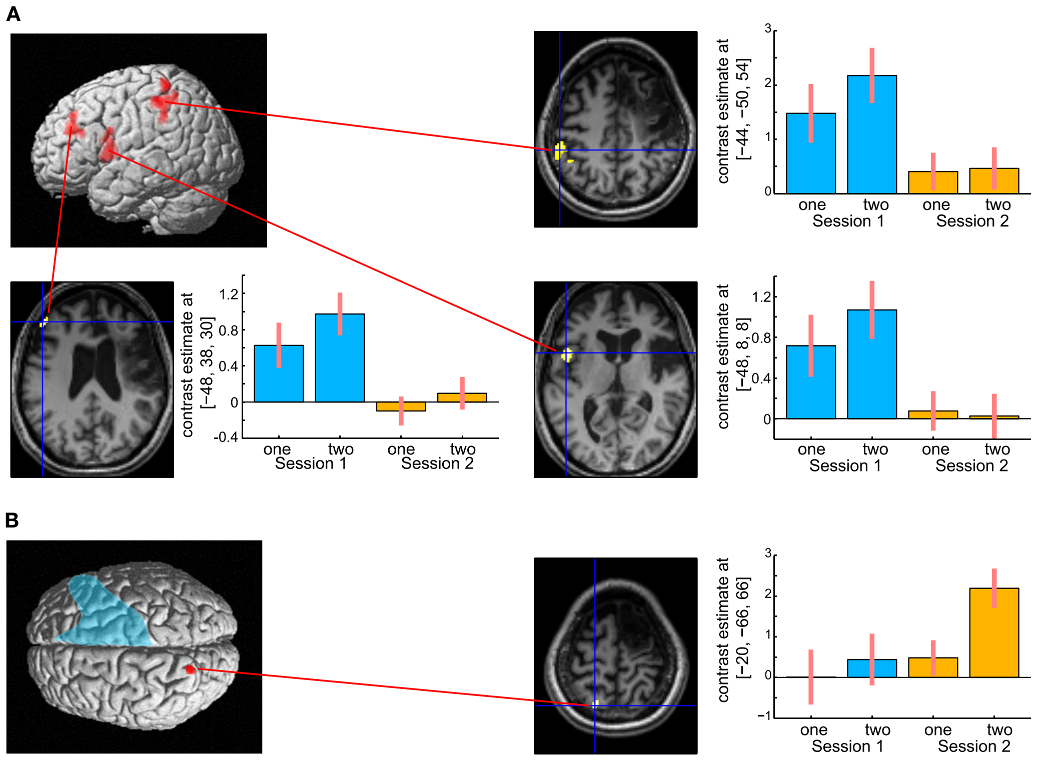


FIGURE 6 | (A) Active areas during the mental calculation task-specific to 6 months after her stroke. These areas include Broca's area, the left dorsolateral prefrontal cortex and the inferior parietal cortex. **(B)** Active areas specific to 13 months after her stroke. Only the left superior parietal cortex was significantly activated on this comparison. The blue color on the rendering structural image indicates the damaged cortical area. Only statistically significantly different activity is shown (voxel level, $p < 0.001$ uncorrected for multiple comparison; cluster level $p < 0.05$ FWE). In each bar graph, the vertical axis indicates contrast estimated values relative to the baseline, whereas the horizontal axis indicates the type of task. One and two in the figure indicates one-digit and two-digit addition task, respectively. Blue and orange color bar indicates session 1 and 2, respectively. The error bars represent standard deviations across scans.

(**Figure 6B**). Activity in the left SPL was observed in the previous functional imaging studies of mental calculation tasks in abacus experts (Hanakawa et al., 2003; Chen et al., 2006; Wu et al., 2009). These functional MRI results were very consistent with the patient's subjective report that she was able to shift the calculation strategy from a phonological – based to a mental abacus – based strategy according to her level of recovery from the stroke.

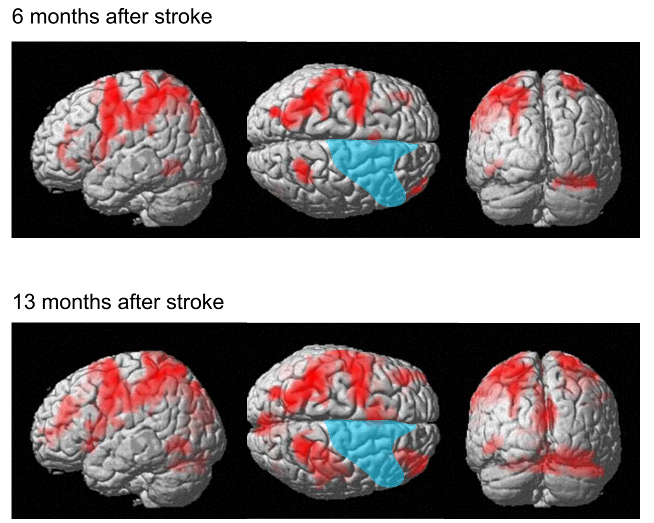


FIGURE 7 | Active areas during the delay period on the digit memory task relative to these during the visual fixation condition 6 and 13 months after her stroke. The blue region indicates the lesion area. Brain activity is overlayed on the spatially normalized rendering images (voxel level, $p < 0.001$ uncorrected for multiple comparison; cluster level $p < 0.05$ FWE). The detailed coordinates and statistical values are listed in **Table A2** in Appendix online.

Digit memory task

The patient correctly answered all trials in both functional MRI sessions. The present analysis of the digit memory task focuses on the brain activities associated with memory retention and thus the brain activities only during the delay period are reported. **Figure 7** shows brain activity associated with the delay interval period during the digit memory tasks relative to the visual fixation condition (see **Table A2** in Appendix online). Overall, brain activity was left lateralized 6 months after her stroke, whereas bilateral activation was observed 13 months after her stroke. These brain regions include the inferior and middle frontal gyrus, insula, supplementary motor area, IPL, SPL, cuneus, fusiform gyrus, inferior temporal gyrus, and cerebellum.

A direct comparison of the brain activities observed during the delay period between the two sessions is shown in **Figure 8**. No brain regions were observed that showed significant regional-specific activities at 6 months compared with those at 13 months after her stroke. In contrast, activities in the bilateral SPL (left $x = -18$, $y = -64$, $z = 64$; $t = 6.36$, cluster size = 223, p corr < 0.05; right $x = 20$, $y = -48$, $z = 70$; $t = 4.93$, cluster size = 132, p corr < 0.05) and the right visual association cortex ($x = 36$, $y = -80$, $z = -14$; $t = 6.77$, cluster size = 529, p corr < 0.05) were significantly greater at 13 months compared with 6 months after her stroke. The bilateral activities in the SPL during the delay period were observed in the previous functional MRI study of abacus experts (Tanaka et al., 2002). Thus, the result suggests that the visuospatial strategy of mental abacus representation might be more dominantly used in the digit memory task,

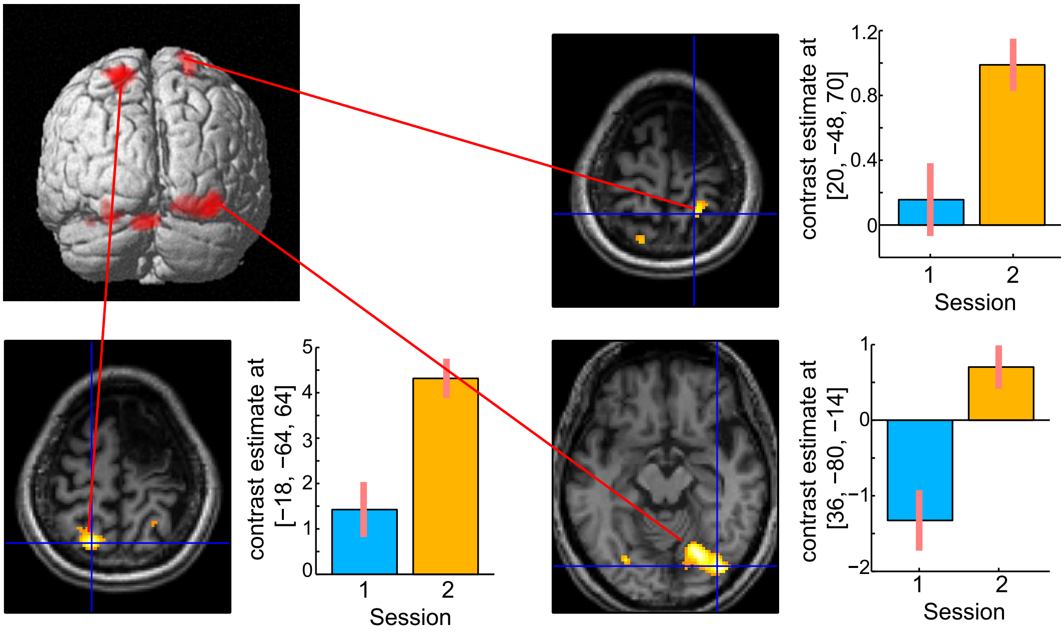


FIGURE 8 | Active areas during the retention period in the digit memory task specific to 13 months after her stroke. The bilateral superior parietal lobule and visual association cortex were significantly activated. Statistical threshold was the same as with the mental calculation task.

the same as in the mental calculation task, at 13 months after her stroke. Again, this was consistent with the patient's subjective report that she was able to utilize the mental abacus strategy 13 months after her stroke.

Verbal fluency and hand grip tasks

Figure 9 shows the results of verbal fluency and hand grip tasks. There was significant task-specific activity mainly in the left DLPFC for the verbal fluency task and in the right primary motor cortex for the left hand grip task, respectively. In contrast, in both tasks, the left SPL, which was dominantly activated during her mental calculation and digit memory tasks, was not significantly activated compared with the visual fixation condition. These findings suggest that activation in the SPL was specific to mental calculation and digit memory tasks 13 months after the stroke.

EXPERIMENT 2

Behavioral experiment

The number of correctly answered trials was 12 for the human face and gray rectangle conditions, compared with 6 for the abacus picture condition (**Figure 10B**). Therefore, the number of the correct trials in the abacus picture condition was clearly fewer than that in the other two distractor conditions. This result showed that the presentation of pictures of abacus figures interfered with the patient's task performance, suggesting her use of a mental abacus on the digit memory task and mental calculations 13 months after her stroke.

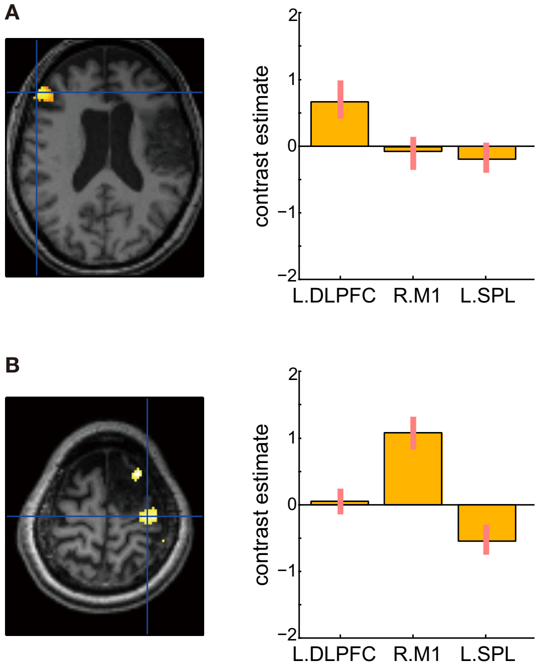


FIGURE 9 | Active areas during verbal fluency and left hand grip tasks 13 months after the stroke. (A) There were significant task-specific activities in the left dorsolateral prefrontal cortex (DLPFC) for verbal fluency task and **(B)** right primary motor cortex (M1) for left hand grip task. The left SPL was not significantly activated. In each bar graph, the vertical axis indicates contrast estimated values relative to the visual fixation task, whereas the horizontal axis indicates each brain region. The error bars represent standard deviations across scans.

DISCUSSION

This is the first case report on the impairment of mental abacus ability by a brain lesion and on recovery-related brain activity. The patient's knowledge and operation of basic arithmetic facts and of a physical abacus were all intact. Her impairment of arithmetic ability was specific to mental calculation and digit memory only based on the mental abacus strategy. Therefore, we consider that this would be a specific case of spatial acalculia (Hécaen et al., 1961; Hartje, 1987; Granà et al., 2006). This is a quite rare case and we have named this "abacus-based acalculia."

The results of the present study show that brain activity during mental calculation at 13 months after her stroke was observed more in an area implicated in visuospatial working memory (Jonides et al., 1993; Mellet et al., 1996; Courtney et al., 1998a,b; Rowe et al., 2001; Tanaka et al., 2005; Oshio et al., 2010), whereas at 6 months after her stroke, brain activity was more predominant in the left hemisphere in areas related to verbal working memory (Paulesu et al., 1993; Fiez et al., 1996; Smith et al., 1998). Brain activity at 13 months after her stroke was observed in the left SPL, whereas that at 6 months after her stroke was observed in Broca's area and the left DLPFC and IPL. This shift of region-specific brain activities is consistent with her subjective report that she was able to shift her calculation strategy from a verbal to a visuospatial strategy according to the level of her recovery from the stroke. In a behavioral experiment using interference paradigms, a visual presentation of an abacus picture, but not a human face picture, interfered with her performance of digit memory, confirming her use of the mental abacus 13 months after her stroke.

The present result is consistent with previous functional imaging studies that reported activation in the SPL during mental calculation and digit memory tasks in abacus users (Tanaka et al., 2002; Hanakawa et al., 2003; Chen et al., 2006; Wu et al., 2009). It is possible that a spatial representation of numbers is developed through abacus practice, which involves rule based visuo motor processing, and utilized in mental calculation and digit memory tasks, because it is more efficient to mentally manipulate large numbers using a spatial representation than a sequentially organized phonological representation (Hatano et al., 1977; Hatano and Osawa, 1983 Hatano et al., 1987; Hatta et al., 1989; Hishitani, 1990; Tanaka et al., 2008; Frank and Barner, 2012). The SPL might be a key brain region for such non-verbal visuospatial representation of numbers.

According to the structural MRI, the lesion area involved the right fronto-parietal regions. Her impairment of mental abacus ability due to her right hemispheric lesion was consistent with previous behavioral and neuroimaging studies that indicate involvement of the right hemisphere in the superior arithmetic abilities of abacus users (Hatta and Ikeda, 1988; Tanaka et al., 2002; Hanakawa et al., 2003; Chen et al., 2006; Wu et al., 2009). More specifically, her lesion area included the right PMd and IPL, which have been repeatedly activated in the previous functional neuroimaging studies of abacus users (Tanaka et al., 2002; Hanakawa

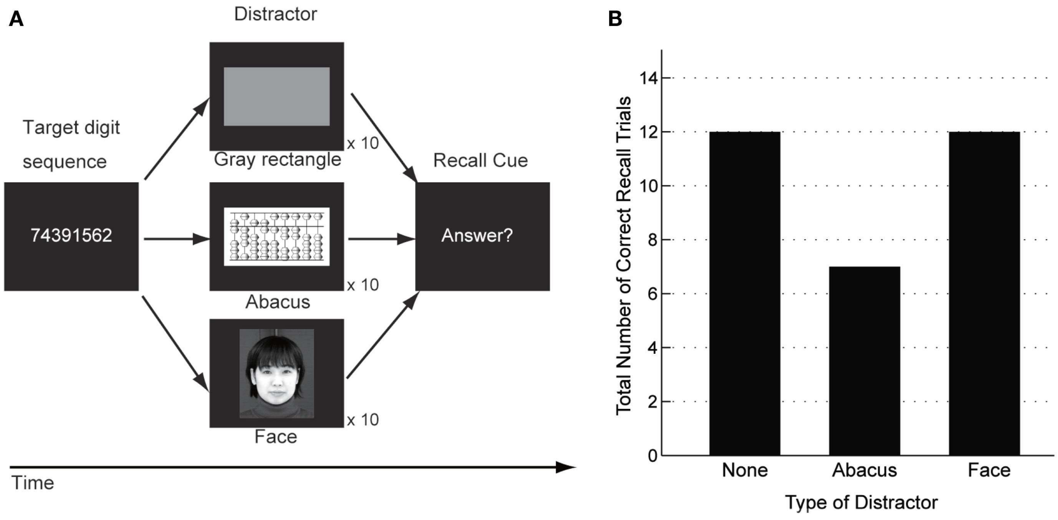


FIGURE 10 | (A) A delayed digit recall task using interference paradigms, based on Hatta et al. (1989). The patient was asked to recall a target digit sequence after a 15-s retention interval. Three different types of visual distractors were presented during the retention interval (pictures of abacus figures, human faces, or gray rectangles). **(B)** Behavioral performance in the delayed digit recall task. The number of correctly answered trials was 12 when a human face or a gray rectangle were presented as distractors. In contrast, the number of correctly answered trials was six when the distractor was pictures of abacus figures. This result indicated that the presentation of abacus figures interfered with the patient's digit memory performance.

et al., 2003; Chen et al., 2006; Wu et al., 2009). Therefore, the present study may suggest the functional relevance of these brain regions to the mental calculation and digit memory of abacus users. However, we should be careful about such interpretations because the lesion area not only covered the PMd and IPL but also included relatively large areas of the right frontal and parietal cortex. A non-invasive brain stimulation study or neuropsychological study of patients with a more focal brain lesion will clarify this issue.

The activation in the SPL was less evident at 6 months compared with 13 months after her stroke. This implies that the damaged regions in the right hemisphere, possibly the PMd and IPL, and the SPL may work as a functional network during abacus-based mental calculation and digit memory. In fact, it is known that there is an anatomical and functional connectivity between the premotor and parietal cortex (Wise et al., 1997; Luppino et al., 1999; Wise and Murray, 2000; Tanaka et al., 2005; Oshio et al., 2010). Damage in one cortical node may induce less activity in another cortical node within the functional network. However, 13 months after her stroke, the SPL might be able to work without the damaged brain regions, possibly because of remote cortical reorganization that may occur within the intact SPL region (Frost et al., 2003; Fridman et al., 2004; Dancause et al., 2006).

In the present study, the significant activity in the SPL was left lateralized in the mental calculation task, whereas bilateral activation was found in the digit memory task. This might be due to differences in task difficulty between the two tasks, based on her subjective report after the experiment. A previous functional MRI study has reported that bilateral SPL activity in abacus users was more evident in the tasks with a higher cognitive demand (Hanakawa et al., 2003). In fact, if a lower statistical threshold was used in the mental calculation task, activation in the bilateral SPL was observed.

Regarding the task-specific activity of the SPL, one might argue that the observed differences in SPL activity among arithmetic and other control tasks (such as verbal fluency and hand grip) might be explained by the difference in task difficulty. However, that would be unlikely because the SPL activity during verbal fluency and hand grip tasks was not significantly different compared with the easiest visual fixation condition in which the subject simply watched the fixation on the screen. If the explanation of activity difference by task difficulty is true, then SPL activity during the verbal fluency and hand grip tasks should be greater than during the visual fixation task. Therefore, it is reasonable to consider that the SPL activity would be specific for her mental abacus use after her stroke recovery.

It has been proposed that the human capacity for mathematical intuition depends on both linguistic competence and visuospatial representations (Dehaene et al., 1999). By a combination of neuropsychological and neuroimaging techniques, the present finding provides evidence for an important role of visual imagery in mental arithmetic operations and also for its underlying neural correlates, the superior parietal cortex. The SPL might be an important cortical structure for non-verbal forms of number representation for calculation. The present finding may contribute to developing our understanding of the relationship between mental imagery and mental arithmetic operations.

There are several limitations for this study. First, this is a single case study and it is difficult to generalize this finding to other populations. Second, the patient was left handed and thus it is difficult to discuss the lateralization of brain activation. For this reason, we did not make any conclusions on the lateralization of brain activity from the present study. Third, the results of the behavioral interference task might be explained by a potential difference in difficulty between the distractors, such as the difference in the visual complexity of stimuli. Thus, in future studies, interference tasks should be matched for difficulty and the subject should be asked to make a behavioral response to the interfering stimuli, to be certain that the subject is actually processing the stimuli. Despite these limitations, however, we believe that this result has important implications regarding the neural substrates underlying the superior arithmetic ability of abacus users, because this is the first neuropsychological case report and also the first longitudinal functional MRI study of abacus users.

In conclusion, the present study reports for the first time a case of "abacus-based acalculia" caused by a brain lesion. Together with previous neuroimaging studies, the present result provides evidence for an important role of the PMd and parietal cortex in the mental calculation and digit memory tasks of abacus users.

ACKNOWLEDGMENTS

This work was supported by grants from the Grants-in-Aid for Scientific Research (KAKENHI) to Satoshi Tanaka (24680061).

REFERENCES

Brett, M., Leff, A. P., Rorden, C., and Ashburner, J. (2001). Spatial normalization of brain images with focal lesions using cost function masking. *Neuroimage* 14, 486–500.

Chen, F., Hu, Z., Zhao, X., Wang, R., Yang, Z., Wang, X., and Tang, X. (2006). Neural correlates of serial abacus mental calculation in children: a functional MRI study. *Neurosci. Lett.* 403, 46–51.

Corbetta, M., Kincade, M. J., Lewis, C., Snyder, A. Z., and Sapir, A. (2005). Neural basis and recovery of spatial attention deficits in spatial neglect. *Nat. Neurosci.* 8, 1603–1610.

Courtney, S. M., Petit, L., Haxby, J. V., and Ungerleider, L. G. (1998a). The role of prefrontal cortex in working memory: examining the contents of consciousness. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 353, 1819–1828.

Courtney, S. M., Petit, L., Maisog, J. M., Ungerkeider, L. G., and Haxby, J. V. (1998b). An area specialized for spatial working memory in human frontal cortex. *Science* 279, 1347–1351.

Dancause, N., Barbay, S., Frost, S. B., Zoubina, E. V., Plautz, E. J., Mahnken, J. D., and Nudo, R. J. (2006). Effects of small ischemic lesions in the primary motor cortex on neurophysiological organization in ventral premotor cortex. *J. Neurophysiol.* 96, 3506–3511.

Dehaene, S., Spelke, E., Pinel, P., Stanescu, R., and Tsivkin, S. (1999). Sources of mathematical thinking: behavioral and brain-imaging evidence. *Science* 284, 970–974.

Evans, A. C., Collins, D. L., Mills, S. R., Brown, E. D., Kelly, R. L., and Peters, T. M. (1993). 3D statistical neuroanatomical models from 305 MRI volumes. *IEEE Nucl. Sci. Symp. Conf. Rec.* 108, 1877–1878.

Fiez, J. A., Raife, E. A., Balota, D. A., Schwarz, J. P., Raichle, M. E., and Petersen, S. E. (1996). A positron emission tomography study of the short-term maintenance of verbal information. *J. Neurosci.* 16, 808–822.

Frank, M. C., and Barner, D. (2012). Representing exact number visually using mental abacus. *J. Exp. Psychol. Gen.* 141, 134–149.

Fridman, E. A., Hanakawa, T., Chung, M., Hummel, F., Leiguarda, R. C., and Cohen, L. G. (2004). Reorganization of the human ipsilesional premotor cortex after stroke. *Brain* 127, 747–758.

Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J. P., Frith, C. D., and Frackowiak, R. S. J. (1994). Statistical parametric maps in functional imaging: a general linear approach. *Hum. Brain. Mapp.* 2, 189–210.

Frost, S. B., Barbay, S., Friel, K. M., Plautz, E. J., and Nudo, R. J. (2003). Reorganization of remote cortical regions after ischemic brain injury: a potential substrate for stroke recovery. *J. Neurophysiol.* 89, 3205–3214.

Granà, A., Hofer, R., and Semenza, C. (2006). Acalculia from a right hemisphere lesion dealing with "where" in multiplication procedures. *Neuropsychologia* 14, 2972–2986.

Hanakawa, T., Honda, M., and Hallett, M. (2004). Amodal imagery in rostral premotor areas. *Behav. Brain Sci.* 27, 406–407.

Hanakawa, T., Honda, M., Okada, T., Fukuyama, H., and Shibasaki, H. (2003). Neural correlates underlying mental calculation in abacus experts: a functional magnetic resonance imaging study. *Neuroimage* 19, 296–307.

Hartje, W. (1987). "The effect of spatial disorders on arithmetical skills," in *Mathematical disabilities: a cognitive neuropsychological perspective*, eds G. Deloche, and X. Seron (Hillsdale, NJ: Erlbaum), 121–125.

Hasegawa, T., Kishi, H., Shigeno, K., Tanemura, J., Kusunoki, K., Kifune, Y., and Yoshida, M. (1984). A study on aphasia rating scale: a method for overall assessment of SLTA results. *High. Brain Funct. Res.* 4, 638–646.

Hatano, G., Amaiwa, S., and Shimizu, K. (1987). Formation of a mental abacus for computation and its use as a memory devices for digits: a developmental study. *Dev. Psychol.* 23, 832–838.

Hatano, G., Miyake, Y., and Bink, M. (1977). Performance of expert abacus operators. *Cognition* 5, 57–71.

Hatano, G., and Osawa, K. (1983). Digit memory of grand experts in abacus-derived mental calculation. *Cognition* 15, 95–110.

Hatta, T., Hirose, T., Ikeda, K., and Fukuhara, H. (1989). Digit memory of soroban experts: evidence of utilization of mental imagery. *Appl. Cogn. Psychol.* 3, 23–33.

Hatta, T., and Ikeda, K. (1988). Hemispheric specialization of abacus experts in mental calculation: evidence from the results of time-sharing tasks. *Neuropsychologia* 26, 877–893.

Hécaen, H., Angerlergues, R., and Houiller, S. (1961). Les variétés cliniques des acalculies au cours des lésions rétrorolandiques: approche statistique du problème. *Rev. Neurol. (Paris)* 105, 85–103.

Heiss, W. D., and Thiel, A. (2006). A proposed regional hierarchy in recovery of post-stroke aphasia. *Brain. Lang.* 98, 118–123.

Hishitani, S. (1990). Imagery expert: how do expert abacus operators process imagery? Appl. *Cogn. Psychol.* 4, 33–46.

Hu, Y., Geng, F., Tao, L., Hu, N., Du, F., Fu, K., and Chen, F. (2011). Enhanced white matter tracts integrity in children with abacus training. *Hum. Brain Mapp.* 32, 10–21.

Jonides, J., Smith, E. E., Koeppe, R. A., Awh, E., Minoshima, S., and Mintun, M. A. (1993). Spatial working memory in humans as revealed by PET. *Nature* 363, 623–625.

Ku, Y., Hong, B., Zhou, W., Bodner, M., and Zhou, Y. D. (2012). Sequential neural processes in abacus mental addition: an EEG and FMRI case study. *PLoS ONE* 7, e36410. doi:10.1371/journal.pone.0036410

Luppino, G., Murata, A., Govoni, P., and Matelli, M. (1999). Largely segregated parietofrontal connections linking rostral intraparietal cortex (areas AIP and VIP) and the ventral premotor cortex (areas F5 and F4). *Exp. Brain Res.* 128, 181–187.

Mellet, E., Tzourio, N., Crivello, F., Joliot, M., Denis, M., and Mazoyer, B. (1996). Functional anatomy of spatial mental imagery generated from verbal instructions. *J. Neurosci.* 16, 6504–6512.

Oshio, R., Tanaka, S., Sadato, N., Sokabe, M., Hanakawa, T., and Honda, M. (2010). Differential effect of double-pulse TMS applied to dorsal premotor cortex and precuneus during internal operation of visuospatial information. *Neuroimage* 49, 1108–1115.

Paulesu, E., Frith, C. D., and Frackowiak, R. S. (1993). The neural correlates of the verbal component of working memory. *Nature* 362, 342–345.

Price, C. J., and Crinion, J. (2005). The latest on functional imaging studies of aphasic stroke. *Curr. Opin. Neurol.* 18, 429–434.

Rowe, J., Toni, I., Josephs, O., Frackowiak, R. S., and Passingham, R. E. (2001). The prefrontal cortex: response selection or maintenance within eorking memory? *Science* 288, 1656–1660.

Smith, E. E., Jonides, J., Marshuetz, C., and Koeppe, R. A. (1998). Components of verbal working memory: evidence from neuroimaging. *Proc. Natl. Acad. Sci. U.S.A.* 95, 876–882.

Stigler, J. (1984). "Mental abacus": the effects of abacus training on Chinese children's mental calculation. *Cogn. Psychol.* 16, 145–176.

Tanaka, S., Hanakawa, T., and Honda, M. (2008). Neural substrates underlying cognitive expertise. *Brain Nerve* 60, 257–262.

Tanaka, S., Honda, M., and Sadato, N. (2005). Modality-specific cognitive function of medial and lateral human Brodmann area 6. *J. Neurosci.* 25, 496–501.

Tanaka, S., Michimata, C., Kaminaga, T., Honda, M., and Sadato, N. (2002). Superior digit memory of abacus experts: an event-related functional MRI study. *Neuroreport* 13, 2187–2191.

Ward, N. S., Brown, M. M., Thompson, A. J., and Frackowiak, R. S. (2003). Neural correlates of motor recovery after stroke: a longitudinal fMRI study. *Brain* 126, 2476–2496.

Wise, S. P., Boussaoud, D., Johnson, P. B., and Caminiti, R. (1997). Premotor and parietal cortex: corticocortical connectivity and combinatorial computations. *Annu. Rev. Neurosci.* 20, 25–42.

Wise, S. P., and Murray, E. A. (2000). Arbitrary associations between antecedents and actions. *Trends Neurosci.* 23, 271–276.

Wu, T. H., Chen, C. L., Huang, Y. H., Liu, R. S., Hsieh, J. C., and Lee, J. J. (2009). Effects of long-term practice and task complexity on brain activities when performing abacus-based mental calculations: a PET study. *Eur. J. Nucl. Med. Mol. Imaging* 36, 436–445.

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: 27 April 2012; accepted: 10 August 2012; published online: 28 August 2012.

Citation: Tanaka S, Seki K, Hanakawa T, Harada M, Sugawara SK, Sadato N, Watanabe K and Honda M (2012) Abacus in the brain: a longitudinal functional MRI study of a skilled abacus user with a right hemispheric lesion. Front. Psychology ***3****:315. doi: 10.3389/fpsyg.2012.00315*

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

APPENDIX

Table A1 | Brain activity during mental calculation task.

Cluster size	Voxel *T*	MNI coordinate			Laterality	Anatomy
		x	*y*	*z* (mm)		
ONE-DIGIT MENTAL CALCULATION > VISUAL FIXATION 6 MONTHS AFTER THE STROKE						
952	6.17	−42	−40	68	L	Postcentral gyrus
169	5.36	−62	−2	36	L	Postcentral gyrus
124	5.02	−64	−22	40	L	Supramarginal gyrus
108	5.00	14	−72	−44	R	Cerebellum (lobule VIIb)
178	4.86	−30	42	−10	L	Middle orbital gyrus
176	4.67	−46	0	10	L	Superior frontal gyrus
90	4.59	18	62	24	R	Superior frontal gyrus
163	4.19	−22	−4	74	L	Cerebellum (lobule VIIb)
TWO-DIGIT MENTAL CALCULATION > VISUAL FIXATION 6 MONTHS AFTER THE STROKE						
6452	9.27	−42	−42	68	L	Postcentral gyrus
299	7.32	16	−74	−48	R	Cerebellum (lobule VIIb)
497	6.15	26	−72	−18	R	Cerebellum
312	4.91	−22	−98	10	L	Middle occipital gyrus
1015	5.42	42	−38	62	R	Postcentral gyrus
134	4.28	−46	−64	−8	L	Inferior temporal gyrus
427	4.12	34	44	34	R	Middle frontal gyrus
TWO-DIGIT > ONE-DIGIT MENTAL CALCULATION 6 MONTHS AFTER THE STROKE						
290	6.42	−44	52	10	L	Middle frontal gyrus
ONE-DIGIT MENTAL CALCULATION > VISUAL FIXATION 13 MONTHS AFTER THE STROKE						
225	7.14	14	−74	−42	R	Cerebellum (lobule VIIb)
175	4.70	24	−70	−12	R	Fusiform gyrus
TWO-DIGIT MENTAL CALCULATION > VISUAL FIXATION 13 MONTHS AFTER THE STROKE						
3494	10.06	14	−74	−42	R	Cerebellum (lobule VIIb)
588	7.91	18	−56	78	R	Superior parietal lobule
1417	7.66	−20	−66	64	L	Superior parietal lobule
320	6.80	−50	−4	58	L	Precentral gyrus
581	6.42	−2	−48	−14	L	Cerebellum (vermis)
547	5.43	−22	−90	12	L	Middle occipital gyrus
266	5.73	−60	2	38	L	Precentral gyrus
136	5.51	32	−36	74	R	Postcentral gyrus
172	4.91	38	−24	48	R	Postcentral gyrus
300	4.73	40	58	4	R	Middle frontal gyrus
337	4.55	22	−98	10	R	Superior occipital gyrus
TWO-DIGIT > ONE-DIGIT MENTAL CALCULATION 13 MONTHS AFTER THE STROKE						
461	6.04	−18	−64	64	L	Superior parietal lobule
414	6.00	16	−60	78	R	Superior parietal lobule
95	4.59	30	−34	74	R	Postcentral gyrus
508	4.31	30	−84	10	R	Middle occipital gyrus
182	4.20	42	56	8	R	Middle frontal gyrus

Table A2 | Brain activity during digit memory task.

Cluster size (number of voxel)	*T* value	MNI coordinates			Laterality	Anatomy
		x	*y*	*z* (mm)		
DELAY PERIOD DURING DIGIT MEMORY TASK > VISUAL FIXATION 6 MONTHS AFTER STROKE						
7021	10.15	−62	−20	38	L	Supramarginal gyrus
656	7.97	26	−70	−20	R	Cerebellum (lobule VI)
432	6.89	26	−52	72	R	Superior parietal lobule
368	6.08	−38	34	−4	L	Inferior frontal gyrus
162	5.42	−2	12	58	L	SMA
215	4.68	−48	−54	−12	R	Inferior temporal gyrus
166	4.75	46	52	4	L	Middle frontal gyrus
178	4.7	−28	14	10	L	Insula
DELAY PERIOD DURING DIGIT MEMORY TASK > VISUAL FIXATION 13 MONTHS AFTER STROKE						
10518	18.9	−42	−44	56	L	Inferior parietal lobule
3206	17.87	24	−72	−16	R	Fusiform gyrus
2988	13.06	22	−50	74	R	Superior parietal lobule
495	10.04	−40	−62	−2	L	Middle temporal gyrus
946	8.67	−44	42	22	L	Middle frontal gyrus
1382	8.24	40	54	4	R	Middle frontal gyrus
581	7.75	−26	−66	−24	L	Cerebellum (lobule VI)
1552	7.62	2	−88	20	R	Cuneus
98	3.88	52	−8	20	R	Rolandic operculum

MNI coordinates (x, y, z) and statistical t-values at the peak anatomical voxel and size of cluster (number of voxels) are listed ($P < 0.001$ uncorrected for multiple comparisons at voxel level, $p < 0.05$ corrected for multiple comparisons with Gaussian random field theory at cluster level). L indicates the left hemisphere whereas R indicates the right hemisphere.