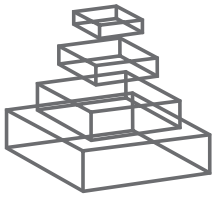
frontiers

RESEARCH TOPICS

# APPROACHES AND ASSUMPTIONS IN HUMAN NEUROSCIENCE

Hosted by
Michael X. Cohen

frontiers in
HUMAN NEUROSCIENCE



# frontiers




Cover image provided by Ibbl sarl, Lausanne CH

**ISSN** 1664-8714
**ISBN** 978-2-88919-015-7
**DOI** 10.3389/978-2-88919-015-7


## 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

# APPROACHES AND ASSUMPTIONS IN HUMAN NEUROSCIENCE

Hosted By
**Michael X. Cohen,** University of Amsterdam, Netherlands

The human brain is arguably the most complex system we know of. Over the past few decades, scientists have developed several methods and theories for studying the functional organization of the brain, and how cognitive/perceptual/emotional processes might arise from the brain's electro-chemical-computational dynamics. These methods facilitated and inspired large literatures on brain-behavior links, and yet there remains a seemingly endless chasm between our simple impoverished models and the unfathomable complexity of the human brain. The purpose of this Research Topic is to ask the question: Are we thinking about thinking about the brain in the right way?

In most scientific publications, researchers describe a broad and established theoretical framework and briefly describe new experimental results consistent with that framework. Here, we encourage authors to express ideas that might be radical, controversial, or different from established theories or methodological approaches. Supportive data are highly encouraged. The aim is to spark discussions about the validity and usefulness of current methodological/theoretical approaches in human cognitive neuroscience, with the goal of inspiring new approaches and ways of thinking.

Neuroscience is a massive field with myriad methodological and theoretical approaches; we focus this special issue on approaches most commonly used in human neuroscience.

# Table of Contents

frontiers in
HUMAN NEUROSCIENCE

EDITORIAL
published: 19 January 2011
doi: 10.3389/fnhum.2011.00002

# It's about time

***Michael X Cohen*[1,2]***

[1] *Department of Psychology, University of Amsterdam, Amsterdam, Netherlands*
[2] *Department of Physiology, University of Arizona, Tucson, AZ, USA*
** Correspondence: mikexcohen@gmail.com*

## SHADOWS IN A CAVE

As cognitive neuroscientists, we want to understand how the dynamics of the brain lead to dynamics in cognition and behavior. But the brain is perhaps the most complex, mysterious, and enigmatic information processing system that we know of, and cognitive processes remain debated in terms of how they should be defined, categorized, and tested. Thus, the problems cognitive neuroscientists try to solve are poorly defined on both the cognitive and neuroscience sides.

In some sense, we have come a long way: Specific cognitive/perceptual/emotional/motor functions have been linked to activity in specific brain regions or circuits of brain regions; neurochemicals have been identified as necessary or relevant for certain aspects of emotion, learning, memory, and action; white matter pathways linking different brain regions have been implicated in specific diseases and behavioral characteristics. Compared to a century ago, when Overton (1897) suggested that thinking is done by "forehead cells," our understanding of the brain has increased tremendously. But even "simple" processes like categorizing a visual stimulus as animal or automobile, maintaining a specific amount of force on a grip, or slowing response time after errors, turn out to have complex neural correlates that remain debated and poorly understood.

In his famous cave allegory, Plato describes prisoners who spend their lives chained to the side of a bridge in the middle of a cave. The bridge is a passageway, and people, animals, and vehicles traverse the bridge to get through the cave. The cave is lit from behind the bridge by a fire. But the prisoners are chained such that they cannot see what is behind them; they see only the flickering shadows cast in front of them on the cave wall. The prisoners do not know what shapes produce the shadows, and, because they spent their entire lives looking at shadows, these shadows – and not the shapes that produce them – are their view of reality.

We are like the prisoners in the cave. There are platonic "biological bases of behavior" that we want to discover (the figures walking on the pathway behind the prisoners), but all we can observe are the shadows cast on the wall (empirical data) by the flame in the back of the cave (methods and technologies); our concept of the nature and shape of the figures (theories) are shaped by past experience, intuition, and, perhaps most importantly, how the light of the flame defines the shadows (**Figure 1**).

However, we have one important advantage over the prisoners in Plato's cave: We can, to some extent, control the flame. We can develop new technologies and methodologies, and we can combine methodologies in interesting, novel, and insightful ways. We can compare the shadows cast on the wall using different materials to fuel the fire.

Here I will argue that too much attention has been focused on investigating neurocognitive function based on attempts to localize processes in space (i.e., functional localization). Instead, fruitful insights might arise from considering *time* to be an important factor in neurocognitive function. Indeed, for some neurocognitive processes, time may be as important, or possibly more important, than space in terms of the underlying neurocomputational mechanisms.

## FUNCTIONAL LOCALIZATION: THE STANDARD APPROACH

The way we as cognitive neuroscientists typically link dynamics of the brain to dynamics of behavior is by correlating increases or decreases of some measure of brain activity with the cognitive or emotional state we hope the subject is experiencing at the time. The primary dependent measure in the majority of these studies is whether the average amount of activity – measured through spiking, event-related-potential or -field component amplitude, blood flow response, light scatter, etc. – in a region of the brain goes up or down. In this approach, the aim is to reduce this complex and enigmatic neural information processing system to two dimensions: Space and activation (up/down). The implicit assumption is that cognitive processes can be localized to specific regions of the brain, can be measured by an increase in average activity levels, and in different experimental conditions, either operate or do not.

It is naïve to think that these two dimensions are sufficient for characterizing neurocognitive function. The range and flexibility of cognitive, emotional, perceptual, and other mental processes is huge, and the scale of typical functional localization claims – on the order of several cubic centimeters – is large compared to the number of cells with unique physiological, neurochemical, morphological, and connectional properties contained in each MRI voxel. Further, there are no one-to-one mappings between cognitive processes and brain regions: Different cognitive processes can activate the same brain region, and activation of several brain regions can be associated with single cognitive processes. In the analogy of Plato's cave, our current approach to understanding the biological foundations of cognition is like looking at shadows cast on a region of the wall of the cave without observing how they change dynamically over time. This makes it difficult to disentangle shadows cast by different but overlapping shapes (**Figure 1**).

## ARE COGNITIVE PROCESSES LOCALIZABLE IN SPACE?

Yes and no, depending on the spatial resolution and the cognitive process in question. At a gross level (e.g., several cubic centimeters), some cognitive functions appear to be localized to specific brain regions when using specific statistical thresholding: The hippocampus seems to be the locus of some aspects of long-term memory formation and retrieval; decoding the visual world occurs largely in occipito-temporal regions, volitional control of the body occurs in cortical motor areas and basal ganglia nuclei.

The fusiform gyrus has received considerable attention for the issue of functional localization. There is a region of the fusiform gyrus that increases in activation for a wide range of visual object categories, but that activates preferentially for faces relative to other

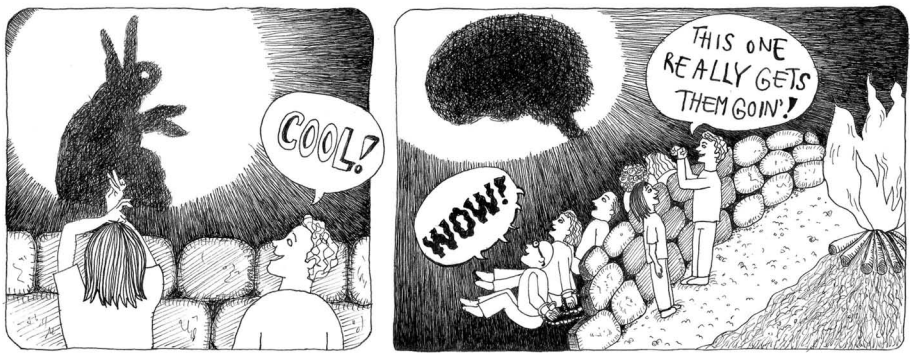



**FIGURE 1 | Like prisoners in Plato's cave, our view of the brain is shaped by how we measure it (i.e., the shadows cast by the flame).** Drawing by Dr. Sanne de Wit.

comparable visual stimuli (Sergent et al., 1992; Haxby et al., 1994; Puce et al., 1996). For this reason, the "fusiform face area" is argued to be a modular region in the brain that specializes in processing faces (Kanwisher et al., 1997; Kanwisher, 2000). It seems clear that some computations related to face perception can be localized to this particular brain area (Kanwisher, 2010). Even in the case of face recognition, however, which is one of the strongest examples of a localized "module" of cognitive processes, it remains unclear what this localization of function entails – alternative accounts suggest that this area is specialized not for faces *per se*, but instead for categories of expertise (Gauthier et al., 2000a,b; Tarr and Gauthier, 2000) or individuation (Gauthier et al., 2000b; Rhodes et al., 2004). Further, distributed and overlapping representations of faces and objects may exist in larger regions of ventral temporal cortex (Haxby et al., 2001).

There are many spatial scales in the brain that differ in size by several orders of magnitude, ranging from single neurons to cortical columns to meso- and macroscopic populations. It is unclear what the appropriate spatial scale is for functional localization, or whether different neurocognitive processes can or should be localized at different spatial scales. Dynamics at some spatial scales may or may not be relevant for dynamics at other spatial scales (Kiebel et al., 2008).

It is tempting to argue that lesion studies provide the most compelling evidence for functional localization: The necessity of a region in a particular cognitive process. However, lesion studies must be interpreted with caution. First, there may be functional and anatomical reorganization/compensation, even within hours or days of damage (Sanes et al., 1988). Second, lesions may cause impairments indirectly by destroying a "way-point" for information flow or fibers-of-passage (Goulet et al., 1998). Third, it is uncertain how the functioning of a damaged (human or non-human animal) brain can be generalized to functioning of a healthy human brain.

Since the emergence of functional neuroimaging, most studies are based on the principle of functional localization – searching for a one-to-one mapping between region of the brain and psychological process. It is no exaggeration to state that standard preprocessing and analysis protocols (mass-univariate statistics and cluster-based thresholding) are specifically designed to find relationships between particular brain locations and particular cognitive functions. The underlying assumptions are that a mass of brain tissue uniformly increases in activity in response to an experiment event, and that this region must be "big enough" to be considered significant. Typically, however, in these studies there is not a single region that is activated, but rather many regions, producing long tables of activated areas. These long tables are difficult to reconcile with the assumptions of the functional localization approach. This lack of one-to-one mappings between brain region and function (Price and Friston, 2005) may in fact be an accurate reflection of the physically separated but functionally linked networks that underlie neurocognitive function (Varela et al., 2001; Guye et al., 2008; Bassett and Bullmore, 2009; Bullmore and Sporns, 2009).

At a spatial scale finer than possible with fMRI, cognitive processes appear even less localizable and more dependent on sparse spatial patterns (Fujisawa et al., 2008; Quian Quiroga and Panzeri, 2009). For example, the homuncular mapping of the body on primary motor area M1 is considerably more distributed and with considerably more overlap among regions than traditionally thought (Schieber, 2001). Even the concept of a cortical "column" – a small-scale functionally homologous unit of cortex – upon close inspection turns out to be a mysterious and fluid idea that oversimplifies the complex mesoscopic organization of the brain; columns are less functionally and anatomically homogeneous, and more dynamic over time and space, than previously thought (Rockland, 2010).

It should be noted that some physiological functions appear to be more reliably localized in the brain, for example the superchiasmatic nucleus may be the "location" of our circadian rhythm. Clearly, these areas influence cognitive function, but the cognitive processes typically under investigation in cognitive neuroscience studies do not appear to be precisely localizable.

In summary, functional localization has been a useful approach and was critical for the development of cognitive neuroscience theories, experiments, and statistical measures. However, its simplicity may be its greatest limitation. It is likely that the brain uses more dimensions for information processing than just space and activation magnitude. This is not meant to imply that space is irrelevant for information coding/processing, or that functional localization is inappropriate or invalid. Rather, after this initial period of studying functional localization and learning about its merits and limitations, it is perhaps useful to consider *time* as an important factor

for neural information processing. Time will help forge new brain-behavior links and may provide insights into human neurocognitive function beyond what can be learned from focusing on space.

## IT'S ABOUT TIME

Time is a factor that is often *though not always* ignored in human cognitive neuroscience, and yet several considerations suggest that neural systems may use time as a factor for information coding, processing, and transmission. Indeed, time may be as important – if not more important – than space for information processing, particularly at the level of small populations of cells (spatial scale of millimeters to a few centimeters). As described below, "time" refers to rapid dynamics in electrochemical signals that are often but not necessarily oscillatory. Time as latency in functional MRI (e.g., the hemodynamic response peaks about 6–8 s after a stimulus) or an event-related component (e.g., the average voltage deflection 300–600 ms after a stimulus) is not taking into account the rich information that appears to be embedded in the temporal dynamics of neural activity.

There are several empirical and theoretical reasons why time may be in important factor in neural information processing.

(1) *There appears to be information carried in the precise timing of activity within and across physically separated areas of the brain that cannot be measured by overall activity levels in any individual brain region.* "Information" here can refer simply to quantifiable measures of brain activity that predict the cognitive state or behavioral response of the subject. In some cases, temporal dynamics of neural activity are significantly related to the task events while the overall amount of activity averaged over time is not. These kinds of results provide direct evidence that information in the brain is embedded in the rich temporal landscape of electrophysiological activity, and is lost when averaging activity over larger periods of time. Examples will be outlined in a subsequent section.

(2) *In an information-theoretic sense, time provides a large number of possibilities for information to be represented and processed continuously, rapidly, and simultaneously (in parallel) in multiple functionally distinct networks that overlap in time and space.* Time provides a rich source of complex multi-dimensional data in which information can be represented and processed. The large amount of information provided by the temporal dynamics of neural activity arises in part because electrophysiological activity of the brain is strongly oscillatory. These oscillations reflect rhythmic fluctuations in the excitability of populations of neurons (Tiesinga et al., 2008; Wang, 2010). Oscillations occur in multiple temporal and spatial scales, ranging from ultra-slow oscillations with a periodicity of tens of seconds over much of the cortex during deep sleep (Steriade, 2006) to ultra-fast oscillations with a periodicity of a few milliseconds within patches of somatosensory cortex (Curio, 2000). Oscillations that seem most relevant for cognitive processes range from delta (~1–4 Hz) and theta (~4–8 Hz) to gamma (~30–100 Hz; for general reviews of neural oscillations, see Varela et al., 2001; Buzsaki and Draguhn, 2004; Traub et al., 2004). Because activity in one frequency band may occur independently of, and in parallel with, activity in other frequency bands, there is considerable bandwidth for information processing. For example, it has been suggested that multiple alpha sub-bands can be functionally dissociated in their roles in memory processes (Klimesch et al., 2007). Thus, if different neurons are "tuned" to different frequency bands (Jacobs et al., 2007), multiple functionally distinct neural networks can spatially coexist and be dissociated according to frequency band or spatiotemporal patterns (Akam and Kullmann, 2010).

Even from the activity recorded from a single electrode, there are already multiple domains of information, including frequency (the speed of the oscillation), power (the amount of the energy in a frequency band at a point in time), and phase angle (the position of the oscillation along the sine wave, driven by the state of excitation of the population of neurons; **Figure 2**; see also Makeig et al., 2004). And because these dimensions are largely independent of each other, a single electrode in a single location in the brain can measure multi-dimensional local neural dynamics. Interactions can occur across these dimensions (e.g., phase–amplitude cross-frequency coupling), suggesting that information may be embedded not only in the dynamics of one of these dimensions, but also in the interactions among dimensions. And because they all occur in the neural population contributing to one electrode, analyses based on spatial localization of average activity levels might be blind to some of these dynamics. Indeed, space is another dimension that increases the potential for information processing and complexity: Interactions can occur across physically separated networks over different frequency bands, and among power and phase.

What this means is that information processing schemes that take advantage of time can utilize many dimensions (information processing possibilities) simultaneously. This is in contrast with the standard functional localization approach, which, as discussed earlier, is limited to two dimensions: Functions occur in specific regions or are indexed by specific ERP components, and are either operating or are not.

(3) *There is arguably selection pressure for individuals and species carrying neural systems that can decode the sensory world, make decisions, and adapt behavior faster and more efficiently.* The fastest known behavioral response that is mediated by a neural connection is the snap closure of the mandible of the *Odontomachus* ant. It takes 8 ms for a hair to trigger receptor cells in the jaw, the message to be sent via the largest axons in insects or vertebrates to the brain, and then another signal to be sent back to the muscles to snap the jaw closed (Holldobler and Wilson, 1998). This mechanism provides the *Odontomachus* an unparalleled ability to attack and to escape (by snapping the jaws against a hard surface, the ant can fly backward over 40 cm). Although this neurally mediated response is rigid, its speed and efficiency give this ant a significant advantage during battle.

As neural systems and the environments in which they operate become more complex and less predictable, flexibility, and adaptability become critical. However, speed of processing has not entirely been compromised. Relatively simple perceptual decisions such as gradient orientation or color discrimination can be done with high accuracy with as little as 10 ms presentation time (Bodelon et al., 2007). Indeed, the macaque brain may

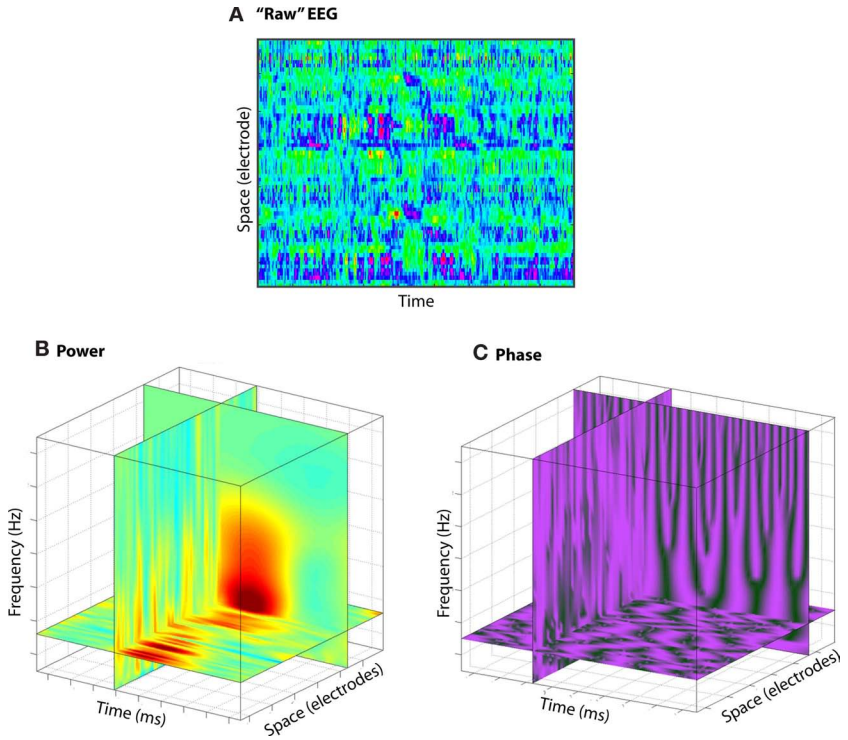



**FIGURE 2 | Although EEG is traditionally conceptualized as two-dimensional [time and space (A)]**, EEG data contain more information, including frequency and power/phase **(B,C)**.

require as little as 30 ms to make simple color discriminations (Stanford et al., 2010). Cortical responses to sensory deviancy occur within 100 ms of stimulus onset (e.g., the mismatch negativity; Garrido et al., 2009; Kujala and Naatanen, 2010). Seemingly complex processes in humans also occur rapidly. Electrical signals generated in the medial frontal cortex that reflect errors or response conflict (when multiple responses are activated but only one can be selected) begin shortly before the button press or stimulus onset, and peak about 80 ms after the button press (van Veen and Carter, 2002; Padilla et al., 2006). When subjects are given external performance feedback, electrical activity over medial frontal cortex begins to distinguish positive from negative feedback as early as 200 ms after feedback presentation (Gehring and Willoughby, 2002; Holroyd and Coles, 2002). In some cases, these signals predict decisions subjects make a few seconds later (Cohen and Ranganath, 2007; Cavanagh et al., 2010). Some of these processes occur in absence of conscious awareness (van Gaal et al., 2008; Cohen et al., 2009d).

Though not conclusive in establishing that there must be information processing/transfer schemes based on temporal coding, these observations indicate that processing information as rapidly as possible while maintaining flexibility and adaptability is ubiquitously needed and often observed. In other words, to the extent that speed and flexibility in responding to unpredictable changes in the environment provides an advantage for survival, animals containing neural systems that take advantage of the processing capabilities afforded by time may have been more likely to reproduce and pass on their genes (and neural processing power and coding schemes).

(4) *Neural activity is inextricably linked to cognition and behavior in time, but not in space.* There is a direct relationship between the timeframe of brain processes and the timeframe of the corresponding cognitive and behavioral processes. A fast neural process implies a fast cognitive process, and this in turn determines the speed of initiating or adjusting behavior. Indeed, our cognitive and neural systems appear well equipped for estimating and attending to time (Ivry, 1996; Fuster, 2001; Coull, 2004; Coull et al., 2004; Nobre et al., 2007; Ivry and Schlerf, 2008). In contrast, the spatial organization of neural processes is arbitrarily related to cognitive processes. Thus, whether a neural process is at one or another location, or distributed throughout the brain, has no implications for the corresponding (location of) behavior, except if physical location can constrain temporal dynamics. For example, would we be any different if our amygdalae were 3 cm more dorsal than they currently are? What about if they took 3,000 instead of 170 ms to respond to a threatening facial expression? The fact that brain activity is time-locked, rather than space-locked, to behavior implies that time will be highly informative about behaviorally relevant neurocognitive mechanisms. The brain does indeed exhibit some spatial relationships with the body (e.g., homuncular organization of sensorimotor regions, retinotopic organization of

visual areas), although these examples still exhibit some arbitrary relations to behavior, such as the left–right and up–down crossovers.

(5) *Controversies develop in cognitive neuroscience over the precise functional role of a specific region, but some of these controversies may be moot because multiple functionally distinct neural networks may coexist in the same space.* Indeed, in these cases, empirical evidence may seem conflicting because different theories can be supported by different experiments. For example, it is widely accepted that the anterior cingulate cortex and surrounding medial frontal cortex is involved in monitoring actions (Ridderinkhof et al., 2004). However, different theoretical accounts have argued over whether this region monitors conflict, errors, or error likelihood (Botvinick, 2007; Carter and van Veen, 2007); signals that behavioral adjustments are needed (Kerns et al., 2004) or implements them directly (Taylor et al., 2007); integrates information about reward history or uncertainty (Rudebeck et al., 2008; Rushworth and Behrens, 2008), and so on. Considering that there is empirical evidence for all of these propositions, it seems likely that functionally different networks can emerge from the same population of anterior cingulate cortex neurons, depending on task demands (Fujisawa et al., 2008). Indeed, event-related potential studies demonstrate that within ~300 ms, medial frontal scalp potentials can code for several different behaviorally relevant experiment parameters including reward magnitude, probability, and significance for learning (Philiastides et al., 2010).

In other words, rather than attempting to resolve a grand-unified-theory for the function of a region of the brain (in this example, the anterior cingulate cortex), attention might be better spent trying to understand how that area may utilize temporal schemes to compute and coordinate the diverse functions suggested by empirical evidence.

## EVENT-RELATED POTENTIALS

Most cognitive EEG studies report event-related potentials. The event-related potential is simply the time-domain average of EEG traces locked to the onset of some experimental event such as stimulus onset or button press. The reasoning behind this approach is that background noise in the EEG is averaged out over many trials, and the remaining fluctuations reflect activation of different cognitive systems. Peaks are named according to voltage sign relative to a pre-stimulus period and approximate peak time (N100, P200, etc.). The components are thought to reflect activation of modular cognitive or perceptual systems (Luck, 2005). The dependent variable is usually the peak amplitude of some component (e.g., the P300), the average voltage over some larger time window, or a peak-to-peak or base-to-peak amplitude difference.

To the extent that neuroelectric dynamics are oscillatory and non-phase-locked to the event, a considerable amount of cognitively relevant information in EEG may be lost in time-domain averaging (Makeig et al., 2004). This is illustrated in **Figure 3**. Although an extreme example, non-phase-locked (and, therefore, not observed in ERP) dynamics are often observed in real data. Indeed, many of the findings reviewed in the next section would not be observed using event-related potentials. For this reason, a lack of differences in event-related potentials between conditions may be difficult to interpret because many aspects of neural dynamics that are not apparent in event-related potentials might differ between conditions, e.g., if the neurocognitive processes under investigation recruits non-phase-locked dynamics or high frequency oscillations.

Thus, event-related potentials are useful for providing a glimpse of the global neural processing, but may be of limited use for elucidating complex electrophysiological dynamics.

## EXAMPLES OF TIME-EMBEDDED INFORMATION IN HUMAN ELECTROPHYSIOLOGICAL ACTIVITY

Considerable work has been done in animals and in computational models regarding how time may be used to encode information. Examples include: The timing of the first post-stimulus action potential in auditory cortex encodes sound amplitude (Heil, 2004); the timing of hippocampal place cells with respect to simultaneous theta phase improves statistical localization of the rats' position based on physiology data (Jensen and Lisman, 2000; Jensen, 2001) and the timing of those action potentials provides independent information compared to average firing rate (Huxter et al., 2003); the timing of specific neuron firing in early visual cortex encodes gradient phase (Aronov et al., 2003); synchronization of local field potentials has been suggested to link disparate sensory modalities to form unified representations (i.e., binding; Engel et al., 1997; Singer, 1999; Fries et al., 2007); computational simulations suggest that the phase of neural activity may be sufficient for pattern completion (Knoblauch and Palm, 2001; Gutierrez-Galvez and Gutierrez-Osuna, 2003).

These and other findings in animals have laid important groundwork for understanding how the brain might use time to encode information, and have inspired many studies in human neuroscience. But because the spatial scale investigated in animals and model simulations is smaller than what is typically available in humans, and because it is not known to what extent human neurocognitive functions operate the same as those of other animals (though presumably some fundamental principles are conserved across species), this section will focus on relevant work in humans.

Although the literature on time-based coding schemes and sophisticated analyses of electrophysiological data is overshadowed by the literature on fMRI-based localization studies, there are too many relevant and insightful findings to discuss all them all here. Instead, this section will highlight three examples of how mathematical analyses of the temporal dynamics of human electrophysiological recordings have shed insight into neurocognitive function. These examples also illustrate cases in which standard localization- and hemodynamic-based analyses would be unlikely to reveal these brain dynamics (e.g., because no overall increase in space-averaged activity occurs).

(1) *Cross-frequency coupling.* Cross-frequency coupling refers to a relationship between activities in two different frequency bands. For example, the power of gamma (~30–80 Hz) oscillations may vary as a function of the phase of theta (~4–8 Hz). Cross-frequency coupling may be used for information coding if the lower frequency oscillations coordinate the activity of sub-

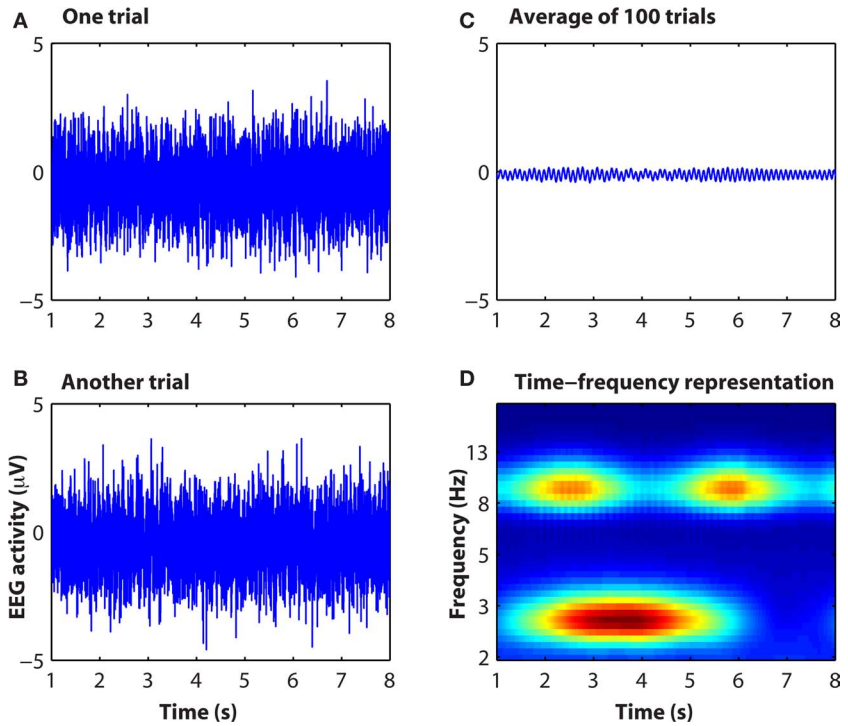



**FIGURE 3 | Simulated data showing how information contained in raw EEG data** [**(A,B)**: single "trials"] is not apparent in the event-related potential **(C)** but is readily observable in the time–frequency representation **(D)**. Matlab code to run this simulation is available from the author.

populations of cells that use higher frequency oscillations to process information. Cross-frequency coupling has been suggested to be a generic brain mechanism for information processing (Lisman, 2005; Jensen and Colgin, 2007), and it likely involves dynamics of multiple neural populations that overlap in time and space. There are several ways in which cross-frequency coupling can be quantified (Mormann et al., 2005; Canolty et al., 2006; Cohen, 2008; Tort et al., 2010); different methods may be suited for different purposes, but all methods generally test for a modulation of activity in one frequency band as a function of activity in another (typically, relatively lower) frequency band.

Cross-frequency coupling has been linked to a variety of human cognitive processes (Canolty et al., 2006; Sauseng et al., 2008; Axmacher et al., 2010; Griesmayr et al., 2010). For example, increases in gamma–theta synchronization strength were reported to increase with working memory load, although there was no reported significant change in overall oscillation power at those specific frequency bands (Axmacher et al., 2010). These findings support a model of working memory that predicts that theta acts to coordinate activation of stimulus representations stored in gamma band activity (Lisman, 2010). The human nucleus accumbens exhibits robust gamma–alpha coupling that differentiates reward from punishment feedback (even though average gamma power does not differentiate these conditions; Cohen et al., 2009a), and predicts, on a trial-by-trial level, the extent to which patients adjusted their decision-making time on subsequent trials (Cohen et al., 2009b). Posterior alpha and gamma power have been linked to frontal theta phase during familiar visual perception (Demiralp et al., 2007), and errors due to lapses in attention (Mazaheri et al., 2009), consistent with alpha–gamma coupling in posterior cortex during visual perception (Osipova et al., 2008). Within the medial frontal cortex, alpha–theta coupling has been linked to reward and punishment evaluation (Cohen et al., 2009c).

Cross-frequency coupling is unlikely to elicit a hemodynamic response as measured with fMRI. The reason is that frequency band-specific activity levels may not change over a timescale measurable with fMRI; rather, it is the precise timing of activity that fluctuates, e.g., 10–20 times per fMRI measurement (**Figure 4A**).

Although there are potential methodological issues to be considered (discussed later), examination of cross-frequency coupling has the potential to shed insight into human neurocognitive function beyond what is possible though fMRI or time-domain averages ERP averaging. Cross-frequency coupling measures a putative mechanism by which spatially overlapping but functionally heterogeneous neural networks can be activated and coordinated in a rapid timescale.

(2) *Inter-regional oscillatory synchronization.* In addition to dynamics across frequency bands within the same region of space, information may be embedded in the temporal relationship of activity over space. Inter-regional phase synchronization (a frequency band-specific measure of functional connectivity) may underlie information transfer and co-processing (Knight, 2007; Womelsdorf et al., 2007). And because changes in phase synchronization may occur without any concomi-

tant changes in power (Heinzle et al., 2007), there might be information embedded in the temporal relationship between areas that is not localized to either region alone.

For example, inter-regional oscillatory synchronization may be the mechanism by which the medial frontal cortex interacts with other brain systems, such as lateral prefrontal cortex to implement cognitive control after errors in speeded reaction-time (Hanslmayr et al., 2008; Cavanagh et al., 2009) or reinforcement learning (Cavanagh et al., 2010) tasks, with occipital cortex to bias sensory processing during go/no-go tasks in which no-go cues were difficult to perceive (Cohen et al., 2009d), or with the nucleus accumbens during reinforcement learning and reward anticipation (Cohen et al., 2009b, 2011). Long-range cortico-cortical phase synchronization has also been linked to conscious visual perception (Melloni et al., 2007), working memory (Palva et al., 2010), and other aspects of top-down control (Engel et al., 2001). Further, several brain disorders ranging from schizophrenia to ADHD to Alzheimer's are associated with aberrant patterns of oscillatory phase synchronization (Uhlhaas and Singer, 2006; Stam, 2010).

Although some measures of phase synchronization are non-directional, meaning it is not possible to determine the direction with which oscillations are traveling, the temporal precision of EEG allows one to estimate directional flow of activity. For example, using spectral Granger causality (Granger, 1969; Cui et al., 2008), we found that directed synchronization from the medial frontal cortex to the occipital cortex increased after response errors in a visually guided go/no-go task in which no-go cues were difficult to perceive (Cohen et al., 2009d). These patterns of inter-regional directional synchrony were not mirrored in the levels of activation (oscillation power) of either frontal or visual cortices, suggesting that there was information contained in the inter-regional interactions that could not be localized to either region alone. Granger causality has also been used to show that consciously perceived words, compared to subliminally presented words, are associated with long-range directional synchronization (Gaillard et al., 2009). Partial directed coherence analyses demonstrate changes in directional cortical wave activity during sleep (Kaminski et al., 1997; De Gennaro et al., 2004), and interactions among motor, frontal, and parietal systems during response switching (Gladwin et al., 2006).

More generally, these findings illustrate an important feature of brain network phenomena: Widespread neural networks may synchronize and desynchronize within hundreds of milliseconds (Varela et al., 2001). This is important because ideas about connectivity and network functioning are becoming increasing popular in human and cognitive neuroscience (Sporns, 2010). From the findings reviewed here, it seems that many instances of brain functional network formation are transient and may be best measured using measurement technologies that take maximal advantage of time. Indeed, transient synchronizations in absence of changes in local power might not elicit a hemodynamic response or event-related potential (**Figure 4B**).

(3) *Microstates and other transient electrophysiological events.* Microstates refer to brief periods of cortical electrophysiological activity that are topographically stable over tens to hundreds of milliseconds (Lehmann et al., 2006). Microstates fluctuate 1–2 orders of magnitude faster than the hemodynamic response, and have been linked to visual perception, error processing, and resting state (Muller et al., 2005; Britz and Michel, 2010). They are sometimes accompanied by hemo-

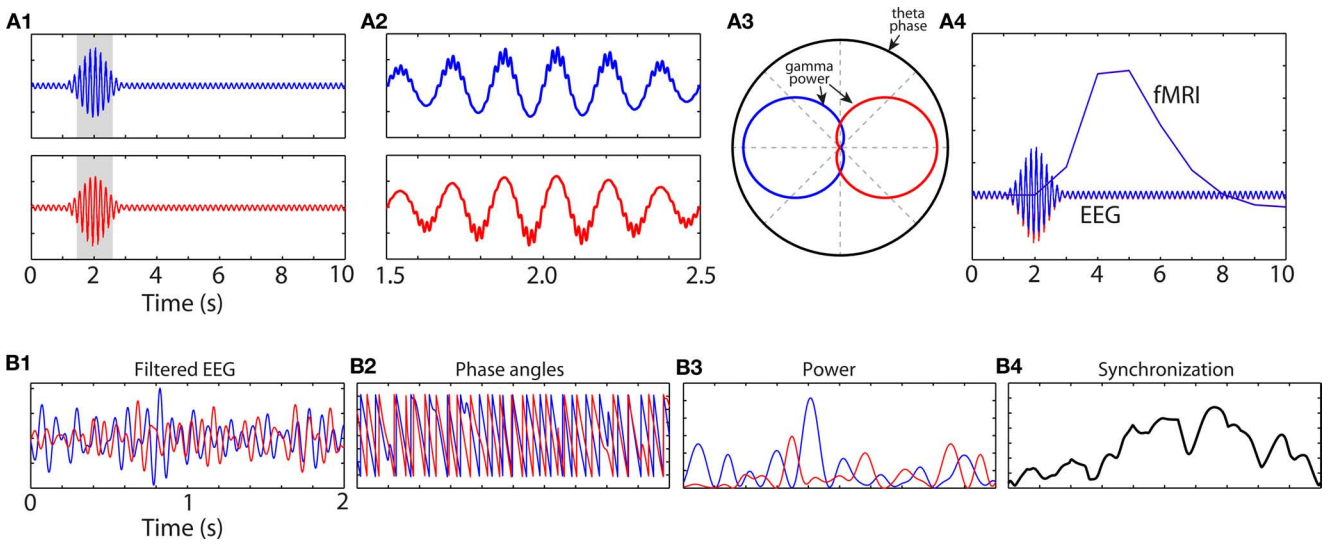



**FIGURE 4 | Simulated data showing how information may be contained in the fine temporal landscape of M/EEG data, and how that information might go undetected using more temporally coarse methods such as fMRI.** **(A1)** An increase in theta-band oscillation power peaking at 2 s is not different between conditions "blue" and "red." However, close inspection of the grayed areas reveals cross-frequency coupling **(A2)** such that gamma power is concentrated at different theta phase regions between the conditions **(A3)**. **(A4)** shows the predicted hemodynamic response (red and blue lines are overlapping). **(B1)** Bandpass filtered EEG data from two channels, and **(B2)** their phase angle time series. **(B3)** shows band-specific power at each site; note that inter-channel synchronization **(B4)** does not occur at times of increased power at either site, showing the independence of local activity **(B3)** and network synchronization **(B4)**.

dynamic responses (Britz et al., 2010; Musso et al., 2010). Other brief cortical events include endogenous "bursts" of frontal alpha asymmetry (Allen and Cohen, 2010) that have been linked to depression. Transient bursts of synchronized electrophysiological activity also occur during sleep, namely spindles and ripples, which have been linked to memory formation and dream recall (Axmacher et al., 2008).

Relatedly, transient pre-stimulus oscillatory dynamics predict performance on the upcoming trial. For example, specific pre-stimulus alpha and theta phases predict performance on perceptual (Mathewson et al., 2009; Busch and VanRullen, 2010), memory (Guderian et al., 2009), cognitive control (Mazaheri et al., 2009; O'Connell et al., 2009; Eichele et al., 2010), and switching (Gladwin et al., 2006) tasks. These and other similar findings demonstrate that there are transient but cognitively meaningful brain states that modulate upcoming task-related performance. Indeed, these pre-stimulus dynamics may be causally involved in perceptual processes, as suggested by TMS manipulations (Romei et al., 2010). Because these pre-stimulus dynamics are transient and driven by phase and not amplitude, they are unlikely to be observed with time-domain averaging or the hemodynamic response (see Mathewson et al., 2009, **Figure 4** for an example with empirical data).

In conclusion, these examples illustrate cases in which cognitively meaningful brain dynamics are related to the subjects' cognitive state, or predict upcoming performance, while overall levels of average activity do not. The point here is not to argue that spatiotemporal averaging or low temporal resolution imaging is invalid or inappropriate; rather, the point is to stress that there are vast and complex neural dynamics that are relevant for understanding neurocognitive function that occur "below the radar" of fMRI and event-related potentials.

## HOW TO STUDY TIME-BASED PROCESSING SCHEMES

In humans, the primary tools for studying electrophysiological activity are EEG and MEG. In some cases, it is possible to record EEG intracranially directly from patients with electrodes implanted for epilepsy or deep brain stimulation. Due to the high temporal resolution of EEG – a sample of electrical/magnetic brain activity can be recorded from each of dozens or hundreds of channels multiple times each millisecond – it is possible to examine the rich temporal landscape of cortical activity, and observe phasic changes in synchronization and desynchronization that occur over tens to hundreds of milliseconds. As discussed earlier, these complex cortical dynamics occur 1–2 orders of magnitude faster than the BOLD response, and may be lost in time-domain EEG averaging during standard event-related potential analyses.

There are several advanced mathematical tools that are appropriate for extracting the fine spatiotemporal oscillation dynamics from EEG data, including but not limited to short-time fast Fourier transform, complex wavelet convolution, multi-taper (in combination with Fourier transform or wavelet convolution), autoregressive coefficients, and bandpass filtering with the Hilbert transform. With careful parameter selection, these methods can produce nearly identical results (Quian Quiroga et al., 2002; Bruns, 2004), although in practice each approach has advantages and limitations. Functional connectivity can be estimated through spectral coherence, phase synchronization, power correlations, Granger causality, partial directed coherence, cross-correlations, etc. Several free toolboxes for analyzing data exist, including but not limited to EEGLAB, fieldtrip, spm8, BIOSIG, and BSMART. Some labs write in-house code for processing and analyses.

The analyses used in this research are not simple, nor are the methods standardized and widely used. New methods and ideas are continuously injected into the field. Although this "wild west" atmosphere provides researchers with the flexibility and freedom to custom-tailor mathematical and statistical approaches that can be optimized for the hypotheses at hand, it also makes entry into the field difficult for scientists who lack the background and experience in signal processing and programming.

However, even without performing relatively complex frequency- or synchronization-based analyses, single-trial analyses (in the time-domain or time–frequency domain) can better link cognitive/behavioral to neural dynamics compared to cross-trial averaging (Debener et al., 2007; Mars et al., 2008; Rousselet et al., 2008), and may be particularly relevant for linking EEG to fMRI (Debener et al., 2005, 2007; De Martino et al., 2010).

One practical issue is that because of volume conduction – the influence of deep and/or distant sources to many electrodes – it may be difficult to distinguish true inter-regional synchronization from artificially high synchronization due to different electrodes recording the same activity. This is a larger issue for EEG than for MEG. Ignoring zero-phase lag synchronizations may combat this issue, although there may be biologically relevant zero-phase lag synchronizations in the brain (Konig et al., 1995; Rajagovindan and Ding, 2008; Vicente et al., 2008). Another approach is to apply a spatial high-pass filter or other spatial transform such as current-source-density or Laplacian, which helps minimize contributions of deep/distant sources that project to many electrodes (Kayser and Tenke, 2006; Srinivasan et al., 2007), or independent components analysis, which estimates unique sources of variance in the brain (Makeig et al., 1997). Finally, one can estimate the cortical generators via beamforming, minimum-norm estimates, or dipole modeling, and then perform analyses in source-space. This approach is also not without drawbacks, because there is no unique inverse solution for any given cortical topography, and different methods may yield different estimates of source activity. Of course, there are advantages and limitations to every methodological approach that one must consider when interpreting results.

## LIMITATIONS OF TIME-BASED INFORMATION CODING/ PROCESSING SCHEMES

Recording electrophysiological or electromagnetic activity is not a perfect measurement of neurocognitive function. M/EEG recordings, like any methodology, have limitations that must be considered.

Mixing in the temporal or spatial domain is a critical issue. Mixing refers to when multiple spatially overlapping populations contribute to the signal recorded at a single electrode. An example of mixing in the temporal domain that cannot be recovered through time–frequency analyses is illustrated in **Figure 5**. Another (extreme) example of mixing is if two populations of pyramidal cells are equally simultaneously active, but aligned in opposing orientation (e.g., on different sides of a sulcus). In this case, their electrical fields will cancel and the researcher may be left with the

misleading conclusion that no neural activity has occurred. Mixing is particularly problematic in EEG due to volume conduction and smearing/smoothing of the signal through the skull. Spatial filters such as current-source-density seem to be appropriate for obtaining relatively finer spatial resolution and linking inter-regional synchronization to cognitive processes (Srinivasan et al., 2007; Winter et al., 2007). Independent components analysis may recover some activity from mixed sources if those sources are temporally differentiable. Complex mixing from spatially overlapping and non-stationary sources may be less mathematically tractable to separate.

Another limitation of M/EEG is that they are limited to recording only certain kinds of activity from certain kinds of neurons (e.g., pyramidal and not interneuron) arranged in certain geometric orientations relative to the skull. Another theoretical limitation is that EEG measures only neural populations that are tangentially aligned to the skull, whereas MEG measures neural populations that are radially aligned to the skull. This is mostly a theoretical argument, however, because in practice many cognitive processes recruit areas of cortex that span gyri and sulci. Finally, due to the large number of dimensions available in EEG data and thus a huge number of possible statistical comparisons across time, space, frequency, and power/phase (and interactions among these dimensions such as inter-electrode cross-frequency coupling), there is a large potential for spurious Type I errors, particularly during exploratory analyses. There is a fine balance between, on the one hand, being driven by and constraining oneself to *a priori* hypotheses based on theory and previous research, and, on the other hand, being open to unexpected and unpredicted but robust patterns of results in the data.

Despite the limitations, examination of neural temporal dynamics has the potential to provide insight into human neurocognitive function beyond what is possible using approaches based on spatial localization (e.g., fMRI) or time-domain averages (e.g., ERP). Arguably, these limitations and considerations reinforce the idea that analyzing the rich temporal dynamics of neural activity bring us closer to the true complexity of brain function.

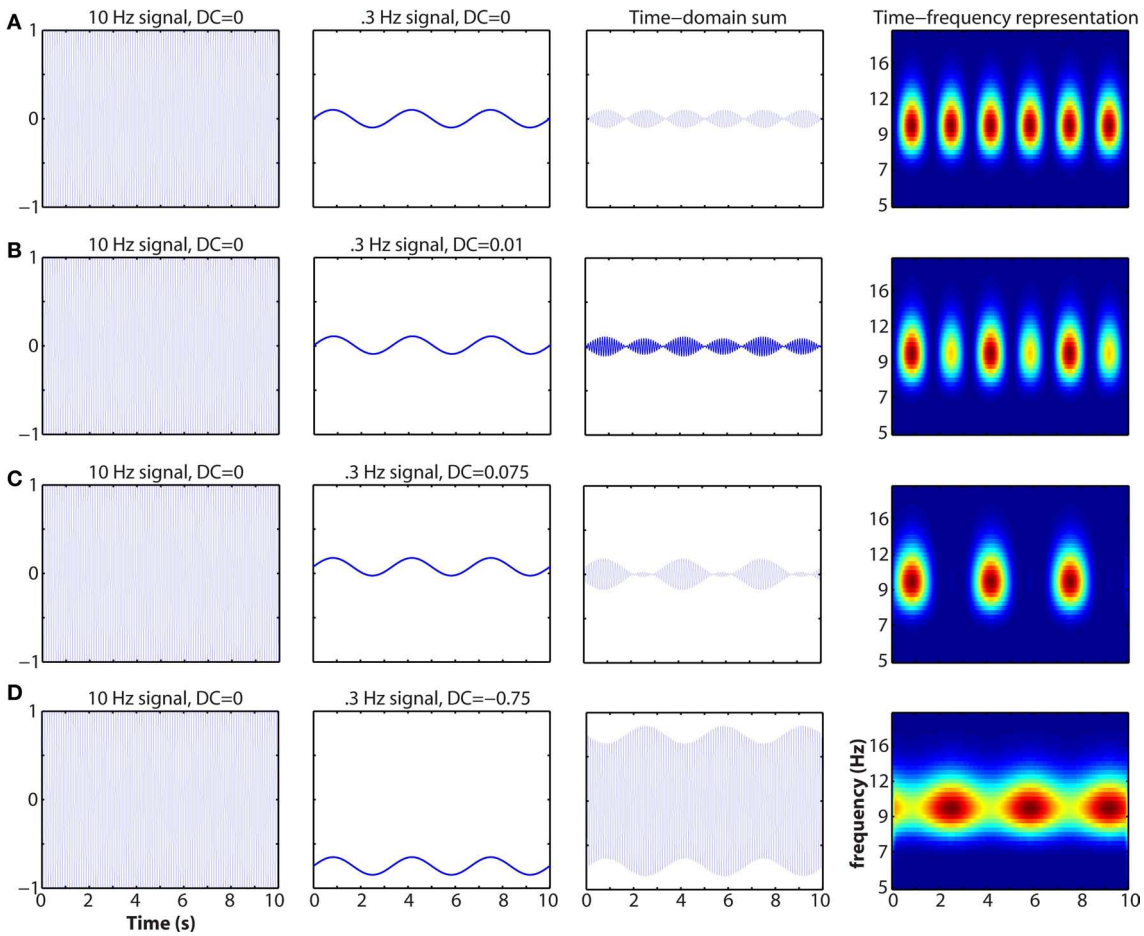



**FIGURE 5 | Example of how mixing in the time-domain can affect the time–frequency representation.** The activities of two spatially overlapping and similarly oriented neural networks (left two columns), one generating a 10-Hz rhythm and the other generating a 0.3-Hz rhythm, sum and are recorded by a single electrode (third column). At right is the time–frequency representation. Note that neither population on its own exhibits cross-frequency coupling. Oscillation baseline shifts and amplitude asymmetries have been described before and linked to cognitively relevant event-related potentials (Nikulin et al., 2007; Mazaheri and Jensen, 2008). This simulation demonstrates that the sensitivity of M/EEG measurements is not infallible, although they may still provide deeper insights into neurocognitive function compared to functional localization.

## WHAT ABOUT SPACE?

Accepting that information in the brain is coded precisely in time but distributed in space does not necessarily imply that space is irrelevant for neural representations and computations. Indeed, a logical consequent of this proposition is that space-based analyses should focus on distributed patterns rather than localization. Space and time may even have similar hierarchical computational organizational principles (Kiebel et al., 2008).

The best spatial resolution currently possible is a few cubic millimeters with high-resolution fMRI, although this resolution refers to the hemodynamic response, which may be spatially dispersed from its neural origin, following vascular features (Disbrow et al., 2000). Nonetheless, to the extent that representations extend over space at the level of several millimeters or centimeters, fMRI seems to be a valid tool for uncovering sparse or distributed representations. Spatial multivariate approaches that analyze patterns of activity over space (voxels) have sometimes proven more sensitive than standard approaches (i.e., testing whether the activity at all voxels is different from zero, or different between conditions) at linking brain states to cognitive states. On the one hand, this might be expected considering that multivariate regressions use more parameters to characterize the data – indeed, it remains to be discussed in the literature what the actual probability of Type I errors are and what preventative statistical measures are appropriate – but nonetheless the pattern of spatial activation sometimes predicts subjects' cognitive state better than the overall amount of activity averaged over space (Haynes and Rees, 2006; Norman et al., 2006). Still, many applications of spatial multivariate approaches continue to rest on functional localization assumptions, for example by considering multivariate patterns only from small clusters of contiguous voxels, and then moving this "spotlight" around the brain (Kriegeskorte et al., 2006), or by selecting voxels for multivariate analyses that exhibit significant modulation by condition in a standard localization-based general linear model (Norman et al., 2006).

Relatively low spatial resolution techniques like EEG and MEG can also be used to examine distributed spatial patterns of electrical activity. For example, spatial multivariate patterns in EEG have been used to dissociate neural computations of magnitude from valence in a feedback-driven learning task (Philiastides et al., 2010) and word categories (Chan et al., 2010). Multivariate pattern analyses may also be informative across frequencies in one brain region. For example, visual gradient orientation can be predicted from frequency multivariate patterns (Duncan et al., 2010), even though the neurons coding for directional gradients are at a spatial scale too small to be resolved by MEG.

Distributed, multivariate spatial analyses in fMRI vs. multivariate time–frequency analyses in M/EEG may provide complementary information: Whereas examining complex spatial patterns in fMRI data may be suitable for understanding how representations are "stored" or activated, examining complex temporal patterns in M/EEG data may be more amenable for understanding the operations/computations performed on those representations, and how those representations are shared or transferred across space and over time.

## OPEN QUESTIONS AND FUTURE DIRECTIONS

Following is a non-exhaustive list of important foci for future research on temporal coding and processing schemes in human neurocognitive function.

(1) *What is the neurobiological meaning of different features of oscillation dynamics*? The different features are power and phase within each frequency range, and all interactions they entail (e.g., power–power correlation, phase–phase synchronization, power–phase synchronization). Within each frequency band, estimates of power and phase are independent of each other (with the exception of zero power, in which case it is not possible to estimate phase, although in practice this is not often observed in real data at neurocognitively relevant frequencies). Sometimes, results obtained from power and from phase are convergent; other times, divergent. How do we interpret results from power, phase, phase-power coherence, phase–phase synchrony, etc., and what do they mean for network dynamics, brain function, and information processing transfer?

It is tempting to speculate that the number of simultaneously active neurons drives power whereas the timing of the activity of those neurons drives phase. However, this is likely overly simplistic. For example, both power and phase information can be used to predict spike-timing (Rasch et al., 2008). For human neuroscience, perhaps the best way to dissociate the roles of power and phase for neurocomputation may come from careful and clever experimental design in which different predictions are made for how measures of power vs. phase are related to different cognitive processes.

(2) *What spatial scales are relevant and how are dynamics at different spatial scales related?* Dynamic and oscillatory neural activity can be measured at a large range of spatial scales, from within a single neuron to populations of millions of neurons (Varela et al., 2001; Kiebel et al., 2008; Moran and Bar-Gad, 2010). What is the appropriate spatial scale for neurocognitive function? Are different scales more appropriate for different cognitive functions? Are multi-spatial-scale interactions relevant for cognition?

There have been few investigations into how electrophysiological measurements at different spatial scales are related to each other. For example, in a study investigating the relationship between single-/multi-unit activity and EEG in a monkey, even the best combination of EEG characteristics (in this case, delta phase and gamma power) recorded from a small electrode accounted for only ~15% of the variance of multi-unit activity (Whittingstall and Logothetis, 2009). Similarly, in an intracranial EEG study in humans, we found that time-domain correlation coefficients and theta-band phase synchrony between Cz and each intracranial electrode in the medial frontal cortex showed significant correlations/synchrony, but the magnitude was low (in the range of 0.1–0.2; Cohen et al., 2008). Thus, surface EEG may reflect a complex mixture of spatiotemporal dynamics from widespread areas. Whether and to what extent the divergence between activities recorded from multiple spatial scales is meaningful for cognitive function deserves more empirical attention.

(3) *How does anatomical connectivity shape functional connectivity*? Synchronous activity across widespread brain regions is believed to reflect functionally unified networks, such that physically separate neural ensembles are co-processing the same information or transferring information back and forth. Presumably, functional interactions – the nature and strength with which different nodes in a brain network communicate with each other – are shaped by the anatomical connections bridging those nodes. But what aspects of functional connectivity are shaped by anatomical connectivity: The strength of connectivity? Frequency range of synchronous interactions? Timing and phase delay? Which aspects of structure–function relationships are relevant for cognitive/behavioral functioning?

This question may be best addressed by linking EEG measures to white matter properties, measured through diffusion tensor imaging (DTI; Johansen-Berg and Rushworth, 2009), which takes advantage of the fact that the diffusion of water molecules in the brain is constrained by white matter fiber bundles. DTI data provides meaningful information about local white matter integrity and also the strength of tracts connecting different brain regions (Johansen-Berg, 2010). For example, visual stimulus-evoked gamma oscillations are correlated across subjects with corpus callosum white matter integrity (Zaehle and Herrmann, 2010). Similar findings have been observed with resting state EEG connectivity (Teipel et al., 2009) and medial frontal cortical responses to errors (Westlye et al., 2009).

(4) *Are oscillations causally involved in neurobiological phenomena*? Establishing causation is critical to science. To date, much of the current work on the role of oscillations in human neurocognitive function has been correlative. Although this is a necessary initial step, once spatial–temporal-frequency characteristics of neurocognitive processes are characterized, oscillation dynamics should be experimentally manipulated, ideally without explicitly manipulating the cognitive process thought to rely on those dynamics. There are several tools for addressing issues of causality, including transcranial magnetic stimulation, which has been shown to transiently perturb ongoing oscillations (Van Der Werf and Paus, 2006) that are dominant to each cortical region (Rosanova et al., 2009), and can impair cognitive processes such as attention that are thought to rely on specific oscillation patterns (Hamidi et al., 2009; Sauseng et al., 2009; Romei et al., 2010). Pharmacological manipulations may also be useful, although pharmacological agents may have complex effects on several brain systems and functions, so it may be difficult to interpret such results solely in the context of oscillations.

Oscillations can also be exogenously manipulated through stimulus flicker: When a visual stimulus is flashed at a particular frequency (like a strobe-light), regions of the brain that process that stimulus begin to oscillate at that frequency. In addition to "tagging" particular stimulus features (Herrmann, 2001; Ding et al., 2006), flicker has been shown to module – in task/frequency band-specific ways – attention and memory processes (Silberstein et al., 2001; Williams, 2001; Ellis et al., 2006; Williams et al., 2006; Wu and Yao, 2007; Mayes et al., 2009). It may also be possible to train subjects to modulate intrinsic oscillatory activity through "neurofeedback," which can modulate cognitive processes (Gruzelier et al., 2006; Keizer et al., 2010; Zoefel et al., 2010) as well as neuroplasticity and corticomuscular excitability (Ros et al., 2010).

(5) *Are all time-based coding/processing schemes oscillatory*? Time–frequency decomposition is often used because of visually observable oscillations in EEG data, the link to animal research examining local field potential oscillations (Buzsaki and Draguhn, 2004), the fact that time–frequency methods are becoming increasingly common in the field, and the continuous advances in computing power, which facilitate analyses. But is all (or even most) time-based information in the brain contained in oscillation dynamics?

In fact, in time–frequency decomposition analyses such as wavelet convolution or Fourier transform, oscillations themselves in different frequency bands are not directly measured; instead, what is measured is the extent to which the time-domain signal correlates with wavelets or sine waves at specific frequency bands with specific windows in time. Because any time-domain signal can be represented as a sum of sine waves of different phases, frequencies, and amplitudes, non-oscillatory responses will be captured by time–frequency decomposition. Yeung et al. (2004) attempted to use this feature of Fourier's theorem to argue that the error-related negativity may not be an oscillatory theta response, although their simulation of a "non-oscillatory response" was a half-sine wave at theta frequency (for more discussion, see Trujillo and Allen, 2007). Yeung et al.'s theoretical point is well taken, however, and there may be non-oscillatory dynamics that appear oscillatory due to time–frequency decomposition. Indeed, even in absence of sharp peaks in EEG power spectra over extended recording periods, frequency band-specific temporal dynamics are apparent (He et al., 2010), and can be characterized using 1/f functions (Miller et al., 2009). Broadband activity also seems to be relevant for some aspects of sensory–motor functioning (Onton and Makeig, 2009; Miller et al., 2010).

One could argue that whether the neural dynamics are truly oscillatory is not important; rather, what is important is that a time–frequency approach to analyzing electrophysiology data may provide new insights into neurocognitive function beyond what could be learned from simple time-domain averaging. However, whether the dynamics are truly oscillatory in nature might be relevant to linking human work to *in vivo* animal recordings, computational models, etc.

There are other ways in which information can be encoded in time that are not necessarily oscillatory. For example, there might be temporal "states" or patterns (Stam and van Dijk, 2002; Osterhage et al., 2007). Dynamics unfolding over time could be decoded using pattern-based analyses like support-vector machines or multivariate regressions, in which activity at different points in time are weighted to produce a linear or non-linear integration of activity over time (and/or frequency; Duncan et al., 2010) that best predicts the subjects' internal mental state or cognitive process.

(6) *What exactly is the information embedded in time*? If the brain uses time-based information coding and processing schemes, what is the nature of this information? Is there a "temporographic" code, such that the precise timing of activity contains information? Or is the "information" simply reflecting passive maintenance of space-based information? Or does it reflect the passive transmission of information from one population of cells to another? Perhaps the oscillations themselves do not code information *per se*, but rather act as an organizing filter or selector for different populations of cells that actually encode/process information through action potentials or other means.

## CONCLUSION

The stomatogastric ganglion is a nucleus in the crustacean stomach that comprises approximately 30 neurons and controls the stomach and parts of the digestive system. The complete circuitry and connectivity of this Lilliputian network is known. And yet, with two central pattern generators, modulation by at least 20 neurochemicals, multiple modes of firing ranging from synchronous bursting to arrhythmic, and pacemaker functions, there is an incomplete understanding as to how stomatogastric ganglion operates, changes neural states, and supports digestion (Hooper and DiCaprio, 2004; Stein, 2009). The human brain, in super-Brobdingnagian contrast, contains perhaps eight orders of magnitude more neurons, even more possible connectivity patterns among these neurons, and controls an impressive and complex array of perceptual, emotional, motor, linguistic, social, and creative processes. In trying to understand how dynamics of the brain lead to dynamics of behavior, it may seem we are doomed to remain prisoners in a platonic cave, forever straining to make out faint and fuzzy shadows cast by the most fantastically complex and enigmatic information processing system we know of.

Here it is argued that considerable neural information is embedded in the rich temporal landscape of electrophysiological dynamics, that much of this information may be lost when confining analyses to spatial dimensions, and that at least some of this information can be extracted non-invasively in humans using EEG and MEG. Approaches and analyses focused on temporal dynamical coding schemes will not render useless other approaches that are based on different (e.g., spatial) assumptions of neurocognitive function. However, ideas about time-based information coding schemes, and the approach of examining the temporal dynamics of brain electrical activity, are an important next step in theoretical and empirical human neuroscience developments. This nascent but growing literature on human neural temporal dynamics will provide a new impetus in uncovering fundamental neurocognitive mechanisms, linking research in humans to that in animals, and improving clinical diagnosis and treatment assessment. It's about time.

## ACKNOWLEDGMENTS

Several of the ideas presented here were outlined at a workshop entitled "Opinions and Discussions in Cognitive Neuroscience: Amsterdam" in October 2009. The author is supported by a VIDI grant from the Dutch Organization for Scientific Research (NWO). Thanks to Dr. Sanne de Wit for **Figure 1**, and to Romke Rouw for comments.

## REFERENCES

Akam, T., and Kullmann, D. M. (2010). Oscillations and filtering networks support flexible routing of information. *Neuron* 67, 308–320.

Allen, J. J., and Cohen, M. X. (2010). Deconstructing the "resting" state: exploring the temporal dynamics of frontal alpha asymmetry as an endophenotype for depression. *Front. Hum. Neurosci.* 4:232. doi: 10.3389/fnhum.2010.00232

Aronov, D., Reich, D. S., Mechler, F., and Victor, J. D. (2003). Neural coding of spatial phase in V1 of the macaque monkey. *J. Neurophysiol.* 89, 3304–3327.

Axmacher, N., Elger, C. E., and Fell, J. (2008). Ripples in the medial temporal lobe are relevant for human memory consolidation. *Brain* 131, 1806–1817.

Axmacher, N., Henseler, M. M., Jensen, O., Weinreich, I., Elger, C. E., and Fell, J. (2010). Cross-frequency coupling supports multi-item working memory in the human hippocampus. *Proc. Natl. Acad. Sci. U.S.A.* 107, 3228–3233.

Bassett, D. S., and Bullmore, E. T. (2009). Human brain networks in health and disease. *Curr. Opin. Neurol.* 22, 340–347.

Bodelon, C., Fallah, M., and Reynolds, J. H. (2007). Temporal resolution for the perception of features and conjunctions. *J. Neurosci.* 27, 725–730.

Botvinick, M. M. (2007). Conflict monitoring and decision making: reconciling two perspectives on anterior cingulate function. *Cogn. Affect. Behav. Neurosci.* 7, 356–366.

Britz, J., and Michel, C. M. (2010). Errors can be related to pre-stimulus differences in ERP topography and their concomitant sources. *Neuroimage* 49, 2774–2782.

Britz, J., Van De Ville, D., and Michel, C. M. (2010). BOLD correlates of EEG topography reveal rapid resting-state network dynamics. *Neuroimage* 52, 1162–1170.

Bruns, A. (2004). Fourier-, Hilbert- and wavelet-based signal analysis: are they really different approaches? *J. Neurosci. Methods.* 137, 321–332.

Bullmore, E., and Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat. Rev. Neurosci.* 10, 186–198.

Busch, N. A., and VanRullen, R. (2010). Spontaneous EEG oscillations reveal periodic sampling of visual attention. *Proc. Natl. Acad. Sci. U.S.A.* 107, 16048–16053.

Buzsaki, G., and Draguhn, A. (2004). Neuronal oscillations in cortical networks. *Science* 304, 1926–1929.

Canolty, R. T., Edwards, E., Dalal, S. S., Soltani, M., Nagarajan, S. S., Kirsch, H. E., Berger, M. S., Barbaro, N. M., and Knight, R. T. (2006). High gamma power is phase-locked to theta oscillations in human neocortex. *Science* 313, 1626–1628.

Carter, C. S., and van Veen, V. (2007). Anterior cingulate cortex and conflict detection: an update of theory and data. *Cogn. Affect. Behav. Neurosci.* 7, 367–379.

Cavanagh, J. F., Cohen, M. X., and Allen, J. J. (2009). Prelude to and resolution of an error: EEG phase synchrony reveals cognitive control dynamics during action monitoring. *J. Neurosci.* 29, 98–105.

Cavanagh, J. F., Frank, M. J., Klein, T. J., and Allen, J. J. (2010). Frontal theta links prediction errors to behavioral adaptation in reinforcement learning. *Neuroimage* 49, 3198–3209.

Chan, A. M., Halgren, E., Marinkovic, K., and Cash, S. S. (2010). Decoding word and category-specific spatiotemporal representations from MEG and EEG. *Neuroimage.* doi:10.1016/j.neuroimage.2010.10.073. [Epub ahead of print].

Cohen, M. X. (2008). Assessing transient cross-frequency coupling in EEG data. *J. Neurosci. Methods* 168, 494–499.

Cohen, M. X., Axmacher, N., Lenartz, D., Elger, C. E., Sturm, V., and Schlaepfer, T. E. (2009a). Good vibrations: cross-frequency coupling in the human nucleus accumbens during reward processing. *J. Cogn. Neurosci.* 21, 875–889.

Cohen, M. X., Axmacher, N., Lenartz, D., Elger, C. E., Sturm, V., and Schlaepfer, T. E. (2009b). Nuclei accumbens phase synchrony predicts decision-making reversals following negative feedback. *J. Neurosci.* 29, 7591–7598.

Cohen, M. X., Elger, C. E., and Fell, J. (2009c). Oscillatory activity and phase-amplitude coupling in the human medial frontal cortex during decision making. *J. Cogn. Neurosci.* 21, 390–402.

Cohen, M. X., van Gaal, S., Ridderinkhof, K. R., and Lamme, V. A. (2009d). Unconscious errors enhance prefrontal–occipital oscillatory synchrony. *Front. Hum. Neurosci.* 3:54. doi: 10.3389/neuro.09.054.2009

Cohen, M. X., Bour, L., Mantione, M., Figee, M., Vink, M., Tijssen, A. J., Rootselaar, F., Munckhof, P., Schuurman, P. R., and Denys, D.

(2011). Top-down directed synchrony from medial frontal cortex to nucleus accumbens during reward anticipation. *Hum. Brain. Mapp.* (in press).

Cohen, M. X., and Ranganath, C. (2007). Reinforcement learning signals predict future decisions. *J. Neurosci.* 27, 371–378.

Cohen, M. X., Ridderinkhof, K. R., Haupt, S., Elger, C. E., and Fell, J. (2008). Medial frontal cortex and response conflict: evidence from human intracranial EEG and medial frontal cortex lesion. *Brain Res.* 1238, 127–142.

Coull, J. T. (2004). fMRI studies of temporal attention: allocating attention within, or towards, time. *Brain Res. Cogn. Brain Res.* 21, 216–226.

Coull, J. T., Vidal, F., Nazarian, B., and Macar, F. (2004). Functional anatomy of the attentional modulation of time estimation. *Science* 303, 1506–1508.

Cui, J., Xu, L., Bressler, S. L., Ding, M., and Liang, H. (2008). BSMART: a Matlab/C toolbox for analysis of multichannel neural time series. *Neural. Netw.* 21, 1094–1104.

Curio, G. (2000). Linking 600-Hz "spike-like" EEG/MEG wavelets "sigma-bursts" to cellular substrates: concepts and caveats. *J. Clin. Neurophysiol.* 17, 377–396.

De Gennaro, L., Vecchio, F., Ferrara, M., Curcio, G., Rossini, P. M., and Babiloni, C. (2004). Changes in fronto-posterior functional coupling at sleep onset in humans. *J. Sleep Res.* 13, 209–217.

De Martino, F., de Borst, A. W., Valente, G., Goebel, R., and Formisano, E. (2010). Predicting EEG single trial responses with simultaneous fMRI and relevance vector machine regression. *Neuroimage*. doi: 10.1016/j.neuroimage.2010.07.068. [Epub ahead of print].

Debener, S., Ullsperger, M., Siegel, M., and Engel, A. K. (2007). Towards single-trial analysis in cognitive brain research. *Trends Cogn. Sci.* 11, 502–503.

Debener, S., Ullsperger, M., Siegel, M., Fiehler, K., von Cramon, D. Y., and Engel, A. K. (2005). Trial-by-trial coupling of concurrent electroencephalogram and functional magnetic resonance imaging identifies the dynamics of performance monitoring. *J. Neurosci.* 25, 11730–11737.

Demiralp, T., Bayraktaroglu, Z., Lenz, D., Junge, S., Busch, N. A., Maess, B., Ergen, M., and Herrmann, C. S. (2007). Gamma amplitudes are coupled to theta phase in human EEG during visual perception. *Int. J. Psychophysiol.* 64, 24–30.

Ding, J., Sperling, G., and Srinivasan, R. (2006). Attentional modulation of SSVEP power depends on the network tagged by the flicker frequency. *Cereb. Cortex* 16, 1016–1029.

Disbrow, E. A., Slutsky, D. A., Roberts, T. P., and Krubitzer, L. A. (2000). Functional MRI at 1.5 tesla: a comparison of the blood oxygenation level-dependent signal and electrophysiology. *Proc. Natl. Acad. Sci. U.S.A.* 97, 9718–9723.

Duncan, K. K., Hadjipapas, A., Li, S., Kourtzi, Z., Bagshaw, A., and Barnes, G. (2010). Identifying spatially overlapping local cortical networks with MEG. *Hum. Brain Mapp.* 31, 1003–1016.

Eichele, H., Juvodden, H. T., Ullsperger, M., and Eichele, T. (2010). Mal-adaptation of event-related EEG responses preceding performance errors. *Front. Hum. Neurosci.* 4:65. doi: 10.3389/fnhum.2010.00065

Ellis, K. A., Silberstein, R. B., and Nathan, P. J. (2006). Exploring the temporal dynamics of the spatial working memory n-back task using steady state visual evoked potentials (SSVEP). *Neuroimage* 31, 1741–1751.

Engel, A. K., Fries, P., and Singer, W. (2001). Dynamic predictions: oscillations and synchrony in top-down processing. *Nat. Rev. Neurosci.* 2, 704–716.

Engel, A. K., Roelfsema, P. R., Fries, P., Brecht, M., and Singer, W. (1997). Role of the temporal domain for response selection and perceptual binding. *Cereb. Cortex* 7, 571–582.

Fries, P., Nikolic, D., and Singer, W. (2007). The gamma cycle. *Trends Neurosci.* 30, 309–316.

Fujisawa, S., Amarasingham, A., Harrison, M. T., and Buzsaki, G. (2008). Behavior-dependent short-term assembly dynamics in the medial prefrontal cortex. *Nat. Neurosci.* 11, 823–833.

Fuster, J. M. (2001). The prefrontal cortex – an update: time is of the essence. *Neuron* 30, 319–333.

Gaillard, R., Dehaene, S., Adam, C., Clemenceau, S., Hasboun, D., Baulac, M., Cohen, L., and Naccache, L. (2009). Converging intracranial markers of conscious access. *PLoS Biol.* 7, e61. doi: 10.1371/journal.pbio.1000061

Garrido, M. I., Kilner, J. M., Stephan, K. E., and Friston, K. J. (2009). The mismatch negativity: a review of underlying mechanisms. *Clin. Neurophysiol.* 120, 453–463.

Gauthier, I., Skudlarski, P., Gore, J. C., and Anderson, A. W. (2000a). Expertise for cars and birds recruits brain areas involved in face recognition. *Nat. Neurosci.* 3, 191–197.

Gauthier, I., Tarr, M. J., Moylan, J., Skudlarski, P., Gore, J. C., and Anderson, A. W. (2000b). The fusiform "face area" is part of a network that processes faces at the individual level. *J. Cogn. Neurosci.* 12, 495–504.

Gehring, W. J., and Willoughby, A. R. (2002). The medial frontal cortex and the rapid processing of monetary gains and losses. *Science* 295, 2279–2282.

Gladwin, T. E., Lindsen, J. P., and de Jong, R. (2006). Pre-stimulus EEG effects related to response speed, task switching and upcoming response hand. *Biol. Psychol.* 72, 15–34.

Goulet, S., Dore, F. Y., and Murray, E. A. (1998). Aspiration lesions of the amygdala disrupt the rhinal corticothalamic projection system in rhesus monkeys. *Exp. Brain Res.* 119, 131–140.

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* 37, 424–438.

Griesmayr, B., Gruber, W. R., Klimesch, W., and Sauseng, P. (2010). Human frontal midline theta and its synchronization to gamma during a verbal delayed match to sample task. *Neurobiol. Learn. Mem.* 93, 208–215.

Gruzelier, J., Egner, T., and Vernon, D. (2006). Validating the efficacy of neurofeedback for optimising performance. *Prog. Brain Res.* 159, 421–431.

Guderian, S., Schott, B. H., Richardson-Klavehn, A., and Duzel, E. (2009). Medial temporal theta state before an event predicts episodic encoding success in humans. *Proc. Natl. Acad. Sci. U.S.A.* 106, 5365–5370.

Gutierrez-Galvez, A., and Gutierrez-Osuna, R. (2003). Pattern completion through phase coding in population neurodynamics. *Neural. Netw.* 16, 649–656.

Guye, M., Bartolomei, F., and Ranjeva, J. P. (2008). Imaging structural and functional connectivity: towards a unified definition of human brain organization? *Curr. Opin. Neurol.* 21, 393–403.

Hamidi, M., Slagter, H. A., Tononi, G., and Postle, B. R. (2009). Repetitive transcranial magnetic stimulation affects behavior by biasing endogenous cortical oscillations. *Front. Integr. Neurosci.* 3:14. doi: 10.3389/neuro.07.014.2009

Hanslmayr, S., Pastotter, B., Bauml, K. H., Gruber, S., Wimber, M., and Klimesch, W. (2008). The electrophysiological dynamics of interference during the Stroop task. *J. Cogn. Neurosci.* 20, 215–225.

Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., and Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 293, 2425–2430.

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.

Haynes, J. D., and Rees, G. (2006). Decoding mental states from brain activity in humans. *Nat. Rev. Neurosci.* 7, 523–534.

He, B. J., Zempel, J. M., Snyder, A. Z., and Raichle, M. E. (2010). The temporal structures and functional significance of scale-free brain activity. *Neuron* 66, 353–369.

Heil, P. (2004). First-spike latency of auditory neurons revisited. *Curr. Opin. Neurobiol.* 14, 461–467.

Heinzle, J., Konig, P., and Salazar, R. F. (2007). Modulation of synchrony without changes in firing rates. *Cogn. Neurodyn.* 1, 225–235.

Herrmann, C. S. (2001). Human EEG responses to 1–100 Hz flicker: resonance phenomena in visual cortex and their potential correlation to cognitive phenomena. *Exp. Brain Res.* 137, 346–353.

Holldobler, B., and Wilson, E. O. (1998). *Journey to the Ants*. Cambridge: Belknap Press of Harvard University Press.

Holroyd, C. B., and Coles, M. G. (2002). The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychol. Rev.* 109, 679–709.

Hooper, S. L., and DiCaprio, R. A. (2004). Crustacean motor pattern generator networks. *Neurosignals* 13, 50–69.

Huxter, J., Burgess, N., and O'Keefe, J. (2003). Independent rate and temporal coding in hippocampal pyramidal cells. *Nature* 425, 828–832.

Ivry, R. B. (1996). The representation of temporal information in perception and motor control. *Curr. Opin. Neurobiol.* 6, 851–857.

Ivry, R. B., and Schlerf, J. E. (2008). Dedicated and intrinsic models of time perception. *Trends Cogn. Sci.* 12, 273–280.

Jacobs, J., Kahana, M. J., Ekstrom, A. D., and Fried, I. (2007). Brain oscillations control timing of single-neuron activity in humans. *J. Neurosci.* 27, 3839–3844.

Jensen, O. (2001). Information transfer between rhythmically coupled networks: reading the hippocampal phase code. *Neural. Comput.* 13, 2743–2761.

Jensen, O., and Colgin, L. L. (2007). Cross-frequency coupling between neuronal oscillations. *Trends Cogn. Sci.* 11, 267–269.

Jensen, O., and Lisman, J. E. (2000). Position reconstruction from an ensemble of hippocampal place cells: contribution of theta phase coding. *J. Neurophysiol.* 83, 2602–2609.

Johansen-Berg, H. (2010). Behavioural relevance of variation in white matter microstructure. *Curr. Opin. Neurol.* 23, 351–358.

Johansen-Berg, H., and Rushworth, M. F. (2009). Using diffusion imaging to study human connectional anatomy. *Annu. Rev. Neurosci.* 32, 75–94.

Kaminski, M., Blinowska, K., and Szclenberger, W. (1997). Topographic analysis of coherence and propagation of EEG activity during sleep and wakefulness. *Electroencephalogr. Clin. Neurophysiol.* 102, 216–227.

Kanwisher, N. (2000). Domain specificity in face perception. *Nat. Neurosci.* 3, 759–763.

Kanwisher, N. (2010). Functional specificity in the human brain: a window into the functional architecture of the mind. *Proc. Natl. Acad. Sci. U.S.A.* 107, 11163–11170.

Kanwisher, N., McDermott, J., and Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J. Neurosci.* 17, 4302–4311.

Kayser, J., and Tenke, C. E. (2006). Principal components analysis of Laplacian waveforms as a generic method for identifying ERP generator patterns: I. Evaluation with auditory oddball tasks. *Clin. Neurophysiol.* 117, 348–368.

Keizer, A. W., Verment, R. S., and Hommel, B. (2010). Enhancing cognitive control through neurofeedback: a role of gamma-band activity in managing episodic retrieval. *Neuroimage* 49, 3404–3413.

Kerns, J. G., Cohen, J. D., MacDonald, A. W. III, Cho, R. Y., Stenger, V. A., and Carter, C. S. (2004). Anterior cingulate conflict monitoring and adjustments in control. *Science* 303, 1023–1026.

Kiebel, S. J., Daunizeau, J., and Friston, K. J. (2008). A hierarchy of time-scales and the brain. *PLoS Comput. Biol.* 4, e1000209. doi: 10.1371/journal.pcbi.1000209

Klimesch, W., Sauseng, P., and Hanslmayr, S. (2007). EEG alpha oscillations: the inhibition-timing hypothesis. *Brain Res. Rev.* 53, 63–88.

Knight, R. T. (2007). Neuroscience. Neural networks debunk phrenology. *Science* 316, 1578–1579.

Knoblauch, A., and Palm, G. (2001). Pattern separation and synchronization in spiking associative memories and visual areas. *Neural. Netw.* 14, 763–780.

Konig, P., Engel, A. K., Roelfsema, P. R., and Singer, W. (1995). How precise is neuronal synchronization? *Neural Comput.* 7, 469–485.

Kriegeskorte, N., Goebel, R., and Bandettini, P. (2006). Information-based functional brain mapping. *Proc. Natl. Acad. Sci. U.S.A.* 103, 3863–3868.

Kujala, T., and Naatanen, R. (2010). The adaptive brain: a neurophysiological perspective. *Prog. Neurobiol.* 91, 55–67.

Lehmann, D., Faber, P. L., Gianotti, L. R., Kochi, K., and Pascual-Marqui, R. D. (2006). Coherence and phase locking in the scalp EEG and between LORETA model sources, and microstates as putative mechanisms of brain temporo-spatial functional organization. *J. Physiol. Paris* 99, 29–36.

Lisman, J. (2005). The theta/gamma discrete phase code occurring during the hippocampal phase precession may be a more general brain coding scheme. *Hippocampus* 15, 913–922.

Lisman, J. (2010). Working memory: the importance of theta and gamma oscillations. *Curr. Biol.* 20, R490–R492.

Luck, S. J. (2005). *An Introduction to the Event-Related Potential Technique*. Cambridge: MIT Press.

Makeig, S., Debener, S., Onton, J., and Delorme, A. (2004). Mining event-related brain dynamics. *Trends Cogn. Sci.* 8, 204–210.

Makeig, S., Jung, T. P., Bell, A. J., Ghahremani, D., and Sejnowski, T. J. (1997). Blind separation of auditory event-related brain responses into independent components. *Proc. Natl. Acad. Sci. U.S.A.* 94, 10979–10984.

Mars, R. B., Debener, S., Gladwin, T. E., Harrison, L. M., Haggard, P., Rothwell, J. C., and Bestmann, S. (2008). Trial-by-trial fluctuations in the event-related electroencephalogram reflect dynamic changes in the degree of surprise. *J. Neurosci.* 28, 12539–12545.

Mathewson, K. E., Gratton, G., Fabiani, M., Beck, D. M., and Ro, T. (2009). To see or not to see: prestimulus alpha phase predicts visual awareness. *J. Neurosci.* 29, 2725–2732.

Mayes, A. K., Pipingas, A., Silberstein, R. B., and Johnston, P. (2009). Steady state visually evoked potential correlates of static and dynamic emotional face processing. *Brain Topogr.* 22, 145–157.

Mazaheri, A., and Jensen, O. (2008). Asymmetric amplitude modulations of brain oscillations generate slow evoked responses. *J. Neurosci.* 28, 7781–7787.

Mazaheri, A., Nieuwenhuis, I. L., van Dijk, H., and Jensen, O. (2009). Prestimulus alpha and mu activity predicts failure to inhibit motor responses. *Hum. Brain Mapp.* 30, 1791–1800.

Melloni, L., Molina, C., Pena, M., Torres, D., Singer, W., and Rodriguez, E. (2007). Synchronization of neural activity across cortical areas correlates with conscious perception. *J. Neurosci.* 27, 2858–2865.

Miller, K. J., Hermes, D., Honey, C. J., Sharma, M., Rao, R. P., den Nijs, M., Fetz, E. E., Sejnowski, T. J., Hebb, A. O., Ojemann, J. G., Makeig, S., and Leuthardt, E. C. (2010). Dynamic modulation of local population activity by rhythm phase in human occipital cortex during a visual search task. *Front. Hum. Neurosci.* 4:197. doi: 10.3389/fnhum.2010.00197

Miller, K. J., Sorensen, L. B., Ojemann, J. G., and den Nijs, M. (2009). Power-law scaling in the brain surface electric potential. *PLoS Comput. Biol.* 5, e1000609. doi: 10.1371/journal.pcbi.1000609

Moran, A., and Bar-Gad, I. (2010). Revealing neuronal functional organization through the relation between multi-scale oscillatory extracellular signals. *J. Neurosci. Methods* 186, 116–129.

Mormann, F., Fell, J., Axmacher, N., Weber, B., Lehnertz, K., Elger, C. E., and Fernandez, G. (2005). Phase/amplitude reset and theta–gamma interaction in the human medial temporal lobe during a continuous word recognition memory task. *Hippocampus* 15, 890–900.

Muller, T. J., Koenig, T., Wackermann, J., Kalus, P., Fallgatter, A., Strik, W., and Lehmann, D. (2005). Subsecond changes of global brain state in illusory multistable motion perception. *J. Neural Transm.* 112, 565–576.

Musso, F., Brinkmeyer, J., Mobascher, A., Warbrick, T., and Winterer, G. (2010). Spontaneous brain activity and EEG microstates. A novel EEG/fMRI analysis approach to explore resting-state networks. *Neuroimage* 52, 1149–1161.

Nikulin, V. V., Linkenkaer-Hansen, K., Nolte, G., Lemm, S., Muller, K. R., Ilmoniemi, R. J., and Curio, G. (2007). A novel mechanism for evoked responses in the human brain. *Eur. J. Neurosci.* 25, 3146–3154.

Nobre, A., Correa, A., and Coull, J. (2007). The hazards of time. *Curr. Opin. Neurobiol.* 17, 465–470.

Norman, K. A., Polyn, S. M., Detre, G. J., and Haxby, J. V. (2006). Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends Cogn. Sci.* 10, 424–430.

O'Connell, R. G., Dockree, P. M., Robertson, I. H., Bellgrove, M. A., Foxe, J. J., and Kelly, S. P. (2009). Uncovering the neural signature of lapsing attention: electrophysiological signals predict errors up to 20 s before they occur. *J. Neurosci.* 29, 8604–8611.

Onton, J., and Makeig, S. (2009). High-frequency broadband modulations of electroencephalographic spectra. *Front. Hum. Neurosci.* 3:61. doi: 10.3389/neuro.09.061.2009

Osipova, D., Hermes, D., and Jensen, O. (2008). Gamma power is phase-locked to posterior alpha activity. *PLoS ONE* 3, e3990. doi: 10.1371/journal.pone.0003990

Osterhage, H., Mormann, F., Wagner, T., and Lehnertz, K. (2007). Measuring the directionality of coupling: phase versus state space dynamics and application to EEG time series. *Int. J. Neural Syst.* 17, 139–148.

Overton, F. (1897). *The Brain. Applied Physiology*. New York: American Book Co., 432.

Padilla, M. L., Wood, R. A., Hale, L. A., and Knight, R. T. (2006). Lapses in a prefrontal-extrastriate preparatory attention network predict mistakes. *J. Cogn. Neurosci.* 18, 1477–1487.

Palva, J. M., Monto, S., Kulashekhar, S., and Palva, S. (2010). Neuronal synchrony reveals working memory networks and predicts individual memory capacity. *Proc. Natl. Acad. Sci. U.S.A.* 107, 7580–7585.

Philiastides, M. G., Biele, G., Vavatzanidis, N., Kazzer, P., and Heekeren, H. R. (2010). Temporal dynamics of prediction error processing during reward-based decision making. *Neuroimage* 53, 221–232.

Price, C. J., and Friston, K. J. (2005). Functional ontologies for cognition: the systematic definition of structure and function. *Cogn. Neuropsychol.* 22, 262–275.

Puce, A., Allison, T., Asgari, M., Gore, J. C., and McCarthy, G. (1996). Differential sensitivity of human visual cortex to faces, letterstrings, and textures: a functional magnetic resonance imaging study. *J. Neurosci.* 16, 5205–5215.

Quian Quiroga, R., Kraskov, A., Kreuz, T., and Grassberger, P. (2002). Performance of different synchronization measures in real data: a case study on electroencephalographic signals. *Phys. Rev. E. Stat. Nonlin. Soft Matter Phys.* 65, 041903.

Quian Quiroga, R., and Panzeri, S. (2009). Extracting information from neuronal populations: information theory and decoding approaches. *Nat. Rev. Neurosci.* 10, 173–185.

Rajagovindan, R., and Ding, M. (2008). Decomposing neural synchrony: toward an explanation for near-zero phase-lag in cortical oscillatory networks. *PLoS ONE* 3, e3649. doi: 10.1371/journal.pone.0003649

Rasch, M. J., Gretton, A., Murayama, Y., Maass, W., and Logothetis, N. K. (2008). Inferring spike trains from local field potentials. *J. Neurophysiol.* 99, 1461–1476.

Rhodes, G., Byatt, G., Michie, P. T., and Puce, A. (2004). Is the fusiform face area specialized for faces, individuation, or expert individuation? *J. Cogn. Neurosci.* 16, 189–203.

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

Rockland, K. S. (2010). Five points on columns. *Front. Neuroanat.* 4:22. doi: 10.3389/fnana.2010.00022

Romei, V., Gross, J., and Thut, G. (2010). On the role of prestimulus alpha rhythms over occipito-parietal areas in visual input regulation: correlation or causation? *J. Neurosci.* 30, 8692–8697.

Ros, T., Munneke, M. A., Ruge, D., Gruzelier, J. H., and Rothwell, J. C. (2010). Endogenous control of waking brain rhythms induces neuroplasticity in humans. *Eur. J. Neurosci.* 31, 770–778.

Rosanova, M., Casali, A., Bellina, V., Resta, F., Mariotti, M., and Massimini, M. (2009). Natural frequencies of human corticothalamic circuits. *J. Neurosci.* 29, 7679–7685.

Rousselet, G. A., Pernet, C. R., Bennett, P. J., and Sekuler, A. B. (2008). Parametric study of EEG sensitivity to phase noise during face processing. *BMC Neurosci.* 9, 98. doi: 10.1186/1471-2202-9-98

Rudebeck, P. H., Behrens, T. E., Kennerley, S. W., Baxter, M. G., Buckley, M. J., Walton, M. E., and Rushworth, M. F. (2008). Frontal cortex subregions play distinct roles in choices between actions and stimuli. *J. Neurosci.* 28, 13775–13785.

Rushworth, M. F., and Behrens, T. E. (2008). Choice, uncertainty and value in prefrontal and cingulate cortex. *Nat. Neurosci.* 11, 389–397.

Sanes, J. N., Suner, S., Lando, J. F., and Donoghue, J. P. (1988). Rapid reorganization of adult rat motor cortex somatic representation patterns after motor nerve injury. *Proc. Natl. Acad. Sci. U.S.A.* 85, 2003–2007.

Sauseng, P., Klimesch, W., Gruber, W. R., and Birbaumer, N. (2008). Cross-frequency phase synchronization: a brain mechanism of memory matching and attention. *Neuroimage* 40, 308–317.

Sauseng, P., Klimesch, W., Heise, K. F., Gruber, W. R., Holz, E., Karim, A. A., Glennon, M., Gerloff, C., Birbaumer, N., and Hummel, F. C. (2009). Brain oscillatory substrates of visual short-term memory capacity. *Curr. Biol.* 19, 1846–1852.

Schieber, M. H. (2001). Constraints on somatotopic organization in the primary motor cortex. *J. Neurophysiol.* 86, 2125–2143.

Sergent, J., Ohta, S., and MacDonald, B. (1992). Functional neuroanatomy of face and object processing. A positron emission tomography study. *Brain* 115(Pt 1), 15–36.

Silberstein, R. B., Nunez, P. L., Pipingas, A., Harris, P., and Danieli, F. (2001). Steady state visually evoked potential (SSVEP) topography in a graded working memory task. *Int. J. Psychophysiol.* 42, 219–232.

Singer, W. (1999). Neuronal synchrony: a versatile code for the definition of relations? *Neuron* 24, 49–65, 111–125.

Sporns, O. (2010). *Networks of the Brain.* Cambridge: MIT Press.

Srinivasan, R., Winter, W. R., Ding, J., and Nunez, P. L. (2007). EEG and MEG coherence: measures of functional connectivity at distinct spatial scales of neocortical dynamics. *J. Neurosci. Methods* 166, 41–52.

Stam, C. J. (2010). Use of magnetoencephalography (MEG) to study functional brain networks in neurodegenerative disorders. *J. Neurol. Sci.* 289, 128–134.

Stam, C. J., and van Dijk, B. W. (2002). Synchronization likelihood: an unbiased measure of generalized synchronization in multivariate data sets. *Physica D* 164, 236.

Stanford, T. R., Shankar, S., Massoglia, D. P., Costello, M. G., and Salinas, E. (2010). Perceptual decision making in less than 30 milliseconds. *Nat. Neurosci.* 13, 379–385.

Stein, W. (2009). Modulation of stomatogastric rhythms. *J. Comp. Physiol. A Neuroethol. Sens. Neural. Behav. Physiol.* 195, 989–1009.

Steriade, M. (2006). Grouping of brain rhythms in corticothalamic systems. *Neuroscience* 137, 1087–1106.

Tarr, M. J., and Gauthier, I. (2000). FFA: a flexible fusiform area for subordinate-level visual processing automatized by expertise. *Nat. Neurosci.* 3, 764–769.

Taylor, P. C., Nobre, A. C., and Rushworth, M. F. (2007). Subsecond changes in top down control exerted by human medial frontal cortex during conflict and action selection: a combined transcranial magnetic stimulation electroencephalography study. *J. Neurosci.* 27, 11343–11353.

Teipel, S. J., Pogarell, O., Meindl, T., Dietrich, O., Sydykova, D., Hunklinger, U., Georgii, B., Mulert, C., Reiser, M. F., Moller, H. J., and Hampel, H. (2009). Regional networks underlying interhemispheric connectivity: an EEG and DTI study in healthy ageing and amnestic mild cognitive impairment. *Hum. Brain Mapp.* 30, 2098–2119.

Tiesinga, P., Fellous, J. M., and Sejnowski, T. J. (2008). Regulation of spike timing in visual cortical circuits. *Nat. Rev. Neurosci.* 9, 97–107.

Tort, A. B., Komorowski, R., Eichenbaum, H., and Kopell, N. (2010). Measuring phase-amplitude coupling between neuronal oscillations of different frequencies. *J. Neurophysiol.* 104, 1195–1210.

Traub, R. D., Bibbig, A., LeBeau, F. E., Buhl, E. H., and Whittington, M. A. (2004). Cellular mechanisms of neuronal population oscillations in the hippocampus in vitro. *Annu. Rev. Neurosci.* 27, 247–278.

Trujillo, L. T., and Allen, J. J. (2007). Theta EEG dynamics of the error-related negativity. *Clin. Neurophysiol.* 118, 645–668.

Uhlhaas, P. J., and Singer, W. (2006). Neural synchrony in brain disorders: relevance for cognitive dysfunctions and pathophysiology. *Neuron* 52, 155–168.

Van Der Werf, Y. D., and Paus, T. (2006). The neural response to transcranial magnetic stimulation of the human motor cortex. I. Intracortical and cortico-cortical contributions. *Exp. Brain Res.* 175, 231–245.

van Gaal, S., Ridderinkhof, K. R., Fahrenfort, J. J., Scholte, H. S., and Lamme, V. A. (2008). Frontal cortex mediates unconsciously triggered inhibitory control. *J. Neurosci.* 28, 8053–8062.

van Veen, V., and Carter, C. S. (2002). The anterior cingulate as a conflict monitor: fMRI and ERP studies. *Physiol. Behav.* 77, 477–482.

Varela, F., Lachaux, J. P., Rodriguez, E., and Martinerie, J. (2001). The brainweb: phase synchronization and large-scale integration. *Nat. Rev. Neurosci.* 2, 229–239.

Vicente, R., Gollo, L. L., Mirasso, C. R., Fischer, I., and Pipa, G. (2008). Dynamical relaying can yield zero time lag neuronal synchrony despite long conduction delays. *Proc. Natl. Acad. Sci. U.S.A.* 105, 17157–17162.

Wang, X. J. (2010). Neurophysiological and computational principles of cortical rhythms in cognition. *Physiol. Rev.* 90, 1195–1268.

Westlye, L. T., Walhovd, K. B., Bjornerud, A., Due-Tonnessen, P., and Fjell, A. M. (2009). Error-related negativity is mediated by fractional anisotropy in the posterior cingulate gyrus – a study combining diffusion tensor imaging and electrophysiology in healthy adults. *Cereb. Cortex* 19, 293–304.

Whittingstall, K., and Logothetis, N. K. (2009). Frequency-band coupling in surface EEG reflects spiking activity in monkey visual cortex. *Neuron* 64, 281–289.

Williams, J., Ramaswamy, D., and Oulhaj, A. (2006). 10 Hz flicker improves recognition memory in older people. *BMC Neurosci.* 7, 21. doi: 10.1186/1471-2202-7-21

Williams, J. H. (2001). Frequency-specific effects of flicker on recognition memory. *Neuroscience* 104, 283–286.

Winter, W. R., Nunez, P. L., Ding, J., and Srinivasan, R. (2007). Comparison of the effect of volume conduction on EEG coherence with the effect of field spread on MEG coherence. *Stat. Med.* 26, 3946–3957.

Womelsdorf, T., Schoffelen, J. M., Oostenveld, R., Singer, W., Desimone, R., Engel, A. K., and Fries, P. (2007). Modulation of neuronal interactions through neuronal synchronization. *Science* 316, 1609–1612.

Wu, Z., and Yao, D. (2007). The influence of cognitive tasks on different frequencies steady-state visual evoked potentials. *Brain Topogr.* 20, 97–104.

Yeung, N., Bogacz, R., Holroyd, C. B., and Cohen, J. D. (2004). Detection of synchronized oscillations in the electroencephalogram: an evaluation of methods. *Psychophysiology* 41, 822–832.

Zaehle, T., and Herrmann, C. S. (2010). Neural synchrony and white matter variations in the human brain – Relation between evoked gamma frequency and corpus callosum morphology. *Int. J. Psychophysiol.* doi:10.1016/j.ijpsycho.2010.06.029. [Epub ahead of print].

Zoefel, B., Huster, R. J., and Herrmann, C. S. (2010). Neurofeedback training of the upper alpha frequency band in EEG improves cognitive performance. *Neuroimage* 54, 1427–1431.

*Received: 01 October 2010; paper pending published: 03 December 2010; accepted: 04 January 2011; published online: 19 January 2011.*

*Citation: Cohen MX (2011) It's about time. Front. Hum. Neurosci.* ***5****:2. doi: 10.3389/fnhum.2011.00002*

frontiers in
HUMAN NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 23 June 2010
doi: 10.3389/fnhum.2010.00052

# Fast optical imaging of human brain function

***Gabriele Gratton*[1,2]* *and Monica Fabiani*[1,2]**

[1] *Department of Psychology, University of Illinois at Urbana-Champaign, Urbana, IL, USA*
[2] *Beckman Institute, University of Illinois at Urbana-Champaign, Urbana, IL, USA*

***Edited by:***
*Michael X. Cohen, University of Amsterdam, Netherlands*

***Reviewed by:***
*Trevor B. Penney, National University of Singapore, Singapore*
*Ron Frostig, University of California, USA*

***Correspondence:**
*Gabriele Gratton, Beckman Institute, University of Illinois, 405 N. Mathews Avenue, Urbana, IL 61801, USA.*
*e-mail: grattong@illinois.edu*

Great advancements in brain imaging during the last few decades have opened a large number of new possibilities for neuroscientists. The most dominant methodologies (electrophysiological and magnetic resonance-based methods) emphasize temporal and spatial information, respectively. However, theorizing about brain function has recently emphasized the importance of rapid (within 100 ms or so) interactions between different elements of complex neuronal networks. Fast optical imaging, and in particular the event-related optical signal (EROS, a technology that has emerged over the last 15 years) may provide descriptions of localized (to sub-cm level) brain activity with a temporal resolution of less than 100 ms. The main limitations of EROS are its limited penetration, which allows us to image cortical structures not deeper than 3 cm from the surface of the head, and its low signal-to-noise ratio. Advantages include the fact that EROS is compatible with most other imaging methods, including electrophysiological, magnetic resonance, and trans-cranial magnetic stimulation techniques, with which can be recorded concurrently. In this paper we present a summary of the research that has been conducted so far on fast optical imaging, including evidence for the possibility of recording neuronal signals with this method, the properties of the signals, and various examples of applications to the study of human cognitive neuroscience. Extant issues, controversies, and possible future developments are also discussed.

**Keywords: event-related optical signal, non-invasive optical imaging, near-infrared spectroscopy, diffusive optical imaging, diffusive optical tomography, cognitive neuroscience**

## INTRODUCTION

The last few decades have seen an enormous increase in the number of studies about human brain function (Barinaga, 1997; Bandettini, 2009). This meteoric increase has been in large part due to the introduction of a series of new methodologies for the non-invasive measurement of brain physiological parameters, which we collectively label "brain imaging methods". These non-invasive techniques have generated a new paradigm, which emphasizes mass activity as a useful level for theorizing about human brain function. The basic assumption is that it is possible to describe the human brain as a collection of macroscopic structures with sizes measurable from a few mm's to several cm's range, performing specialized functions over times ranging between tenths of ms to seconds or even longer, and whose synergistic interactions result in the emergence of overall brain states and behavioral outcomes. Ideally, therefore, brain imaging methods should be capable of describing brain activity with a level of spatial and temporal resolution consistent with these temporal and spatial parameters (mm and ms, respectively). In practice, however, the most commonly used techniques (functional magnetic resonance imaging, or fMRI, and the event-related brain potential, or ERP) only reach this level of resolution in one dimension (respectively, space and time) but not the other. To obviate this problem, researchers have proposed combining these two techniques, which however leads to a number of practical problems (e.g., Luck, 1999). Alternatively, the problem can be addressed by other methods, capable of reaching target values of both spatial and temporal resolution. In this paper we describe the use of fast optical signals, and in particular of a methodology called the Event-Related Optical Signal (or EROS, Gratton et al., 1995; Gratton and Fabiani, 2009) to achieve this result.

Fast optical signals refer to changes in optical scattering that occur in neural tissue when the tissue is active (depolarized or hyperpolarized), compared to when it is not. They were first described in the 1940's (Hill and Keynes, 1949) in isolated nerves, and subsequently reported in hippocampal (Frostig et al., 1990; MacVicar and Hochman, 1991; Andrew and MacVicar, 1994) and brainstem slices (Momose-Sato et al., 1998), as well as integral brain preparations in both invertebrates (Stepnoski et al., 1991) and vertebrates (Rector et al., 1997, 2005). As the response is blocked by tetrodoxin (Lee and Kim, 2010), it appears that opening and closing of ion channels is critical for its presence. Current research (Foust and Rector, 2007) suggests that the biological basis of the phenomenon is swelling (in the case of depolarization, Buchheim et al., 1999; Lee and Kim, 2010) or shrinking (in the case of hyperpolarization, Momose-Sato et al., 1998) of neurites (mostly dendrites) due to movement of water across the membrane associated with ion transport (see also Lee and Kim, 2010 for a biophysical model of scattering phenomena associated with neural function).

Because of their association with neuronal activity (rather than hemodynamic responses, which follow several seconds later), fast optical signals can potentially yield measures of brain activity with high (ms-level) temporal resolution, comparable to those of electrophysiological methods (as demonstrated by the work of Stepnoski et al., 1991; Rector et al., 1997, 2005). In the last several years our lab and others have worked to determine whether (a) fast

optical signals can be recorded non-invasively in humans; and (b) whether the data can provide the combination of spatial and temporal resolution required for research in cognitive neuroscience.

Although recording of optical changes in exposed cortex has been carried out for several decades yielding images with exquisite spatial resolution (e.g., Grinvald et al., 1986), non-invasive measurement presents some challenges, which clearly limit the spatial resolution that can be achieved. Head tissues, such as the skin, skull, and meninges both absorb (mostly in the visible range) and scatter light. Light absorption is mostly due to the hemoglobin present in the blood, and can be minimized by using light in the far red and near-infrared (NIR) spectrum (Frostig et al., 1990). At these wavelengths, the major limitation to cortical imaging is due to scattering, which is mostly due to mitochondria, membranes, and other vesicles present in the tissue (Beauvoit et al., 1995). To image deep tissues a successful approach has been diffusive optical imaging (Jobsis, 1977; Gratton et al., 1997). Although initially this approach produced images with low spatial resolution, current methodology (based on the use of a large density of recording channels and of particular measurement procedures described in the section Materials and Methods) has pushed the spatial resolution to a sub-cm level (Gratton and Fabiani, 2003). With this methodology, our group (Gratton et al., 1995; for a review, see Gratton and Fabiani, 2009), as well as others (e.g., Steinbrink et al., 2000; Wolf et al., 2002; Franceschini and Boas, 2004; Lebid et al., 2005; Tse et al., 2006; Kubota et al., 2008; Medvedev et al., 2008) have shown that fast optical signals can be recorded consistently, with a combination of spatial and temporal resolution that is adequate for the type of neuroimaging research considered above. However, other groups have questioned this possibility (Steinbrink et al., 2005; Radakrishnan et al., 2009). In the remainder of this paper we present our methodology for recording fast optical signals (EROS) (see Materials and Methods) and review the evidence in favor and against the non-invasive recording of fast optical signals and current data about its spatial and temporal resolution (see Results). In the Discussion we evaluate the current status of the research and possible future directions.

## MATERIALS AND METHODS

### PROPAGATION OF LIGHT THROUGH TISSUE

Using light to generate images of brain function is a well-tested approach (Frostig, 2009). However, with diffuse illumination (as in normal lighting situations), penetration is reduced to a few mm, because of the strongly scattering and absorption properties of head tissues. To circumvent this problem and obtain images of cortical activity without opening the skull and/or lesioning the brain, it is necessary to use a different approach, called diffusive optical imaging. This approach is based on separating the locations from which light is inserted into the tissue (sources) from those at which it is measured (detectors) (see **Figure 1**). Point sources of NIR light are positioned at various locations on the surface of the head. Because of the scattering properties of the tissue, the light propagates in a random fashion around the source. A detector located at some distance will pick up some of this light. The combination of sources and detectors identifies spindles through which photons are likely to travel during their random motion. As the head is bounded by a non-scattering surface, photons that reach the surface in their random motion will penetrate a non-scattering medium, and therefore move in a straight line and permanently exit the head. Therefore photons that travel too close to the surface of the medium are not likely to reach distant detectors. This determines a "curvature" of the statistical spindles describing the motion of the photons. The maximum spindle depth is located approximately half-way between the source and the detector; the actual penetration depends, among other factors, on the source-detector distance. In practice, source-detector distances varying between 3–5 cm afford depths of the central portion of the spindle varying between 1–3 cm; this range of depths is sufficient to explore large areas of the cortical surface (such as large parts of the occipital, parietal, temporal, and frontal cortex), although some areas (such as medial temporal cortex, basal ganglia, the ventral surface of the cortex and deep regions inside sulci) are not accessible to this type measurement.

### MEASUREMENT OF LIGHT PARAMETERS

Two basic technologies for diffusive optical imaging are currently available. The first technology, called continuous-wave (or CW) method, is based on constant (or slowly oscillating, <10 kHz) sources of light. With this approach, only the total amount of light emitted by sources and reaching the detector can be measured (*intensity* or *DC intensity*). In contrast, the other technology uses rapidly varying sources of light, which afford measurement not only of the amount of variations in light intensity, but also of the time required by photons to move between a source and a detector (time of flight). Because of the diffusive nature of the photon movement and of the high refraction index of tissue, this time is in the ps or ns range. Photons' time of flight is most conveniently measured in the frequency domain (FD), as the phase delay of a photon density wave moving between the source and the detector. Currently, commercially available FD devices use sources that are modulated at frequencies ranging between 100 and 300 MHz. They can be used to derive three types of measures: average (or DC) intensity, amplitude of the intensity changes at the modulation frequency (AC intensity), and phase delay. In principle, even more information about the distribution of time-of-flight of individual photons can be obtained by time-resolved (or TR) method, which use pulsating time-sources, but these devices are more expensive and difficult to use.

Continuous-wave and frequency domain methods each offer advantages and disadvantages. CW methods are typically less expensive. However, phase delay measures may provide important data that are not available using intensity data alone. For example, phase data can be used to compute *absolute* absorption and scattering coefficients for a particular region. Although this may be very important for clinical studies, this measurement is complex, and is rarely conducted in fast signal studies. It is very important to understand that intensity and phase data are fundamentally different. Intensity measures are essentially counts of the photons reaching the detector. Phase delay data instead measure the average arrival time of photons. This implies that photons are weighted by their time-of-arrival, with increments or decrements in the number of photons that have times-of-flight shorter than the mean producing opposite effects than increments or decrements in the number of photons that have times-of-flight longer than the mean. Photons traveling deep inside the head typically travel longer distances (and have therefore longer times-of-flights) than photons traveling more

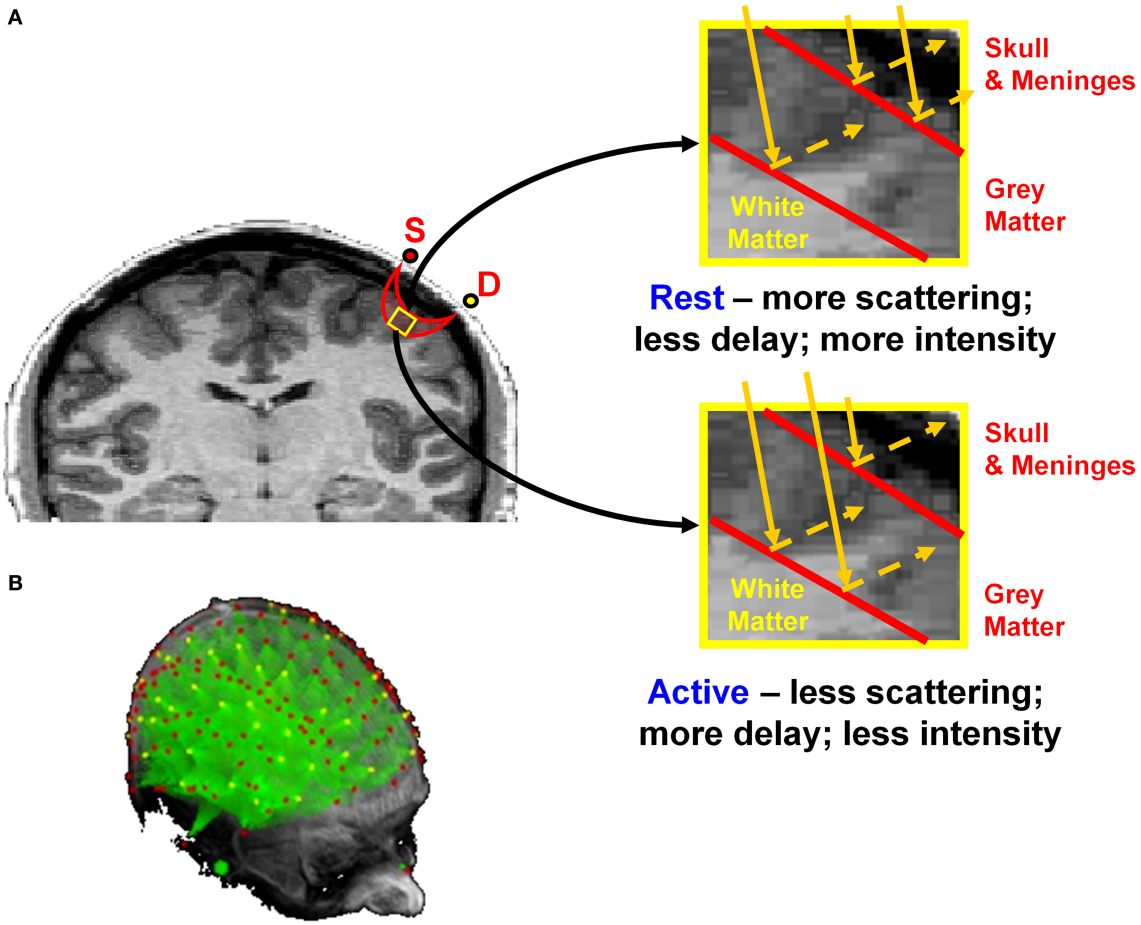



**FIGURE 1 | (A)** Schematic representation of the back-scattering of photons under conditions of rest and activity in the cortex. S = Source (red dot); D = detector (yellow dot). As shown in the right panels, changes in transparency of the cortex are associated with changes in photon penetration and path length. Adapted from Figure 7 in Gratton and Fabiani (2003). **(B)** Projections (in green) of the areas investigated by a large number of optical sources (red dots) and detectors (yellow dots) onto 3D renditions of an MR anatomical image.

superficially. As a consequence, variations in the number of photons traveling deep will have an *opposite* effect on the phase delay parameter than variations in the number of photons traveling superficially. In other words, the net effects of absorption or scattering changes on the photon delay parameter are critically dependent on *where* the changes occur. This has three critical consequences for fast optical imaging: (a) phase delay measures have a very different depth-sensitivity compared to intensity measures, with greater sensitivity for deeper locations (this is due to the fact that photons traveling a very long path have very long times-of-flight, and greater influence on the mean value measured by phase delay)); (b) phase delay measures have a greater spatial resolution than intensity measures (this is because even small changes in the relative number of photons traveling long or short paths may have a large effect on the phase value); and (c) phase delay measures are largely insensitive to the total amount of light (and variations thereof) injected into the tissue, or picked up at the detector (this because these changes will equally influence photons traveling long and short paths, and therefore have no net effect on the phase parameter). This makes phase delay measures largely insensitive to movements (Schneider-Garces et al., 2009), which cause changes at the interface between the sources and detectors and the head. These movement-related changes instead cause large artifacts on intensity measures, which need to be dealt with appropriate procedures (Sato et al., 2006; Medvedev et al., 2008; Huppert et al., 2009; Schneider-Garces et al., 2009; Robertson et al., 2010). These three factors make the phase delay parameter particularly interesting for deriving images of brain activity with high spatial and temporal resolution.

## CO-REGISTRATION AND IMAGE RECONSTRUCTION

Optical data provide little anatomical information. As a consequence, to determine the area of the brain to which measurements relate it is important to provide an anatomical frame of reference. In initial studies, this frame of reference was based on surface scalp measures (e.g., inion, vertex, e.g., Gratton et al., 1995). However, this frame of reference is very approximate. Therefore, we now routinely co-register the locations of sources and detectors to structural MR images obtained on the same subjects, using both fiducial alignment and fitting methods (Whalen et al., 2008). These co-registration methods lead to errors of less than 3–4 mm.

To produce images, it is necessary to obtain measurements from a number of locations. CW and FD systems allowing for concurrent recording from multiple locations are currently commercially available. In general, because of the scattering and absorption properties of tissue, a detector can only pick up light emitted from source less than 5–6 cm away. In addition, even much closer sources can be separated from each other if they are time-, frequency- or wavelength-multiplexed. This permits to achieve a very high spatial sampling, with a channel (defined as a combination of a source and a detector) every few squares cm. Wolf et al. (2000) showed that using a high spatial sampling and multiple channels with overlapping "curved spindles" may result in large increases in the spatial resolution of optical imaging data. According to our computations (e.g., Gratton and Fabiani, 2003), the spatial resolution thus obtained can be as high as 15–20 mm for intensity measures and 5–10 mm for phase delay measures (which, as mentioned above, have a greater spatial resolution than intensity measures). This spatial resolution is comparable to that of most fMRI studies.

The majority of the fast optical imaging studies reported so far are based on surface-projected images. These images are typically obtained by back-projection, a technique in which the values obtained from a particular channel are represented as changes occurring at particular locations (or areas) on the brain surface in between the source and the detector. If the same location lies in between multiple combinations of sources and detectors, the corresponding values are combined (typically by arithmetic averaging for phase measures or geometric averaging for intensity measures; Wolf et al., 2000; Gratton and Fabiani, 2009). Statistical analysis can then be conducted pixel-wise in a manner analogous to what is done for fMRI, using appropriate corrections for multiple comparisons (Kiebel et al., 1999). Three-dimensional reconstruction algorithms for intensity data have also been proposed (Roggan et al., 1994; Zhu et al., 1997; Boas et al., 2001; Intes et al., 2002, 2004; Fukui et al., 2003). We have recently developed three-dimensional reconstruction algorithms that can be used for both intensity and phase data; these methods are currently under evaluation.

### ARTIFACT TREATMENT AND NOISE REDUCTION

Fast signals are very small (of the order of 1/1000 for intensity measures and ps or fractions thereof for phase delay measures). These signals need to be separated from noise that may be several orders of magnitude larger. Two types of noise are important: large, non-random noise, which occurs sporadically or it is clearly of non-brain origin (artifacts), and small, random noise, which occurs continuously and it may be generated by a host of factors including brain and other physiological phenomena (background noise).

There are two major sources of artifacts in fast optical imaging: movements and changes related to hemodynamics. As mentioned above, intensity data are particularly sensitive to changes in the interface between the optical instrumentation and the head surface, which may occur during large head movements. These changes are particularly significant when the connection between the instruments and the head is not very tight. Therefore, use of appropriate holders for the optical devices may greatly reduce this type of artifact. In addition, investigators have proposed a number of approaches for dealing with this artifact, including discarding trials with evidence of movement artifact (using algorithms for identifying such artifacts on regular channels – Sato et al., 2006; Huppert et al., 2009; Schneider-Garces et al., 2009; Robertson et al., 2010), or correcting the estimated effect of artifact either using regression methods (this requires estimating the occurrence of movements through specialized channels with source-detector distances so short to not be likely to be influenced by brain effects – e.g., Medvedev et al., 2008) or statistically-based methods (in which the principal components of data are assumed to be due to movements and discarded from the data – Huppert et al., 2009; Robertson et al., 2010). In any case, for the reasons explained earlier, phase data appear largely impervious from movement artifacts (Schneider-Garces et al., 2009), so that movement correction does not appear necessary in this case. The second source of artifacts is related to hemodynamic effects, which cause changes in absorption that may overlap with the fast effects. Some of these changes, such as those due to the vasodilation effects that form the basis of the fMRI Blood-Oxygenation-Level-Dependent (BOLD) signal, are very slow and can be easily separated from fast signals through frequency filtering. Others, due to the oscillations occurring during a cardiac cycle, are very rapid (about 1 Hz). To eliminate these artifacts we (Gratton and Corballis, 1995) developed an adaptive filtering algorithm that is effective in greatly reducing the pulse artifact.

Background noise can be reduced significantly using signal averaging. This procedure, however, introduces the problem of having to record a large number of trials, which may impose limits in the experimental design. Another approach that is often used in combination with averaging is filtering. Three types of filters have been most commonly used: frequency filters (Maclin et al., 2003), spatial filters and statistically-based adaptive filtering (e.g., application of independent component analysis, or ICA, Morren et al., 2004; Medvedev et al., 2008). The first two approaches have the limitation of reducing the temporal or spatial resolution of data. The third requires making specific assumptions about what constitutes signal and what should be considered as noise. In any case, large improvements in signal-to-noise ratio (SNR) have been obtained using these approaches (Maclin et al., 2003; Medvedev et al., 2008).

Finally, a practical problem in most brain optical imaging studies is the presence of hair, which, depending on the color, may absorb a large proportion of the light used for the measurement, thus greatly limiting our ability to image brain activity. It is therefore essential to comb the hair away from sources and detectors. This greatly improves the ability to record optical data without participants' exclusion. Note that similar procedures are also routinely applied for EEG recording.

## RESULTS

### DETECTION OF THE FAST OPTICAL SIGNAL

To demonstrate the existence of the fast optical signal, our lab and others have shown that (a) a fast response with *latency* compatible with a known signal is consistently present in optical data (e.g., Gratton et al., 1997, 2001; Rinne et al., 1999; DeSoto et al., 2001; Tse et al., 2006, 2007); (b) this response does *not* occur in appropriately selected control conditions (e.g., Gratton et al., 1995, 2000, 2009; DeSoto et al., 2001; Wolf et al., 2002; Franceschini and Boas, 2004; Tse et al., 2006; Medvedev et al., 2008). These control conditions need to be chosen so as to eliminate any account that is not due to neural function.

There are now a large number (>30) of published studies that satisfy these requirements (see Gratton and Fabiani, 2009 for a recent review). These studies come from several different laboratories, involve both intensity and phase delay measures, and are based on a number of different paradigms. For example, in the study providing the first description of the fast optical signal (Gratton et al., 1995), four different visual stimulation conditions (one for each quadrant of the visual field) were used. Recordings were made from 12 locations over the occipital area chosen so as to yield a retinotopic map of primary visual cortex (V1). Thus, for each of the four stimulation conditions, a different quadrant of V1 was expected to be stimulated, with the control condition provided by the same quadrant of V1 when other visual field quadrants were stimulated. In this fashion each area provided a control for itself. The results indicated that the response (an increase of the phase delay with a latency of 50–100 ms from stimulation consistent with the initial response in V1 measured with electrophysiological methods) occurred only in the stimulated condition but not in the others, for each of the four quadrants of V1. Although the original study was based on a small *N*, this finding was replicated in a subsequent study with a larger *N*, a higher sampling rate (20 ms), and a higher spatial sampling (Gratton and Fabiani, 2003). This study showed a response characterized by an increase in phase delay peaking at a latency of approximately 80 ms from stimulation. Results from this replication study are reported in **Figure 2**. Similarly, Gratton et al. (2000) presented stimuli varying in stimulus eccentricity (1, 2, 4, and 8 degrees). According to the known retinotopic organization of primary visual cortex, it was predicted that the location of the response in V1 should occur at progressively deeper locations depending on the eccentricity of the visual stimulus (this was controlled using an fMRI study on the same subjects). In this case, the prediction was that the response should occur in optical channels with progressively greater source-detector distance, reflecting their change in penetration inside the head. As for the previous study, each set of channels with different source-detector distance would serve as both the one in which the response should be observed and as a control when it should not. The results were consistent with the predictions: the source-detector distance of channels showing the maximum response (as in the other studies, an increase in phase delay with a latency of 80 ms from stimulation) varied systematically as a function of stimulus eccentricity. As in the previous example, in this case the finding was replicated in a recent study (see **Figure 3**, Maclin et al., 2008).

Notwithstanding the large number of controlled studies supporting the detection of an optical response, there have been a handful of studies that have presented apparently contrasting results. Of these studies, all based on intensity measures, one was conducted in humans (Steinbrink et al., 2005) and one intradurally in monkeys (Radakrishnan et al., 2009). The Steinbrink et al. (2005) is based on somatosensory stimulation. The authors reported the occurrence of a significant fast optical response only when the stimulation was above threshold for motor response: this led them to consider likely that the observed response was based on movement artifacts, thus contradicting an earlier report from the same group (Steinbrink et al., 2000). However, it should be noted that movements of the head were not monitored, and a reduction of the neuronal response as a function of a weaker stimulation could have been predicted on the basis of extant knowledge. The authors also reported failure to identify responses in the case of visual stimulation, but the recording conditions appear to vary significantly across subjects. In both the somatosensory and visual conditions, very few recording channels were used (1 or 2), which is clearly sub-optimal, and no alignment with anatomical structures was used. The Radakrishnan et al. (2009) study reports absence of a fast optical response in the

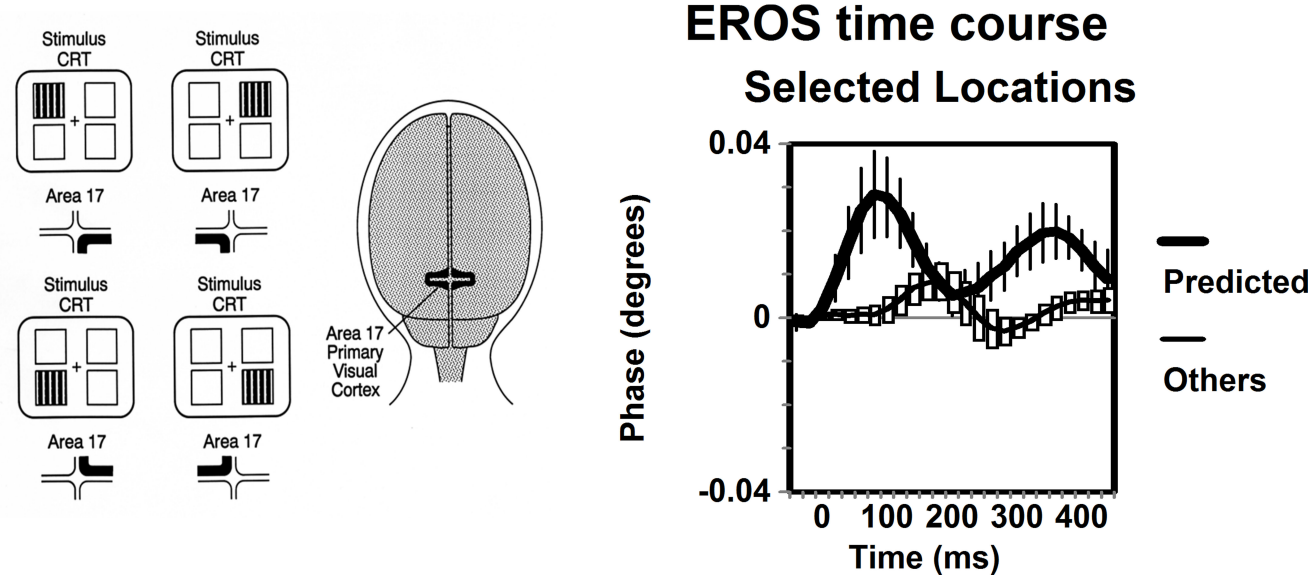



**FIGURE 2 | Adapted from Gratton et al. (2003).** The left panel depicts the stimulation conditions used and the cortical regions that are predicted to carry the response. The right panel indicated the EROS time course from the predicted location averaged across the four stimulation conditions (thick lines), and the responses from the same locations when the other stimulation conditions where presented (thin lines). Error bars are based on the standard error of the mean ($N = 8$).

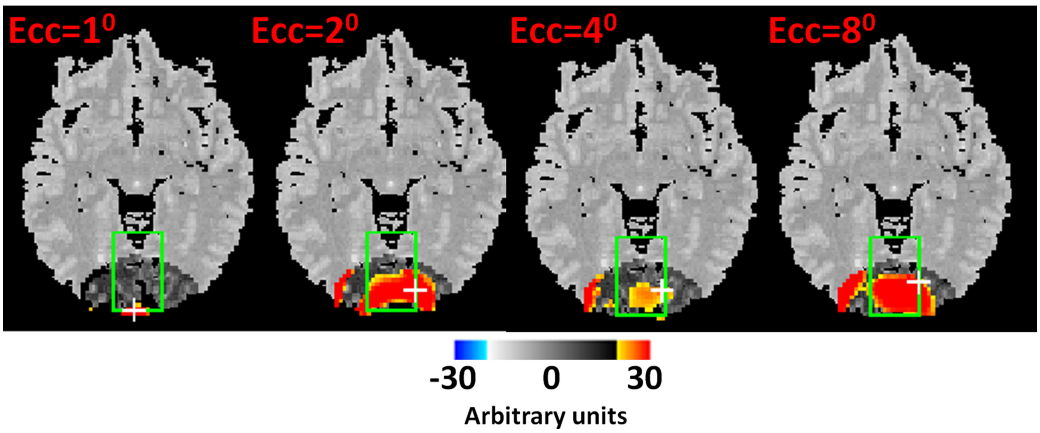



**FIGURE 3 | From Maclin et al. (2008).** Three-dimensional reconstruction of EROS phase delay data for each stimulus eccentricity condition (Ecc = stimulus eccentricity). The data are grand average maps ($N = 14$) of an axial slice ($z = -10$ in MNI space) obtained at a latency of 76 ms from stimulation (sampling rate was 25.6 ms). The white crosses indicate the peak points. The data are displayed in arbitrary phase units.

primary visual cortex of a macaque monkey, thus apparently contradicting not only the human studies, but also the animal studies reviewed earlier in this paper. A problem with this study is that the sources and detectors were located at a distance of approximately 1 cm from each other, in very close proximity to the surface of the visual cortex. As the thickness of V1 in macaque is less than 2 mm, it is quite possible that the curved spindles describing the optical sensitivity would have overshot the active area of the brain. Although they report the occurrence of slow, hemodynamic responses with CW methods, these responses are known to have a lower spatial resolution (see Frostig et al., 1990 for a demonstration of the spread of the slow hemodynamic responses), and thus they may originate in different areas, not investigated with fast responses. If this proved to be the case, it would suggest that fast optical signals need dense and extended spatial sampling for appropriate recording. The Radakrishnan et al. (2009) paper also reports FD measurements, but neither fast nor slow responses were reliably observed with these methods, suggesting the occurrence of methodological problems. Both studies also report Monte Carlo simulations indicating that, if existent, the fast optical response should have been observable, but only with intensity and not with phase measures. However, the results of these simulations contradict other simulations presented previously by one of the two groups (Franceschini and Boas, 2004). Further, these simulations only explore a relatively small set of conditions, not necessarily covering real life situations.

In summary, the vast majority of the published studies support the detection of fast optical responses in a variety of different paradigms. The fast optical response is typically characterized by a reduction in light intensity and an increase in phase delay at a latency corresponding to the expected time of cortical activation for each paradigm. Fast optical responses have been reported in visual (e.g., Gratton et al., 1995, 2000, 2001, 2003, 2006), auditory (Rinne et al., 1999; Fabiani et al., 2006; Tse et al., 2006; Tse and Penney, 2007, 2008), somatosensory (Steinbrink et al., 2000; Maclin et al., 2004; Franceschini and Boas, 2004) and motor cortex (DeSoto et al., 2001; Wolf et al., 2002; Morren et al., 2004), under appropriate modality-specific conditions. Further, fast optical responses have been observed in higher-order cortical regions including prefrontal and parietal cortex (Low et al., 2006, 2009; Tse et al., 2006, 2007; Medvedev et al., 2008; Gratton et al. 2009) in various paradigms using cognitive manipulations.

## SPATIAL AND TEMPORAL RESOLUTION OF FAST OPTICAL IMAGING

The spatial resolution of fast optical imaging is partly dependent on the methods used for its measurement. Three sets of studies have specifically investigated this issue, all using tasks involving the activation of different segments of functionally organized cortical regions. They include the two visual stimulation experiments described above (Gratton et al., 1995; Gratton and Fabiani, 2003), and a study using finger tapping by Wolf et al. (2002). All these studies have shown that fast optical imaging of phase data can distinguish between the activities of areas as close as 5–10 mm from each other, provided that sufficient spatial sampling is attained (at least 1 channel/square cm). However, the latter study also compared the spatial resolution of phase and intensity data, showing that phase data are much more localized than intensity data. An alternative approach is based on the statistical computation of the size of "resels" (volumes with uncorrelated error terms) obtained using the random field theory (Worsley et al., 1992; Kiebel et al., 1999). The results of this analysis are consistent with the spatial resolution estimates obtained with functional manipulations (5–10 mm spatial resolution for phase data, 10–20 mm for intensity data). The temporal resolution of optical data also appears to depend on the sampling rate adopted, at least for frequencies up to 100 Hz. In a series of studies in which ERPs and EROS were recorded simultaneously, the timing of the optical response appears to coincide with that of ERP responses, at least at the level of the sampling rate adopted (e.g., Tse et al., 2007). Note that current limits in the temporal and spatial resolution of EROS could be improved with hardware and software development. For example, current equipment requires multiplexing, thus limiting the temporal resolution

of the technique. Recording hardware updates allowing for more parallel data collection would bring the temporal resolution on par with that obtained with ERPs and MEG.

## DISCUSSION

Taken together, the work published so far provides a strong support for the idea that fast optical signals (EROS) can be detected from surface recording, providing a tool for studying rapid changes in brain activity. The temporal and spatial resolutions of EROS depend in part on the methods used. However, when phase measures and high spatial sampling are adopted, the spatial resolution can be as high as 5–10 mm, in the range of most published fMRI work in cognitive neuroscience. The temporal resolution appears approximately similar to that obtained with EEG and ERPs, and again depends on the sampling rate.

When compared to other brain imaging methods, EROS has several advantages and some limitations. The major limitations include the restricted depth of penetration (a few cm from the head surface), and the low SNR, which renders it necessary to accrue data across a number of trials[1]. The major advantages are contained cost and relative portability (when compared to MRI, positron emission tomography – PET, and magnetoencephalography – MEG) and ease of concurrent recording with other measures. In fact, EROS can and has been recorded simultaneously with ERPs (e.g., DeSoto et al., 2001; Gratton et al., 2001), fMRI (Toronov et al., 2005; Zhang et al., 2005), and, in a current pilot study in our lab, Transcranial Magnetic Stimulation (TMS) without any evidence of interference in either direction. This is potentially a great asset as fast optical data may make an ideal bridge technology for neuroimaging data fusion (Barinaga, 1997).

The types of data obtained with EROS allow for different types of analyses, including study of the rapid interaction of different cortical regions. For example, Rykhlevskaia et al. (2006) showed that EROS can be used to identify the flow of information across cortical regions, as well as excitation and inhibition between cortical regions. They also showed that this interaction may be related to the anatomical connections between the areas (see also Gratton et al., 2009). These types of data may be critical for our theorizing about how different regions of the brain cooperate to perform cognitive functions.

Other methods can also be used to generate dynamic images of brain activity. For instance, MEG can be used to derive images combining spatial and temporal resolution comparable to those obtained with EROS (Hari et al., 2000). However, this is predicated on the use of complex analytical methods to address the inverse problem. These methods typically require simplifying assumptions about the properties of the signal to be modeled (cortical activity). Reliance on these simplifying assumptions is much less critical for EROS. Further, MEG is more expensive than EROS.

Another approach to achieve dynamic imaging of brain activity is to combine data from different imaging modalities (e.g., ERPs and fMRI, Luck, 1999). A problem with this approach is that this relationship is not well understood and probably subject to many variations. As mentioned earlier, because of its temporal similarity with ERPs and spatial affinity with fMRI, EROS can provide a bridge between these two modalities (e.g., Gratton et al., 1997), and serve to test and/or validate the presumed relationship between them. The combined recording of different modalities can also afford new intriguing views into brain function. For example, the combined use of EROS and source-localized ERP measures may help distinguish between the activity of inter-neurons and pyramidal cells. This may be possible because EROS is likely sensitive to both interneuron and pyramidal cell activity, whereas ERPs are selective for pyramidal neurons only, albeit with limited spatial resolution.

At present, EROS is still in its infancy. Applications to various areas of cognitive neuroscience, such as attention (Gratton, 1997; DeSoto et al., 2001), sensory (Rinne et al., 1999; Fabiani et al., 2006; Tse et al., 2006; Tse and Penney, 2007, 2008) and working (Low et al., 2006; Medvedev et al., 2008) memory, executive and preparatory (Gratton et al., 2008) processes, and language (Tse et al., 2007; Kubota et al., 2008) are beginning to appear in the last few years, as well as applications involving not only young adults but also older adults (Fabiani et al., 2006; Gratton et al., 2009). However, the methodology is still evolving. Current methodological research includes work on the development of more efficient hardware and data collection software (including both acquisition systems that are faster, allow for the recording from a larger number of channels, and have higher SNR and, more suitable devices for securely connecting the measurement system to the head), three-dimensional reconstruction algorithms for both intensity and phase data, filtering and data extraction systems, more efficient procedures to deal with artifact, and so on. It is likely that application of these methodological advancements will result in significant improvement in the quality of the data, including higher spatial and temporal resolution and higher SNR. Hopefully these advancements will lead to a new level of analysis of human brain function.

[1]Typical recordings aim at obtaining at least 50–100 trial per condition per subject. However, the exact number needed depends on the size of the signal, which varies somewhat in different cortical regions and as a function of task conditions. For example, visual experiments with repeated stimulation may require many more trials (several hundreds) than recordings from frontal cortex during attention-demanding tasks (a few dozen).

## ACKNOWLEDGMENTS

The work presented in this paper was supported by NIA grant 1RC1AG035927 to Monica Fabiani and NIMH grant # MH080182 to Gabriele Gratton.

## REFERENCES

Andrew, R. D., and MacVicar, B. A. (1994). Imaging cell volume changes and neuronal excitation in the hippocampal slice. *Neuroscience* 62, 371–383.

Bandettini, P. A. (2009). What's new in neuroimaging methods? *Ann. N. Y. Acad. Sci.* 1156, 260–293.

Barinaga, M. (1997). New imaging methods provide a better view into the brain. *Science* 276, 1974–1976.

Beauvoit, B., Evans, S. M., Jenkins, T. W., Miller, E. E., and Chance, B. (1995). Correlation between the light scattering and the mitochondrial content of normal tissues and transplantable rodent tumors. *Anal. Biochem.* 226, 167–174.

Boas, D. A., Culver, J. P., Stot, J. J., and Dunn, A. K. (2001). Three dimensional Monte Carlo code for photon migration through complex heterogeneous media including the adult human head. *Opt. Express* 10, 159–170.

Buchheim, K., Schuchmann, S., Siegmund, H., Gabriel, H. J., Heinemann, U., and Meierkord, H. (1999). Intrinsic optical signal measurements reveal characteristic features during different forms of spontaneous neuronal hyperactivity associated with ECS shrinkage *in vitro*. *Eur. J. Neurosci.* 11, 1877–1882.

DeSoto, M. C., Fabiani, M., Geary, D. L., and Gratton, G. (2001). When in doubt, do it both ways: Brain evidence of the simultaneous activation of conflicting

responses in a spatial Stroop task. *J. Cogn. Neurosci.* 13, 523–536.
Fabiani, M., Low, K. A., Wee, E., Sable, J. J., and Gratton, G. (2006). Reduced suppression or labile memory? Mechanisms of inefficient filtering of irrelevant information in older adults. *J. Cogn. Neurosci.* 18, 637–650.
Foust, A. J., and Rector, D. M. (2007). Optically teasing apart neural swelling and depolarization. *Neuroscience* 145, 887–899.
Franceschini, M. A., and Boas, D. A. (2004). Noninvasive measurement of neuronal activity with near-infrared optical imaging. *Neuroimage* 21, 372–386.
Frostig, R. D. (Ed.) (2009). *In Vivo Optical Imaging of Brain Function,* 2nd edn. (Boca Raton, FL: CRC Press).
Frostig, R. D., Lieke, E. E., Ts'o, D. Y., and Grinvald, A. (1990). Cortical functional architecture and local coupling between neuronal activity and the microcirculation revealed by *in vivo* high- resolution optical imaging of intrinsic signals. *Proc. Natl. Acad. Sci. U.S.A.* 87, 6082–6086.
Fukui, Y. Ajichi, Y., and Okada, E. (2003). Monte Carlo prediction of near-infrared light propagation in realistic adult and neonatal head models. *Appl. Opt.* 42, 2881–2887.
Gratton, E., Fantini, S., Franceschini, M. A., Gratton, G., and Fabiani, M. (1997). Measurements of scattering and absorption changes in muscle and brain. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* 352, 727–735.
Gratton, G. (1997). Attention and probability effects in the human occipital cortex: an optical imaging study. *Neuroreport* 8, 1749–1753.
Gratton, G., Brumback, C. R., Gordon, B.A., Pearson, M.A., Low, K. A. and Fabiani, M. (2006). Effects of measurement method, wavelength, and source-detector distance on the fast optical signal. *Neuroimage*, 32, 1576–1590.
Gratton, G., and Corballis, P. M. (1995). Removing the heart from the brain: compensation for the pulse artifact in the photon migration signal. *Psychophysiology* 32, 292–299.
Gratton, G., Corballis, P. M., Cho, E., Fabiani, M., and Hood, D. (1995). Shades of gray matter: Noninvasive optical images of human brain responses during visual stimulation. *Psychophysiology* 32, 505–509.
Gratton, G., Goodman-Wood, M. R., and Fabiani, M. (2001). Comparison of neuronal and hemodynamic measures of the brain response to visual stimulation: an optical imaging study. *Human Brain Mapping* 13, 13–25.
Gratton, G., and Fabiani, M. (2003). The event-related optical signal (EROS) in visual cortex: replicability, consistency, localization and resolution. *Psychophysiology* 40, 561–571.
Gratton, G., and Fabiani, M. (2009). "Fast optical signals: Principles, methods, and experimental results," in *In Vivo Optical imaging of Brain Function*, 2nd edn, ed. R. D. Frostig (Boca Raton, FL: CRC Press), 435–460.
Gratton, G., Low, K. A., and Fabiani, M. (2008). "Time course of executive processes: data from the event-related optical signal (EROS)," in *Perspectives on Rule-Guided Behavior*, eds S. A. Bunge and J. D. Wallis (New York, NY: Oxford University Press) 197–223.
Gratton, G., Rykhlevskaia, E., Wee, E., Leaver, E., and Fabiani, M. (2009). Does white matter matter? Spatiotemporal dynamics of task switching in aging. *J. Cogn. Neurosci.* 21, 1380–1395.
Gratton, G., Sarno, A. J., Maclin, E., Corballis, P. M., and Fabiani, M. (2000). Toward non-invasive 3-D imaging of the time course of cortical activity: Investigation of the depth of the event-related optical signal (EROS). *Neuroimage* 11, 491–504.
Grinvald, A., Lieke, E., Frostig, R. D., Gilbert, C. D., and Wiesel, T. N. (1986). Functional architecture of cortex revealed by optical imaging of intrinsic signals. *Nature* 324, 361–364.
Hari, R., Levänen, S., and Raij, T. (2000). Timing of human cortical functions during cognition: role of MEG. *Trends Cogn. Sci.* 4, 455–462.
Hill, D. K., and Keynes, R. D. (1949). Opacity changes in stimulated nerve. *J. Physiol.* 108, 278–281.
Huppert, T. J., Diamond, S. G., Franceschini, M. A., and Boas, D. A. (2009). HomER: A review of time-series analysis methods for near-infrared spectroscopy of the brain. *Appl. Opt.* 48, 280–298.
Intes, X., Maloux, C., Guven, M., Yazici, M., and Chance, B. (2004). Diffuse optical tomography with physiological and spatial a priori constraints. *Phys. Med. Biol.* 49, 155–163.
Intes, X, Ntziachristos, V, Culver, J. P., Yodh, A., and Chance, B. (2002). Projection access order in algebraic reconstruction technique for diffuse optical tomography. *Phys. Med. Biol.* 47, 1–10.
Jobsis, F. F. (1977). Noninvasive, infrared monitoring of cerebral and myocardial oxygen sufficiency and circulatory parameters. *Science* 198, 1264.
Kiebel, S. J., Poline, J.-B., Friston, K. J., Holmes, A. P., and Worsley, K. J. (1999). Robust smoothness estimation in statistical parametric maps using standardized residuals from the general linear model. *Neuroimage* 10, 756–766.
Kubota, M., Inouchi, M., Dan, I., Tsuzuki, D., Ishikawa, A., and Scovel, T. (2008). Fast (100–175 ms) components elicited bilaterally by language production as measured by three-wavelength optical imaging. *Brain Res.* 1226, 124–133.
Lebid, S., O'Neill, R. O., Markham, C., Ward, T., and Cole, S. (2005). "Multitimescale measurements of brain responses in visual cortex during stimulation using time-resolved spectroscopy," in *Proceedings of the SPIE conference*, Bruges, Belgium.
Lee, J., and Kim, S. J. (2010). Spectrum measurement of fast optical signal of neural activity in brain tissue and its theoretical origin. *Neuroimage* 51, 713–722.
Low, K. A., Leaver, E., Kramer, A. F., Fabiani, M., and Gratton, G. (2006). Fast optical imaging of frontal cortex during active and passive oddball tasks. *Psychophysiology* 43, 127–136.
Low, K. A., Leaver, E., Kramer, A. F., Fabiani, M., and Gratton, G. (2009). Share or compete? Load-dependent recruitment of prefrontal cortex during dual-task performance. *Psychophysiology* 46, 1–11.
Luck, S. J. (1999). Direct and indirect integration of event-related potentials, functional magnetic resonance images, and single-unit recordings. *Hum. Brain Mapp.* 8, 115–120.
Maclin, E. L., Gratton, G., and Fabiani, M. (2003). Optimum filtering for EROS measurements. *Psychophysiology* 40, 542–547.
Maclin, E. L., Low, K. A., Sable, J. J., Fabiani, M., and Gratton, G. (2004). The Event Related Optical Signal (EROS) to electrical stimulation of the median nerve. *Neuroimage* 21, 1798–1804.
Maclin, E. L., Tse, C.-Y., Fabiani, M., and Gratton, G. (2008). Probing the depth of optical signals: a visual eccentricity study. *Psychophysiology* 45, S93.
MacVicar, B. A., and Hochman, D. (1991). Imaging of synaptically evoked intrinsic optical signals in hippocampal slices. *J. Neurosci.* 11, 1458–1469.
Medvedev, A. V., Kainerstorfer, J., Borisov, S. V., Barbour, R. L., and VanMeter, J. (2008). Event-related fast optical signal in a rapid object recognition task: Improving detection by the independent component analysis. *Brain Res.* 1236, 145–158.
Momose-Sato, Y., Sato, K., Hirota, A., and Kamino, K. (1998). GABA-Induced intrinsic light-scattering changes associated with voltage-sensitive dye signals in embryonic brain stem slices: coupling of depolarization and cell shrinkage. *J. Neurophysiol.* 79, 2208–2217
Morren, G., Wolf, M., Lemmerling, P., Wolf, U., Choi, J. H., Gratton, E., De Lathauwer, L., and Van Huffel, S. (2004). Detection of fast neuronal signals in the motor cortex from functional near infrared spectroscopy measurements using independent component analysis. *Med. Biol. Eng. Comput.* 42, 92–99.
Radakrishnan, H., Vanduffel, W., Deng, H. P., Ekstrom, L., Boas, D. A., and Franceschini, M. A. (2009). Fast optical signal not detected in awake behaving monkeys. *Neuroimage* 45, 419–419.
Rector, D. M., Carter, K. M., Volegov, P. L., and George, J. S. (2005). Spatio-temporal mapping of rat whisker barrels with fast scattered light signals. *Neuroimage* 26, 619–627.
Rector, D. M., Poe, G. R., Kristensen, M. P., and Harper, R. M. (1997). Light scattering changes follow evoked potentials from hippocampal Schaeffer collateral stimulation. *J. Neurophysiol.* 78, 1707–1713.
Rinne, T., Gratton, G., Fabiani, M., Cowan, N., Maclin, E., Stinard, A., Sinkkonen, J., Alho, K., and Näätänen, R. (1999). Scalp-recorded optical signals make sound processing from the auditory cortex visible. *Neuroimage* 10, 620–624.
Robertson, F., Douglas, T., and Meintjes, E. (2010). Motion artefact removal for functional near infrared spectroscopy: a comparison of methods. *IEEE Trans. Biomed. Eng.* 57, 1377–1387.
Roggan, A., Minet, O. Schroeder, C., and Mueller, G. (1994). Determination of optical tissue properties with double integrating sphere technique and Monte Carlo simulations. *Proc. SPIE* 2100, 42–56.
Rykhlevskaia, E., Fabiani, M., and Gratton, G. (2006). Lagged covariance structure models for studying functional connectivity in the brain. *Neuroimage* 30, 1203–1218.
Sato, H., Tanaka, N., Uchida, M., Hirabayashi, Y., Kanai, M., Ashida, T., Konishi, I, and Maki, A. (2006). Wavelet analysis for detecting body-movement artifacts in optical topography signals. *Neuroimage* 33, 580–587.
Schneider-Garces, N., Maclin, E., Gratton, G., and Fabiani, M. (2009). Movement correction in optical data: comparison between feature-based and principal component-based approaches. *Psychophysiology* 46, S108.
Steinbrink, J., Kempf, F. C., Villringer, A., and Obrig, H. (2005). The fast optical signal – robust or elusive when non-invasively measured in the human adult? *Neuroimage* 26, 996–1008.
Steinbrink, J., Kohl, M., Obrig, H., Curio, G., Syré, F., Thomas, F., Wabnitz, H., Rinneberg, H., and Villringer, A. (2000). Somatosensory evoked fast optical intensity changes detected non-invasively in the adult human head. *Neurosci. Lett.* 291, 105–108.
Stepnoski, R. A., LaPorta, A., Raccuia-Behling, F., Blonder, G. E., Slusher, R. E., and Kleinfeld, D. (1991). Noninvasive detection of changes in membrane

potential in cultured neurons by light scattering. *Proc. Natl. Acad. Sci. U.S.A.* 88, 9382–9386.

Toronov V. Y., Zhang, X., Fabiani, M., Gratton, G., and Webb, A. G. (2005). "Signal and image processing techniques for functional near-infrared imaging of the human brain," in *Complex Dynamics and Fluctuations in Biomedical Photonics II*, ed. V. V. Tuchin (Proc. SPIE Vol. 5696), San Jose, California, USA.

Tse, C.-Y., Lee, C.-L., Sullivan, J., Garnsey, S., Dell, G. S., Fabiani, M., and Gratton, G. (2007). Imaging the cortical dynamics of language processing with the event-related optical signal. *Proc. Natl. Acad. Sci. U.S.A.* 104, 17157–17162.

Tse, C. Y., and Penney, T. B. (2007). Preattentive change detection using the event-related optical signal. *IEEE Eng. Med. Biol. Mag.* 26, 52–58.

Tse, C. Y., and Penney, T. B. (2008). On the functional role of temporal and frontal cortex activation in passive detection of auditory deviance. *Neuroimage* 41, 1462–1470.

Tse, C. Y., Tien, K. R., and Penney, T. B. (2006). Event-related optical imaging reveals the temporal dynamics of right temporal and frontal cortex activation in pre-attentive change detection. *Neuroimage* 29, 314–320.

Whalen, C., Maclin, E. L., Fabiani, M., and Gratton, G. (2008). Validation of a method for coregistering scalp recording locations with 3D structural MR images. *Hum. Brain Mapp.* 29, 1288–1301.

Wolf, M., Wolf, U., Choi, J. H., Gupta, R., Safonova, L. P., Paunescu, L. A., Michalos, A., and Gratton, E. (2002). Functional frequency-domain near-infrared spectroscopy detects fast neuronal signal in the motor cortex. *Neuroimage* 17, 1868–1875.

Wolf, U., Wolf, M., Toronov, V., Michalos, A. Paunescu, L. A., and Gratton, E. (2000). "Detecting cerebral functional slow and fast signals by frequency-domain near-infrared spectroscopy using two different sensors," in *Paper Presented at OSA Meeting in Optical Spectroscopy and Imaging and Photon Migration,* Miami, April 2–5, 2000.

Worsley, K. J., Evans, A. C., Marrett, S., and Neelin, P. (1992). A three-dimensional statistical analysis for CBF activation studies in human brain. *J. Cereb. Blood Flow Metab.* 12, 900–918.

Zhang, X., Toronov V. Y., Fabiani, M., Gratton, G. and Webb, A. G. (2005). The study of cerebral hemodynamic and neuronal response to visual stimulation using simultaneous NIR optical tomography and BOLD fMRI in humans. In K. E. Bartels, L. S. Bass, W. T. de Riese, K. W. Gregory, H. Hirschberg, A. Katzir, N. Kollias, S. J. Madsen, R. S. Malek, K. M. McNally-Heintzelman, L. P. Tate, Jr., E. A. Trowers, B. J.-F. Wong (Eds.). *Proc. SPIE Vol. 5686, Photonic Therapeutics and Diagnostics* (pp. 566–572),San Jose, California.

Zhu, W., Wang, Y., Yao, Y., Chang, J., Graber, H. L., and Barbour, R. L. (1997). Iterative total least-squares image reconstruction algorithm for optical tomography by the conjugate gradient method. *J. Opt. Soc. Am.* 14, 799–807.

**Conflict of Interest Statement:** The authors declare no conflict of interest. Specifically, they have no ownership or financial interests in ISS, Inc.

*Received: 28 April 2010; paper pending published: 19 May 2010; accepted: 02 June 2010; published online: 23 June 2010.*
*Citation: Gratton G and Fabiani M (2010) Fast optical imaging of human brain function. Front. Hum. Neurosci. **4**:52. doi: 10.3389/fnhum.2010.00052*

frontiers in
HUMAN NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 10 August 2010
doi: 10.3389/fnhum.2010.00065

# Mal-adaptation of event-related EEG responses preceding performance errors

***Heike Eichele[1], Hilde T. Juvodden[1], Markus Ullsperger[2,3] and Tom Eichele[1,4]****

[1] *Department of Biological and Medical Psychology, University of Bergen, Bergen, Norway*
[2] *Max Planck Institute for Neurological Research, Cologne, Germany*
[3] *Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, Netherlands*
[4] *Mind Research Network, Albuquerque, NM, USA*

***Edited by:***
*Michael X. Cohen, University of Amsterdam, Netherlands*

***Reviewed by:***
*James F. Cavanagh, Université Paris Descartes, France*
*Ali Mazaheri, University of California at Davis, USA*

****Correspondence:***
*Tom Eichele, Department of Biological and Medical Psychology, University of Bergen, Jonas Lies Vei 91, 5009 Bergen, Norway.*
*e-mail: tom.eichele@psybp.uib.no*

Recent EEG and fMRI evidence suggests that behavioral errors are foreshadowed by systematic changes in brain activity preceding the outcome by seconds. In order to further characterize this type of error precursor activity, we investigated single-trial event-related EEG activity from 70 participants performing a modified Eriksen flanker task, in particular focusing on the trial-by-trial dynamics of a fronto-central independent component that previously has been associated with error and feedback processing. The stimulus-locked peaks in the N2 and P3 latency range in the event-related averages showed expected compatibility and error-related modulations. In addition, a small pre-stimulus negative slow wave was present at erroneous trials. Significant error-preceding activity was found in local stimulus sequences with decreased conflict in the form of less negativity at the N2 latency (310–350 ms) accumulating across five trials before errors; concomitantly response times were speeding across trials. These results illustrate that error-preceding activity in event-related EEG is associated with the performance monitoring system and we conclude that the dynamics of performance monitoring contribute to the generation of error-prone states in addition to the more remote and indirect effects in ongoing activity such as posterior alpha power in EEG and default mode drifts in fMRI.

**Keywords: performance monitoring, errors, event-related EEG, independent component analysis, single trial analysis, deconvolution**

## INTRODUCTION

> *"It is hardly surprising to find that the organism's response to "identical" stimuli is in flux. The nervous system is not a passive recipient of inputs that are obediently switched to outputs; rather it is a dynamic system that continuously generates hypotheses about the environment"*
>
> (Squires et al., 1976).

It is trivial to state that human behavior and human brain activity are highly variable, yet more than thirty years after presenting the astonishing EEG data that led Squires and colleagues to write these lines, echoing James' prescient perception (James, 1890, the digitized text is freely available at http://www.archive.org/) the core assumption in cognitive neuroscience and its neuroimaging methods still maintains that there is a *deterministic* event-related signal and *random* noise. This assumption is used to justify averaging the rich and complex information in EEG and fMRI measurements for denoising and data reduction (in event-related EEG simply by taking the mean across trials, in fMRI through first-level modeling with fixed predictors). This happens, we assume, for analytical convenience and traditional reasons, although the average signal does not at all reflect a great deal of variability in the raw data (roughly about 10%), and is not necessarily representative of the single trials at which some behavior occurred (Arieli et al., 1996; Raichle and Snyder, 2007). The challenge for cognitive neuroscience in a time when processors are fast and memory is cheap is to utilize the many elegant methods for single-trial analysis that are (mostly freely) available and peek into the exciting dynamics of brain activity on a moment-to-moment basis. In order to make the case, we present an example from event-related EEG where a combination of data decomposition with independent component analysis (ICA), multiple regression and deconvolution is used to derive a heretofore unknown electrophysiological precursor of behavioral errors.

Since the observation of the medial frontal negativity in the response-locked event-related potential associated with error commission (Falkenstein et al., 1991; Gehring et al., 1993), brain activity patterns and behavioral changes that are caused by errors have received a great deal of attention and helped in understanding how the brain responds to errors, and shapes subsequent behavioral adaptation. A neural system located mainly within the rostral cingulate zone (RCZ), the pre-supplementary motor area (pre-SMA) and the anterior insular cortex has been identified to support this function by signaling the need for increased control, whenever the action goal is not achieved or the risk to fail is high (Ridderinkhof et al., 2004; Debener et al., 2005; Klein et al., 2007).

In contradistinction, changes in brain activity and behavior preceding errors have received much less attention, although these antecedent conditions may help to understand how brain states affect behavioral accuracy, and activity patterns preceding errors might lend themselves to prediction of upcoming performance, which would possibly have a number of interesting real-world applications. To date, only a limited amount of studies have investigated EEG activity immediately preceding errors; these have employed different EEG features for analysis and the evidence points to

different, but quite possibly interrelated sources of performance errors. Briefly summarized, in trials preceding errors event-related EEG shows more positive response-locked activity (Ridderinkhof et al., 2003; Allain et al., 2004; Hajcak et al., 2005), reduced theta power (Cavanagh et al., 2009), decreased amplitudes of stimulus-preceding contingent negative variation and the stimulus-following P300 components (O'Connell et al., 2009), while α and μ rhythm power increase (Mazaheri et al., 2009).

In addition, changes expanding on a longer timescale across trials before errors can be observed in behavioral, electrophysiological, and hemodynamic measures: responses are executed increasingly more quickly (Smith and Brewer, 1995; Gehring and Fencsik, 2001), frontal and insular regions associated with attention regulation and effort show a graded decline of hemodynamic activation (Eichele et al., 2008), while event-related responses in regions within the default mode network (Raichle et al., 2001) show reduced deactivation (Li et al., 2007; Eichele et al., 2008). Similarly, an increase in right parietally localized α-power in the scalp EEG evolves across the same timescale (O'Connell et al., 2009).

We have previously suggested that such trends in the trial-by-trial dynamics of hemodynamic activity preceding errors stem from a (mal-)adaptive system rather than spontaneous variability. As such, we posit that error-preceding activity more or less directly represents parts of the systems that mediate error monitoring and more general cognitive control functions (Ridderinkhof et al., 2004; Ullsperger and von Cramon, 2004). In the context of cognitive control models (Yeung et al., 2004; Yu et al., 2009) this account predicts that variations of the stimulus sequence that impact on conflict (and attention) are continuously employed to adapt the system, and thus are reflected in the dynamics of event-related responses generated by the system. We imply more generally that a representation of predictive information transferred in the stimulus history is continuously utilized to optimize brain responses (Sutton et al., 1965; Squires et al., 1976; Eichele et al., 2005; Friston, 2005; Raichle, 2006; Botvinick, 2007; Mars et al., 2008). This optimization yields adjustments of overt response speed and commensurate corticospinal excitability when events are repetitive and subjectively predictable (Huettel et al., 2002, 2005; Bestmann et al., 2008) and consequently affect the trade-off between speed and accuracy (Forstmann et al., 2008; Ivanoff et al., 2008; van Veen et al., 2008).

Here, we employed a modified Eriksen flanker task (Eriksen and Eriksen, 1974; Gratton et al., 1992) and the event-related responses in this task are modulated by conflict in particular during the latency of the N2 from about 250 ms after stimulus onset, extending well into the latency of the P3 up to around 500 ms (Kopp et al., 1996; Yeung et al., 2004; Folstein and Van Petten, 2008). Errors occur relatively more often to incompatible trials, in particular when these follow low conflict compatible trials and yield a large error-related negativity (ERN) in the response-locked EEG at about 80 ms with a subsequent positivity (Falkenstein et al., 2000; Debener et al., 2005).

The present study focuses on the dynamics of a medial frontal EEG source with a generator in the RCZ that previously has been associated with error and feedback processing (Roger et al., 2010; Debener et al., 2005; Gentsch et al., 2009), providing the scalp electrophysiological surrogate for the performance monitoring system. We aimed to show that this source reflects mal-adaptive event-related activity preceding erroneous performance at the time of conflict/compatibility processing, in particular we expect gradually lower amplitudes of event-related responses similar to our previous fMRI findings. This finding would link error precursors more directly to performance monitoring, rather than to unspecific ongoing activity producing spurious noise signals that lead to random lapses in performance.

In order to separate the medial frontal source from other event-related responses and background activity we employed temporal ICA (Bell and Sejnowski, 1995) and identified for each participant the source that jointly topographically and functionally reflected conflict and error responses (IE) and then analyzed the trial-by-trial by dynamics surrounding errors. Due to its role in (pre-response) conflict processing and temporal overlap with the ERN we expected in particular the N2 amplitude to show a gradually lesser negativity prior to errors. However, we inspected the entire time period surrounding stimulus onset to pick up trends localized to preceding responses (Ridderinkhof et al., 2003), as well as stimulus-preceding negativities and P3 (O'Connell et al., 2009).

## MATERIALS AND METHODS

### PARTICIPANTS

Seventy participants (29 male, 41 female) are included in this study and were recruited from psychology and medicine undergraduate classes at the University of Bergen. Written informed consent was obtained from all participants and the study was approved by Regional Committee for Medical Research Ethics, West-Norway. Included participants had normal, or corrected to normal vision, and no history of neurological or psychiatric disorders and present use of psychotropic medication. Eight of the participants were left-handed individuals. The mean age of the participants was 22.02 years (±3.01). Data from five additional participants were discarded due to excessive non-stereotyped EEG artifacts.

### EXPERIMENTAL DESIGN

After verbal and written instruction and a training sequence, participants performed a modified visual Eriksen flanker task implemented in E-prime 2 (Psychology software tools). At the center of a PC screen, participants were presented a fixation dot. Trials began with the presentation of six horizontal flanker-arrows appearing below the fixation. Participants were instructed to respond as fast as possible and as accurate as possible with either a left or a right mouse button press following the direction of a central target arrow that appeared 100 ms after the flankers. The central target arrow pointed either into the same direction as the flanker-arrows in compatible trials (<<<<<<<, >>>>>>>) or in the opposite direction in incompatible trials (<<<>>>>, >>><<<). The target and flanker-arrows remained on screen until a response was registered. Trials were terminated by the motor response and were followed by a fixed 800 ms interval before the onset of the next trial. Stimuli were presented in five blocks with 200 trials that were pseudo-randomized separately for each participant. No performance feedback was given during the experiment. The overall probability of compatible and incompatible trials, as well as left and right responses was kept at 0.5, respectively. Local probability manipulations of compatibility, response and stimulus-type were embedded in the stimulus sequences: twenty four participants received stimulus sequences in

which compatibility was parametrically varied in short sequences of 40 trials in which the ratio between the compatible and incompatible trials was 1/9, 3/7, 5/5, 7/3 or 9/1, respectively, while the probability of a left/right response was kept at 0.5. Twenty three participants received stimulus sequences in which response side was similarly varied in sequences of 40 trials with ratios between left and right of 1/9, 3/7, 5/5, 7/3 or 9/1, respectively, while conflict was kept at 0.5. Twenty three participant received sequences with sequences of 50 trials in which one of the trial type appeared at $p = 0.7$, and the other three stimuli at $p = 0.1$ respectively. Here, we only focus on error-preceding activity which is present in all three experiments, and will report the main effects of parametric manipulations elsewhere.

### TRIAL SEQUENCE

Errors in the flanker task are not randomly distributed but occur more frequently at high conflict incompatible trials. In order to visualize and analyze the relationship between errors and the preceding stimulus sequence the compatibility was coded in a binary vector containing all trials, where −1 denoted compatible, and +1 denoted incompatible trials, respectively. These were then used to derive the average occurrences from five trials prior to five trials after errors (see **Figure 4**, top panel).

### BEHAVIORAL ANALYSIS

Response time (RT) and response accuracy (RACC) averages were generated for all possible outcomes. Responses faster than 100 ms (0.03%) and slower than 1000 ms (0.48%) were not considered. Errors were defined as incorrect key presses to compatible and incompatible trials. The categorical effects of compatibility, stimulus/response repetition and accuracy were removed from the single-trial RTs by means of multiple linear regression (Notebaert and Verguts, 2007). The RACC data for each participant were coded as a binary vector with 0 representing correct outcomes, and 1 errors, and were then used to derive the residual modulation from five trials prior to five trials after error commission using the deconvolution method we have introduced for hemodynamic response estimation previously (Eichele et al., 2008). In brief, we take the Moore-Penrose pseudo-inverse of the convolution matrix containing the −5…+5 lagged versions of the accuracy predictor and multiply this with the RT/EEG vector yielding the residual modulation surrounding the error as the output. An illustration of the method with a simulation example is provided in **Figure 1**, for further details and areas of application see our previous work (Eichele et al., 2008, 2009).

Deconvolution enables inclusion of overlapping trial sequences with errors and is preferable to within-subject averaging here. The deconvolved RT modulation was tested with point-wise one-sample

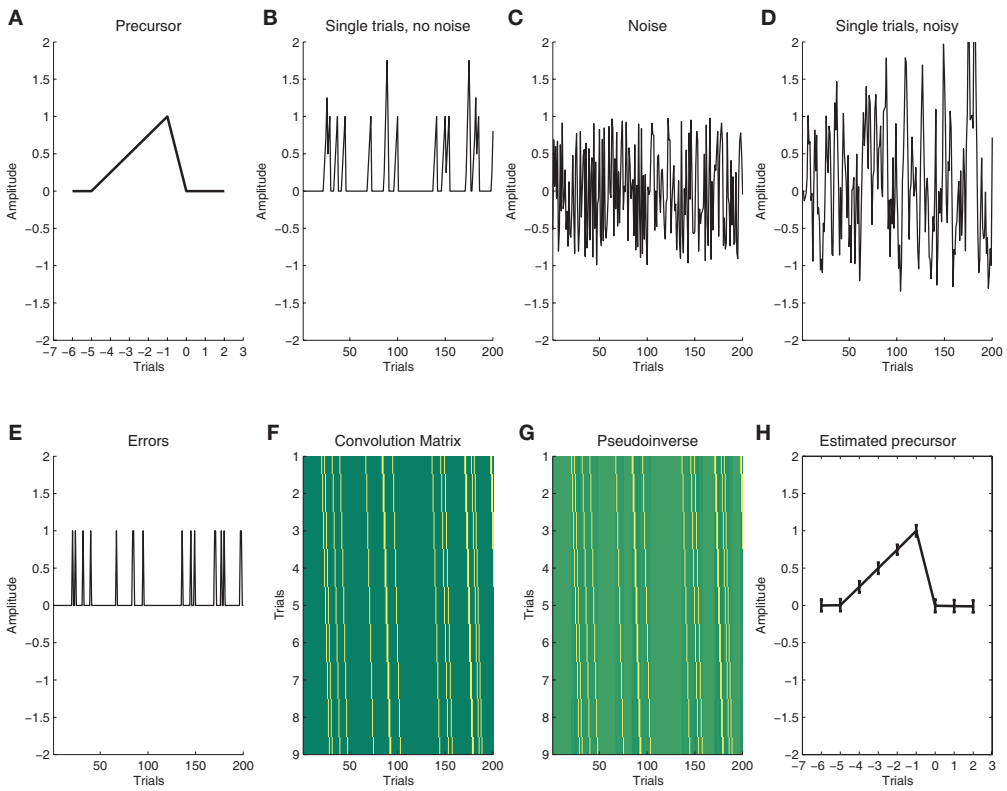



**FIGURE 1 | Illustration of deconvolution of latent sequential effects. (A)** We assume a latent precursor signal that gradually evolves across trials and precedes each error; **(B)** In order to illustrate a noiseless sequence of single trials we convolve a vector with 10% errors at random instances, note the summation of overlapping sequences; **(C)** for real data, we assume additional noise; **(D)** Convolution of the noisy data **(C+E)** with **(A)** yields the simulated single-trial data that are used for deconvolution; **(E)** In order to deconvolve the data, we use a "Stick"-function that describes the occurrence of errors; **(F)** stacking the stick function at different lags is the convolution matrix; **(G)** we then take the pseudo-inverse of **(F)** and multiply the resulting matrix with the noisy data **(D)**; **(H)** The product is an estimate of the precursor. In this figure, we show the average across 100 runs, error bars indicate the ±1 standard deviation around the mean.

*t*-tests against zero mean; in addition a linear slope was fitted to the five error-preceding trials using linear regression. The resulting scaling factors were also subjected to a one-sample *t*-test to provide "random" effects population inferences (**Figure 4**, middle panel).

### EEG ACQUISITION

EEGs were recorded continuously inside an electromagnetically and acoustically shielded chamber (Rainford EMC Systems, Wigan, UK) at 1-kHz sampling frequency (low cutoff at 0.1 Hz and a high cutoff at 250 Hz) with BrainAmp MR plus X2 amplifiers (BrainProducts, Munich, Germany). Participants were fitted with an elastic cap (Braincap, FMS, Falk Minow Services, Herrsching, Germany) containing 61 Ag/AgCl electrodes placed at Fp1, Fpz, Fp2, AF7, AF3, AFz, AF4, AF8, F7, F5, F3, F1, Fz, F2, F4, F6, F8, FT7, FC5, FC3, FC1, FCz, FC2, FC4, FC6, FT8, T7, C5, C3, C1, Cz, C2, C4, C6, T8, TP7, CP5, CP3, CP1, CPz, CP2, CP4, CP6, TP8, P7, P5, P3, P1, Pz, P2, P4, P6, P8, PO7, PO3, POz, PO4, PO8, O1, Oz, O2, TP9 and TP10. Vertical eye movements were monitored with a bipolar derivation between Fp1 and an additional electrode placed below the left eye. Additionally, ECG was monitored. Channels were referenced to TP9 with a ground on the right cheek and impedances were kept below 10 kΩ.

### EEG PROCESSING

The EEG data were offline re-referenced to common average reference, filtered from 0.5 to 45 Hz (12 db), and decimated to 500 Hz sampling rate. The data were then divided into epochs spanning from before the preceding response (−900 ms relative to target onset) to after the current response (+1100 ms after target onset), the mean value of the entire epoch served as baseline. Segments containing large, non-stereotyped artifacts with amplitudes exceeding ±300 μV on any of the channels were rejected, and padded with the average of adjacent trials. Hereafter, the 61 scalp channels from each dataset were subjected to temporal ICA using infomax (Bell and Sejnowski, 1995), implemented in EEGLAB (Delorme and Makeig, 2004), estimated 30 components after PCA compression. For an overview about ICA and its applications to neuroimaging data see (Makeig et al., 2004; Stone, 2004; Calhoun et al., 2009; Eichele et al., 2009). In order to cluster components for further analysis and identify the most relevant to conflict and error processing we used automated sorting routines. Firstly, correlation with spatial templates for blink and lateral eyes movements was used to identify ocular artifacts (Viola et al., 2009). Secondly, muscular and other artifacts mainly localized to single electrodes were identified through the spatial standard deviation of the topography and correlation with a spectral template. From the remaining components, one was extracted for each participant that best matched the expected fronto-central topography, with large amplitude error-related activity and a conditional difference between compatible and incompatible trials (Roger et al., 2010; Debener et al., 2005; Gentsch et al., 2009). In most participants this procedure yielded a single matching component. In one participant the topography did not match the spatial template. In about 10% of the sample, two or three components matched the criteria, and we selected the one with the largest correlation score. Topographies of the selected components were assigned a positive maximum, and the component activations were then back projected onto sensor level to recover the sign and amplitude of the scalp recorded potentials.

### INFERENCE

In order to test the consistency of topographies across participants, the channel weights were entered into one-sample *t*-tests under the assumption of zero magnitude. For visualization, topographies were also averaged, and a random sample of six individual topographies is provided in **Figure 2**.

Amplitude differences of individual component averages for compatible and incompatible correct (IC) trials were tested with paired *t*-tests at all time points to derive the latency and magnitude of the conflict effect; similarly, the difference between component averages for IC trials and error trials were subjected to paired *t*-tests to test for error-related responses.

In order to estimate error-preceding activity not predicted by the categorical effects of compatibility, stimulus/response repetition and accuracy, which themselves exert significant influences on the event-related response we removed the variability associated with these predictors from the single-trial EEG at each time point by means of multiple linear regression. In effect, the N2–P3 amplitude difference due to conflict, the error-specific response (ERN/Pe), and unspecific repetition priming (Mayr et al., 2003; Ullsperger et al., 2005) were subtracted from the data before further analysis. As for the behavioral data, the RACC vectors were used to deconvolve the residual event-related EEG modulation from five trials prior to five trials after error commission for the average amplitude in the latency window around the maximal N2 conflict effect (310–350 ms). The residual event-related modulation was tested with point-wise one-sample *t*-tests against zero mean; in addition a linear gradient was fitted to the five error-preceding trials using linear regression. The resulting scaling factors were also subjected to a one-sample *t*-test. Additional latency windows around stimulus onset, the early sensory evoked responses, P3 and post-response slow waves were also explored but yielded no significant residual effects preceding errors (data not shown).

Effects of all statistical tests were significant at an uncorrected *t*-threshold of $p < 0.002$ for two-tailed tests.

## RESULTS

### STIMULUS SEQUENCE

Eighty two percent of all errors occurred to incompatible trials, an expected outcome in the flanker task. Error-preceding trials showed an increased frequency of compatible trials (63%), the linear fit across five preceding trials resulted in an average 2% per trial increase of compatible trials ($t_{69} = -6.27$). Following erroneous responses, the probability of compatible/incompatible outcomes expectedly returns to 0.5. **Figure 4** (top panel) shows the modulation of compatibility surrounding error trials.

### BEHAVIOR

The RT and RACC results yielded a pattern typical for the Flanker task. Compatible correct (CC) responses were associated with fast RT (362 ms ± 0.62) and low error rate (2.0% ± 0.02), while incompatible trials showed slower RT (461 ms ± 0.72) and more frequent errors (9.2% ± 0.08) with fast RT (309 ms ± 0.89). Residual RTs were on average 11 ms (±2.42), faster in trials immediately prior to errors ($t_{69} = -4.76$). Across the five error-preceding trials, the linear fit indicated average speeding by 2.5 ms (±0.56) per trial ($t_{69} = -4.51$). After error trials post-error slowing of responses by 42 ms (±6.15) was

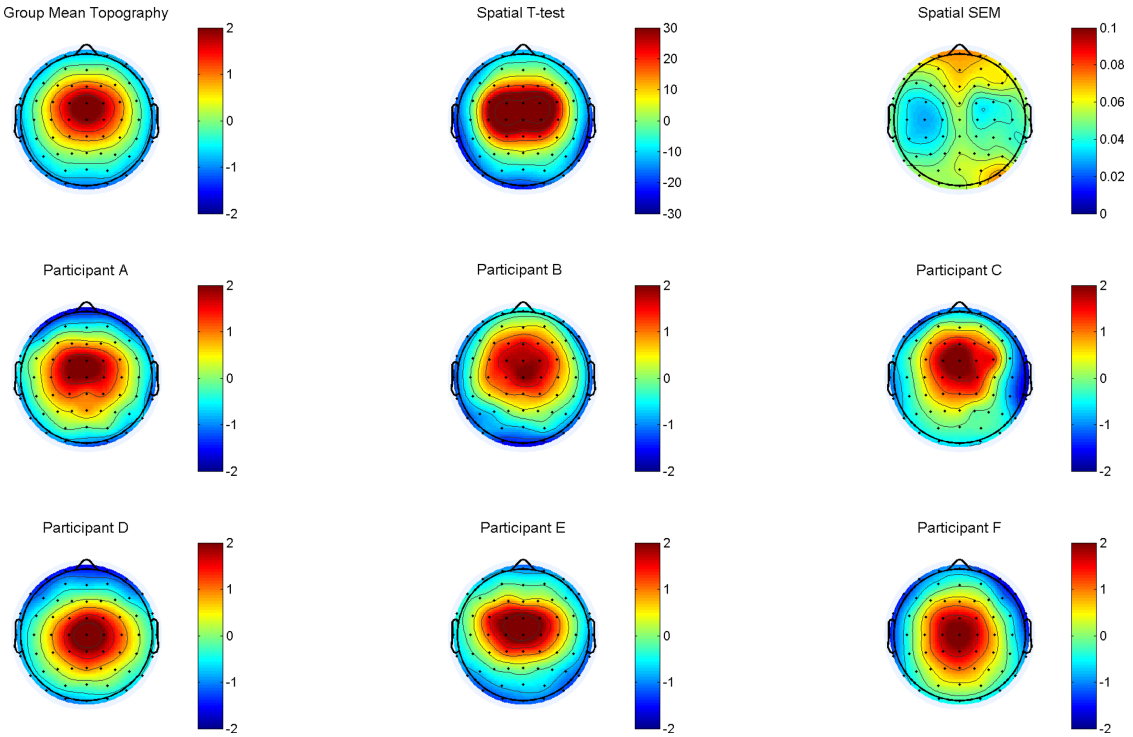



**FIGURE 2 | Top row, left: the group average topography, scaled from −2 to +2 μV shows a clear maximum at FCz, extending to the neighboring electrodes FC, FC2, C1, Cz and C2.** The spatial T-map (top row, middle) and the corresponding map of the standard error of the mean (SEM) illustrate the robustness of the selected topography in the sample. The middle and bottom rows (participants A–F) provide six randomly drawn single subject replications.

observed ($t_{69} = 6.87$). In subsequent trials sustained slowing on the order of 15 ms was present (all trials $p < 0.001$). The RT modulation around errors is shown in **Figure 4** in the middle panel.

### EEG

The component topographies were fairly stable across participants, with an average maximum at FCz (2.31 ± 0.06, $t_{69} = 40.97$); the average of the spatial correlation between all individual topographies was 0.83.

In the component time courses the largest differences between CC and IC responses were seen during the latency of the N2 at 316 ms post stimulus with a 0.95 μV more negative peak in incompatible trials ($t_{69} = -7.12$); at 430 ms the P3 peak was 1.21 μV larger ($t_{69} = 6.96$). Error trials yielded a large negativity at 350 ms post stimulus (−3.76 μV, $t_{69} = -12.01$), corresponding to the response-locked ERN, and a subsequent positivity at 546 ms (3.54 μV, $t_{69} = 12.71$). In addition, we observed a small negativity (−0.40 μV, $t_{69} = -5.29$) that preceded stimulus onsets with subsequent IE by approximately 200–300 ms, and which sustained through the early P1 and N1 peaks of the component average.

These effects are visible in the independent component event-related potentials (ICERPs) in the bottom sections of **Figure 3**, to the left separately for CC, IC, and IE, to the right as difference waves.

In addition to ICERPs we visualized the dependency of the stimulus-locked response on RT by sorting the epochs by RT (**Figure 3**, top left, gray line represents the average sorted RT), and averaging the resulting single-trial image across participants. For orientation, we also overlaid the relative frequency of incompatible trials (blue line), and errors (red line). The figure illustrates in particular N2 and P3 preceding the response in mainly correct trials with slower RT (top half of the image), and ERN and PE following the response at error trials (bottom quarter of the image). Correspondingly, the top right section of the figure presents the average RACC across the participants when the data are sorted by the component amplitude at each time point and shows higher than mean error rates in red, and lower than mean error rates in blue. The dominant feature in the figure is the scaling of error rate with increasing negative amplitudes during the N2/ERN latency and during the pre-stimulus negativity, and inversely during the P3/Pe latency range.

The residual amplitude at the N2 latency was on average 0.34 μV (±0.08) more positive in trials immediately prior to errors ($t_{69} = 4.11$). Correspondingly, a positive-going trend of 0.07 μV (±0.02) was estimated across the five error-preceding trials ($t_{69} = 2.99$). Following errors, N2 estimates returned to baseline. The EEG modulation around errors is shown in **Figure 4** in the bottom panel.

## DISCUSSION

This study addressed the question whether errors are associated with antecedent (mal-)adaptation of the performance monitoring system effective during conflict processing indexed by the N2 amplitude modulation in a medial frontal independent source (Roger et al., 2010; Debener et al., 2005; Gentsch et al., 2009).

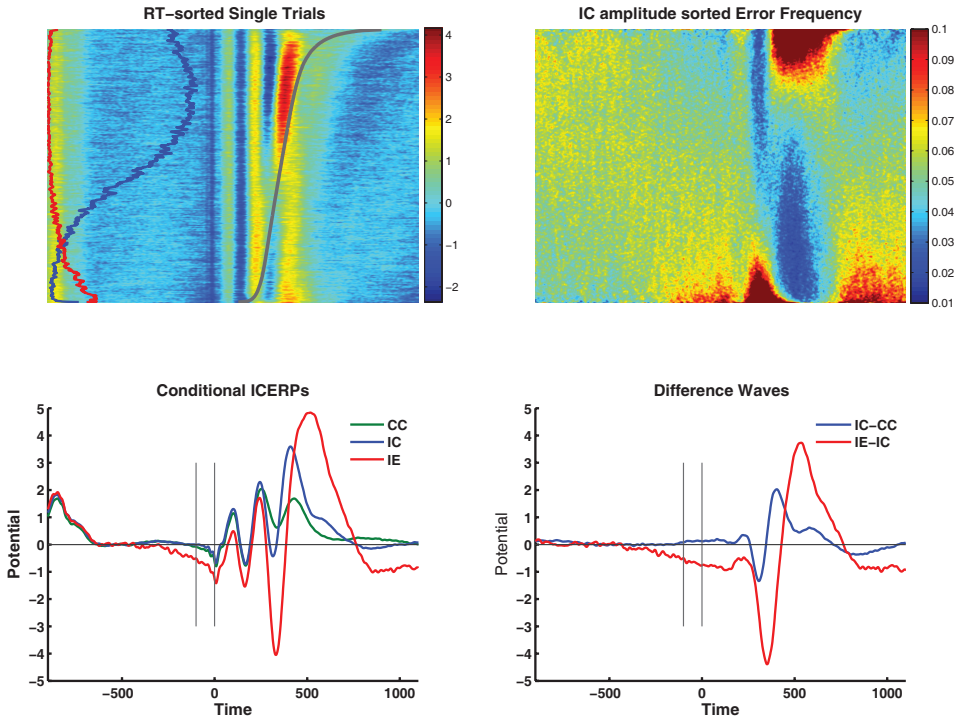



**FIGURE 3 | Top left: Group average RT-sorted single-trial ICERP image showing the dependency of the stimulus-locked response on RT (gray line).** The relative frequency of incompatible trials is given on the left of the image in blue, the proportion of errors in red. The figure illustrates in particular N2 and P3 preceding the response in mainly correct trials with slower RT (top half of the image), and ERN and PE following the response at error trials (bottom quarter of the image). Top right: Shows the average accuracy across the group sorted by the component amplitude at each time point and shows higher error rates in red, and lower than error rates in blue (equivalent to "vincentizing"). The dominant feature in the figure is the scaling of error rate with increasing negative amplitudes during the N2/ERN latency and during the pre-stimulus negativity, and inversely during the P3/Pe latency range. Bottom left: conditional ICERPs for compatible correct (CC, green), incompatible correct (IC, blue), and error responses (IE, red). Bottom right: difference waves for incompatible correct *minus* compatible correct (blue) and incompatible errors *minus* incompatible correct (red).

In order to *unmix* the scalp EEG correlates of performance monitoring from unrelated event-related responses and background rhythms we employed blind source separation with infomax ICA in the time domain and used correlation-based clustering of individual component topographies and time courses to identify this source across a large sample of individual datasets. From inspection and statistical analysis of the data presented in **Figures 1 and 2** we assume that the decomposition allows focusing selectively on the best surrogate for performance monitoring and affords a single-trial analysis of this source without latent confounds from other concurrent processes (Makeig et al., 2004; Onton et al., 2006), thus the error-preceding activity changes reported here should be largely unrelated to lapses that relate to spontaneous fluctuations of intrinsic activity.

The ICERPs showed a typical pattern of results with an N2–P3 conflict effect (Forster et al., 2010; Folstein and Van Petten, 2008) and a large (albeit stimulus-locked) ERN/Pe to error trials (Falkenstein et al., 1991; Gehring et al., 1993). Of note, erroneous trials showed a sustained negativity that significantly deviated from baseline already before flanker/target onset (**Figure 3**, bottom left), suggesting that (premature) error commission indeed starts before the stimulus has arrived. In terms of topography and timing this negativity may be related to preparatory potentials such as the (typically much larger) contingent negative variation that have sources in the medial walls of the frontal lobes (Nagai et al., 2004). Alternatively, this deflection may represent asymmetric oscillations around stimulus onset (Mazaheri and Jensen, 2008), possibly in the theta-range which also has been associated with this topography during task processing (Onton et al., 2005) and relaxed resting (Scheeringa et al., 2008). Although a long trend across trials was not present this negativity deserves further study.

The results shown in **Figure 4** yield three error-preceding phenomena. First: an increasing frequency of compatible trials prior to mainly incompatible errors. Second: speeding of RTs similar to previous reports (Smith and Brewer, 1995; Gehring and Fencsik, 2001). Third: a positive-going trend in the N2 latency range of the IC event-related response. RT and N2 effects were present in averages and deconvolution output estimated from the raw data (data not shown), and remained significant in the residual activity after removal of confounding effects, that is, the error-preceding activity cannot be solely accounted for by skewed averaging of compatible/incompatible trials or stimulus repetition priming.

What then is the mechanism behind error precursors? Predictive information conveyed by the stimulus history triggers neural and consequently behavioral adaptation. Local sequences with increased occurrence of a particular stimulus reduce the surprise

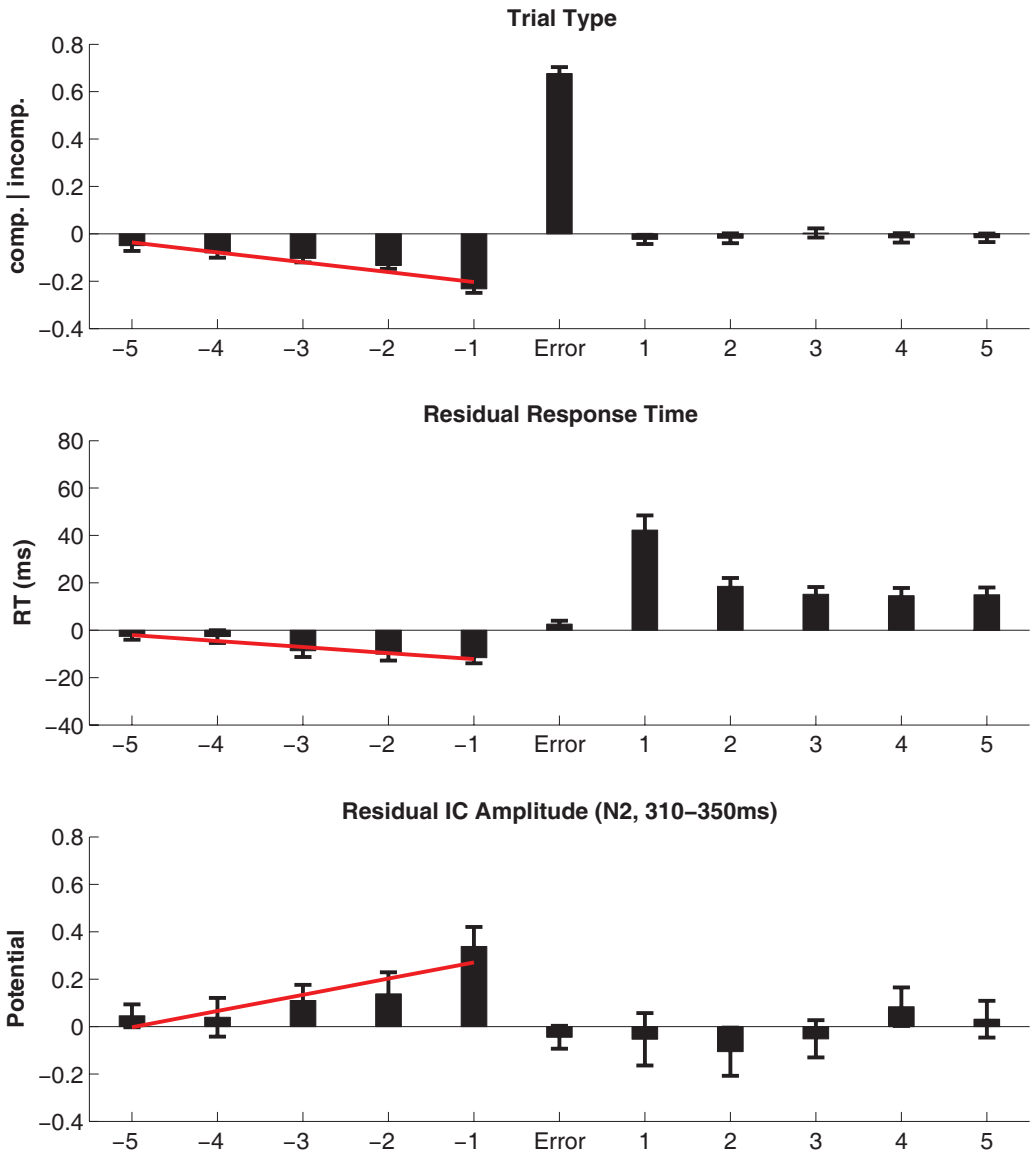



**FIGURE 4 | Top panel: Most errors occurred to incompatible trials, and error-preceding trials showed an increased frequency of compatible trials.** Following errors, the probability of compatible/incompatible outcomes is equal (as it is across the entire experiment). Middle panel: Residual RTs were on average 11 ms faster in trials immediately prior to errors and were speeding by 2.5 ms per trial from –5 to –1. Post-error slowing of responses by 42 ms was observed and in subsequent trials sustained slowing on the order of 15 ms was present. Bottom panel: The residual amplitude at the N2 latency was less negative in trials prior to errors and a trend was present across the five error-preceding trials. Following errors, N2 estimates returned to baseline. Note that, in contradistinction to **Figure 3**, conflict-related N2 modulations as well as the error-related amplitude increase equivalent to the ERN/N2 have been removed by multiple linear regression.

about the more frequent stimulus, reduce the local entropy in the stimulus sequence and thus increase the subjective predictability of upcoming stimuli (Huettel et al., 2002; Bestmann et al., 2008; Mars et al., 2008) and the possibility for "strategic modulation" (Gratton et al., 1992). As alluded to by the opening quote from the work of Squires and colleagues (Squires et al., 1976), generating predictions about upcoming stimuli/responses is a pervasive and automatic process that permits adaptive, optimized, fast and accurate responding to upcoming stimuli and has been shown on many levels of the cortical hierarchy (Squires et al., 1976; Llinas, 2001; Huettel et al., 2002; Eichele et al., 2005; Friston, 2005; Raichle, 2006; Bestmann et al., 2008). In the special case where low conflict compatible trials accumulate in the flanker task the mechanism becomes mal-adaptive, in the sense that RTs are speeding while conflict-monitoring-driven recruitment of effort by activity in the medial frontal wall is lowered concurrently, yielding a more error-prone trade-off between speed and accuracy (Forstmann et al., 2008; Ivanoff et al., 2008; van Veen et al., 2008). According to current theories of cognitive control the performance monitoring system provides signals for adaptive optimization of goal-directed

behavior and signals the need for adjustments required in responses after errors and, more generally, whenever the action outcome is at risk (Botvinick et al., 2004; Ridderinkhof et al., 2004; Rushworth et al., 2007). Here, we show that the basic function of this system makes it equally susceptible to mal-adaptive over-optimization in the presence of random event series where local sequences convey low conflict and low uncertainty about future outcomes (Botvinick, 2007; Eichele et al., 2008). This is in line with the notion that the posterior medial frontal cortex plays a major role in effort-based cost-benefit valuations in humans (Croxson et al., 2009) and other animals (Schweimer et al., 2005; Rudebeck et al., 2006). In this way the gradual decline in N2 amplitude across trials preceding errors appears to reflect a reduction of attentional effort, which by itself is driven by increasing predictability.

Additionally, we should note that components in the N2 latency range serve as a match/mismatch detectors that are reduced by stimulus predictability and increased by surprise/prediction error (Eichele et al., 2005; Jongsma et al., 2006) in a similar way as the later P3 (Sutton et al., 1965; Mars et al., 2008), such that conflict-related and mismatch-related subcomponents, if represented together in the source analyzed here, might be reduced jointly (Forster et al., 2010; Folstein and Van Petten, 2008).

Considering the present results together with recent literature, it appears that apart from behavioral changes (Smith and Brewer, 1995; Gehring and Fencsik, 2001) and regional changes observed prior to errors in fMRI (Li et al., 2007; Eichele et al., 2008) there are a variety of features in the EEG signal that provide predictive information about upcoming RACC (Hajcak et al., 2005; Padilla et al., 2006; Cavanagh et al., 2009; Mazaheri et al., 2009; O'Connell et al., 2009). Future studies should clarify whether the diverse phenomenology describes aspects of the same underlying functional system or reflects multiple concurrent effects. For error prediction in real-world situations, it would be desirable to understand the effects of performance (mal-) adaptation on event-related and induced activities as well as connectivity to have multiple redundant features available that can be employed for classification.

In summary, we note that a combination of data mining and modeling on a trial-by-trial level helps to better understand the dynamics of the cognitive process under investigation and we hope this approach to analysis and inference becomes the rule rather than the exception in cognitive neuroscience and neuroimaging (Makeig et al., 2004; Debener et al., 2006; Onton et al., 2006). In this particular case we show that one antecedent of errors is indeed generated by the branch of the cognitive control system that is in charge of monitoring them.

## ACKNOWLEDGMENTS

This paper provides a detail report of some of the data presented by the authors at the Opinions and Discussions on Cognitive Neuroscience symposium held in Amsterdam (10/26-10/27 2009). The work was supported by a Nevronor grant from the Norwegian Research council (189661) to Tom Eichele. Tom Eichele and Markus Ullsperger conceived the experiments; Heike Eichele, Hilde T Juvodden and undergraduate RA's collected the data; Heike Eichele and Tom Eichele analyzed the data; Heike Eichele, Hilde T Juvodden, Markus Ullsperger and Tom Eichele co-wrote the paper. We thank our undergraduate research assistants Aleksander H. Erga, Eva Sunde, Katrine V. Granli, Christer Lunde Gjerstad, Ingrid Langøen Bull-Tornøe, Agnethe Smith Wiker, Pia E. Lyse, Branca Vojnovic, Camilla Lindgren and Marius Kalsås Worren for their help with data collection. We are indebted to Marco Steinhauser for helpful discussions. The data and analysis functions for trial sequence analysis are available from the corresponding author upon request.

## REFERENCES

Allain, S., Carbonnell, L., Falkenstein, M., Burle, B., and Vidal, F. (2004). The modulation of the Ne-like wave on correct responses foreshadows errors. *Neurosci. Lett.* 372, 161–166.

Arieli, A., Sterkin, A., Grinvald, A., and Aertsen, A. (1996). Dynamics of ongoing activity: explanation of the large variability in evoked cortical responses. *Science* 273, 1868–1871.

Bell, A. J., and Sejnowski, T. J. (1995). An information-maximization approach to blind separation and blind deconvolution. *Neural. Comput.* 7, 1129–1159.

Bestmann, S., Harrison, L. M., Blankenburg, F., Mars, R. B., Haggard, P., Friston, K. J., and Rothwell, J. C. (2008). Influence of uncertainty and surprise on human corticospinal excitability during preparation for action. *Curr. Biol.* 18, 775–780.

Botvinick, M. M. (2007). Conflict monitoring and decision making: reconciling two perspectives on anterior cingulate function. *Cogn. Affect. Behav. Neurosci.* 7, 356–366.

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

Calhoun, V. D., Liu, J., and Adali, T. (2009). A review of group ICA for fMRI data and ICA for joint inference of imaging, genetic, and ERP data. *Neuroimage* 45:S163–S172.

Cavanagh, J. F., Cohen, M. X., and Allen, J. J. (2009). Prelude to and resolution of an error: EEG phase synchrony reveals cognitive control dynamics during action monitoring. *J. Neurosci.* 29, 98–105.

Croxson, P. L., Walton, M. E., O'Reilly, J. X., Behrens, T. E., and Rushworth, M. F. (2009). Effort-based cost-benefit valuation and the human brain. *J. Neurosci.* 29, 4531–4541.

Debener, S., Ullsperger, M., Siegel, M., and Engel, A. K. (2006). Single-trial EEG-fMRI reveals the dynamics of cognitive function. *Trends Cogn. Sci.* 10, 558–563.

Debener, S., Ullsperger, M., Siegel, M., Fiehler, K., von Cramon, D. Y., and Engel, A. K. (2005). Trial-by-trial coupling of concurrent electroencephalogram and functional magnetic resonance imaging identifies the dynamics of performance monitoring. *J. Neurosci.* 25, 11730–11737.

Delorme, A., and Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* 134, 9–21.

Eichele, T., Calhoun, V. D., and Debener, S. (2009). Mining EEG-fMRI using independent component analysis. *Int. J. Psychophysiol.* 73, 53–61.

Eichele, T., Debener, S., Calhoun, V. D., Specht, K., Engel, A. K., Hugdahl, K., von Cramon, D. Y., and Ullsperger, M. (2008). Prediction of human errors by maladaptive changes in event-related brain networks. *Proc. Natl. Acad. Sci. U.S.A.* 105, 6173–6178.

Eichele, T., Specht, K., Moosmann, M., Jongsma, M. L., Quiroga, R. Q., Nordby, H., and Hugdahl, K. (2005). Assessing the spatiotemporal evolution of neuronal activation with single-trial event-related potentials and functional MRI. *Proc. Natl. Acad. Sci. U.S.A.* 102, 17798–17803.

Eriksen, B. A., and Eriksen, C. W. (1974). Effects of noise letters upon identification of a target letter in a nonsearch task. *Percept. Psychophys.* 16, 143–149.

Falkenstein, M., Hohnsbein, J., Hoormann, J., and Blanke, L. (1991). Effects of crossmodal divided attention on late ERP components. II. Error processing in choice reaction tasks. *Electroencephalogr. Clin. Neurophysiol.* 78, 447–455.

Falkenstein, M., Hoormann, J., Christ, S., and Hohnsbein, J. (2000). ERP components on reaction errors and their functional significance: a tutorial. *Biol. Psychol.* 51, 87–107.

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.

Forster, S. E., Carter, C. S., Cohen, J. D., and Cho, R. Y. (2010). Parametric manipulation of the conflict signal and control-state adaptation. *J. Cogn. Neurosci.* (inpress).

Forstmann, B. U., Dutilh, G., Brown, S., Neumann, J., von Cramon, D. Y., Ridderinkhof, K. R., and Wagenmakers, E. J. (2008). Striatum and pre-SMA facilitate decision-making under time pressure. *Proc. Natl. Acad. Sci. U.S.A.* 105, 17538–17542.

Friston, K. J. (2005). A theory of cortical responses. *Philos. Trans. R Soc. Lond. B Biol. Sci.* 360, 815–836.

Gehring, W. J., and Fencsik, D. E. (2001). Functions of the medial frontal cortex in the processing of conflict and errors. *J. Neurosci.* 21, 9430–9437.

Gehring, W. J., Goss, B., Coles, M. G. H., Meyer, D. E., and Donchin, E. (1993). A neural system for error-detection and compensation. *Psychol. Sci.* 4, 385–390.

Gentsch, A., Ullsperger, P., and Ullsperger, M. (2009). Dissociable medial frontal negativities from a common monitoring system for self- and externally caused failure of goal achievement. *Neuroimage* 47, 2023–2030.

Gratton, G., Coles, M. G., and Donchin, E. (1992). Optimizing the use of information: strategic control of activation of responses. *J. Exp. Psychol. Gen.* 121, 480–506.

Hajcak, G., Nieuwenhuis, S., Ridderinkhof, K. R., and Simons, R. F. (2005). Error-preceding brain activity: robustness, temporal dynamics, and boundary conditions. *Biol. Psychol.* 70, 67–78.

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

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

Ivanoff, J., Branning, P., and Marois, R. (2008). fMRI evidence for a dual process account of the speed-accuracy tradeoff in decision-making. *PLoS ONE* 3, e2635. doi:10.1371/journal.pone.0002635.

James, W. (1890). *The Principles of Psychology*. New York: Holt.

Jongsma, M. L., Eichele, T., Van Rijn, C. M., Coenen, A. M., Hugdahl, K., Nordby, H., and Quiroga, R. Q. (2006). Tracking pattern learning with single-trial event-related potentials. *Clin. Neurophysiol.* 117, 1957–1973.

Klein, T. A., Endrass, T., Kathmann, N., Neumann, J., von Cramon, D. Y., and Ullsperger, M. (2007). Neural correlates of error awareness. *Neuroimage* 34, 1774–1781.

Kopp, B., Rist, F., and Mattler, U. (1996). N200 in the flanker task as a neurobehavioral tool for investigating executive control. *Psychophysiology* 33, 282–294.

Li, C. S., Yan, P., Bergquist, K. L., and Sinha, R. (2007). Greater activation of the "default" brain regions predicts stop signal errors. *Neuroimage* 38, 640–648.

Llinas, R. R. (2001). *I of the Vortex: From Neurons to Self*. Cambridge: The MIT Press.

Makeig, S., Debener, S., Onton, J., and Delorme, A. (2004). Mining event-related brain dynamics. *Trends Cogn. Sci.* 8, 204–210.

Mars, R. B., Debener, S., Gladwin, T. E., Harrison, L. M., Haggard, P., Rothwell, J. C., and Bestmann, S. (2008). Trial-by-trial fluctuations in the event-related electroencephalogram reflect dynamic changes in the degree of surprise. *J. Neurosci.* 28, 12539–12545.

Mayr, U., Awh, E., and Laurey, P. (2003). Conflict adaptation effects in the absence of executive control. *Nat. Neurosci.* 6, 450–452.

Mazaheri, A., and Jensen, O. (2008). Asymmetric amplitude modulations of brain oscillations generate slow evoked responses. *J. Neurosci.* 28, 7781–7787.

Mazaheri, A., Nieuwenhuis, I. L., van Dijk, H., and Jensen, O. (2009). Prestimulus alpha and mu activity predicts failure to inhibit motor responses. *Hum. Brain Mapp*. 30, 1791–1800.

Nagai, Y., Critchley, H. D., Featherstone, E., Fenwick, P. B., Trimble, M. R., and Dolan, R. J. (2004). Brain activity relating to the contingent negative variation: an fMRI investigation. *Neuroimage* 21, 1232–1241.

Notebaert, W., and Verguts, T. (2007). Dissociating conflict adaptation from feature integration: a multiple regression approach. *J. Exp. Psychol. Hum. Percept. Perform.* 33, 1256–1260.

O'Connell, R. G., Dockree, P. M., Robertson, I. H., Bellgrove, M. A., Foxe, J. J., and Kelly, S. P. (2009). Uncovering the neural signature of lapsing attention: electrophysiological signals predict errors up to 20 s before they occur. *J. Neurosci.* 29, 8604–8611.

Onton, J., Delorme, A., and Makeig, S. (2005). Frontal midline EEG dynamics during working memory. *Neuroimage* 27, 341–356.

Onton, J., Westerfield, M., Townsend, J., and Makeig, S. (2006). Imaging human EEG dynamics using independent component analysis. *Neurosci. Biobehav. Rev.* 30, 808–822.

Padilla, M. L., Wood, R. A., Hale, L. A., and Knight, R. T. (2006). Lapses in a prefrontal-extrastriate preparatory attention network predict mistakes. *J. Cogn. Neurosci.* 18, 1477–1487.

Raichle, M. E. (2006). Neuroscience. The brain's dark energy. *Science* 314, 1249–1250.

Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., and Shulman, G. L. (2001). A default mode of brain function. *Proc. Natl. Acad. Sci. U.S.A.* 98, 676–682.

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; discussion 1097–1089.

Ridderinkhof, K. R., Nieuwenhuis, S., and Bashore, T. R. (2003). Errors are foreshadowed in brain potentials associated with action monitoring in cingulate cortex in humans. *Neurosci. Lett.* 348, 1–4.

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

Roger, C., Benar, C. G., Vidal, F., Hasbroucq, T., and Burle, B. (2010). Rostral Cingulate Zone and correct response monitoring: ICA and source localization evidences for the unicity of correct- and error-negativities. *NeuroImage* 51, 391–403.

Rudebeck, P. H., Walton, M. E., Smyth, A. N., Bannerman, D. M., and Rushworth, M. F. (2006). Separate neural pathways process different decision costs. *Nat. Neurosci.* 9, 1161–1168.

Rushworth, M. F., Buckley, M. J., Behrens, T. E., Walton, M. E., and Bannerman, D. M. (2007). Functional organization of the medial frontal cortex. *Curr. Opin. Neurobiol.* 17, 220–227.

Scheeringa, R., Bastiaansen, M. C., Petersson, K. M., Oostenveld, R., Norris, D. G., and Hagoort, P. (2008). Frontal theta EEG activity correlates negatively with the default mode network in resting state. *Int. J. Psychophysiol.* 67, 242–251.

Schweimer, J., Saft, S., and Hauber, W. (2005). Involvement of catecholamine neurotransmission in the rat anterior cingulate in effort-related decision making. *Behav. Neurosci.* 119, 1687–1692.

Smith, G. A., and Brewer, N. (1995). Slowness and age: speed-accuracy mechanisms. *Psychol. Aging* 10, 238–247.

Squires, K. C., Wickens, C., Squires, N. K., and Donchin, E. (1976). The effect of stimulus sequence on the waveform of the cortical event-related potential. *Science* 193, 1142–1146.

Stone, J. V. (2004). *Independent Component Analysis: A Tutorial Introduction*. Cambridge, MA: MIT press.

Sutton, S., Braren, M., Zubin, J., and John, E. R. (1965). Evoked-potential correlates of stimulus uncertainty. *Science* 150, 1187–1188.

Ullsperger, M., Bylsma, L. M., and Botvinick, M. M. (2005). The conflict adaptation effect: it's not just priming. *Cogn. Affect. Behav. Neurosci.* 5, 467–472.

Ullsperger, M., and von Cramon, D. Y. (2004). Neuroimaging of performance monitoring: error detection and beyond. *Cortex* 40, 593–604.

van Veen, V., Krug, M. K., and Carter, C. S. (2008). The neural and computational basis of controlled speed-accuracy tradeoff during task performance. *J. Cogn. Neurosci.* 20, 1952–1965.

Viola, F. C., Thorne, J., Edmonds, B., Schneider, T., Eichele, T., and Debener, S. (2009). Semi-automatic identification of independent components representing EEG artifact. *Clin. Neurophysiol.* 120, 868–877.

Yeung, N., Botvinick, M. M., and Cohen, J. D. (2004). The neural basis of error detection: conflict monitoring and the error-related negativity. *Psychol. Rev.* 111, 931–959.

Yu, A. J., Dayan, P., and Cohen, J. D. (2009). Dynamics of attentional selection under conflict: toward a rational Bayesian account. *J. Exp. Psychol. Hum. Percept. Perform.* 35, 700–717.

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

*Received: 18 May 2010; paper pending published: 02 June 2010; accepted: 21 July 2010; published online: 10 August 2010.*

*Citation: Eichele H, Juvodden HT, Ullsperger M and Eichele T (2010) Mal-adaptation of event-related EEG responses preceding performance errors. Front. Hum. Neurosci.* ***4****:65. doi: 10.3389/fnhum.2010.00065*

frontiers in
HUMAN NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 08 October 2010
doi: 10.3389/fnhum.2010.00181

# Individuating faces and common objects produces equal responses in putative face-processing areas in the ventral occipitotemporal cortex

***Frank Haist[1,2]*, Kang Lee[3,4] and Joan Stiles[2,5]***

[1] *Department of Psychiatry, University of California, San Diego, La Jolla, CA, USA*
[2] *Center for Human Development, University of California, San Diego, La Jolla, CA, USA*
[3] *Institute of Child Study, University of Toronto, Toronto, ON, Canada*
[4] *Department of Psychology, University of California, San Diego, La Jolla, CA, USA*
[5] *Department of Cognitive Science, University of California, San Diego, La Jolla, CA, USA*

***Edited by***
*Michael X. Cohen, University of Amsterdam, Netherlands*

***Reviewed by:***
*Christine Schiltz, Université Catholique de Louvain, Belgium*
*Romke Rouw, University of Amsterdam, Netherlands*

***Correspondence:**
*Frank Haist, Developmental Neuroimaging Lab, Department of Psychiatry and Center for Human Development, University of California, San Diego, 9500 Gilman Drive, MC 0115, La Jolla, CA 92093-0115, USA.*
*e-mail: fhaist@ucsd.edu*

Controversy surrounds the proposal that specific human cortical regions in the ventral occipitotemporal cortex, commonly called the fusiform face area (FFA) and occipital face area (OFA), are *specialized* for face processing. Here, we present findings from an fMRI study of identity discrimination of faces and objects that demonstrates the FFA and OFA are equally responsive to processing stimuli at the level of individuals (i.e., individuation), be they human faces or non-face objects. The FFA and OFA were defined via a passive viewing task as regions that produced greater activation to faces relative to non-face stimuli within the middle fusiform gyrus and inferior occipital gyrus. In the individuation task, participants judged whether sequentially presented images of faces, diverse objects, or wristwatches depicted the identical or a different exemplar. All three stimulus types produced equivalent BOLD activation within the FFA and OFA; that is, there was no face-specific or face-preferential processing. Critically, individuation processing did not eliminate an object superiority effect relative to faces within a region more closely linked to object processing in the lateral occipital complex (LOC), suggesting that individuation processes are reasonably specific to the FFA and OFA. Taken together, these findings challenge the prevailing view that the FFA and OFA are face-specific processing regions, demonstrating instead that they function to individuate – i.e., identify specific individuals – within a category. These findings have significant implications for understanding the function of brain regions widely believed to play an important role in social cognition.

**Keywords: fusiform face area, face processing, functional MRI, visual processing, occipital face area**

## INTRODUCTION

As a key element in social cognition, visual face processing receives extensive study within cognitive neuroscience. This research has led to considerable debate about the function of two particular cortical regions located in the ventral occipitotemporal lobe (VOT), the so-called fusiform face area (FFA; Kanwisher et al., 1997) and the occipital face area (OFA; Clark et al., 1996; Gauthier et al., 2000; Haxby et al., 2000), although most of the debate has focused on the role of the FFA (Kanwisher and Yovel, 2006). There are two primary threads in this debate. One asks whether brain activity to stimulus input reflects a distributed representation of features (Haxby et al., 2001; Hanson et al., 2004; O'Toole et al., 2005) or a modular, category specific representation (Kanwisher et al., 1997; Yovel and Kanwisher, 2004; Kanwisher and Yovel, 2006). A second asks whether brain areas respond to specific stimulus inputs (Kanwisher et al., 1997; Haxby et al., 2001; Hanson et al., 2004; Yovel and Kanwisher, 2004; O'Toole et al., 2005; Kanwisher and Yovel, 2006) or to processing demands and operations (Tarr and Gauthier, 2000). Here, we present findings relevant to this second line of debate demonstrating that the FFA and OFA respond to a specific type of visual processing operation and thus are not just face-specific processing regions.

Neuromodular accounts posit that biologically dedicated regions of cortex respond selectively to specific categories of input such as faces, objects, or spatial locations (Kanwisher et al., 1997; Yovel and Kanwisher, 2004; Kanwisher and Yovel, 2006). Functional magnetic resonance imaging (fMRI) studies provide support for this hypothesis with numerous demonstrations that faces produce greater brain activations compared to other categories of visual stimuli in both the FFA and OFA, particularly when observers view the stimuli passively or covertly (Puce et al., 1995; Kanwisher et al., 1997; Aguirre et al., 1998; Downing et al., 2001). In contrast, studies of patients with severe deficits specific to facial recognition, known as prosopagnosia, challenge the proposal that face selectivity in the FFA is necessary for overt face recognition. Specifically, patients with prosopagnosia with lesions that spare the middle fusiform gyrus can produce normal FFA regions in terms of intensity of blood oxygenation level-dependent (BOLD) signal and extent of activation (Hadjikhani and Gelder, 2002; Rossion et al., 2003; Avidan et al., 2005; Steeves et al., 2006, 2009).

**Abbreviations:** fMRI, functional magnetic resonance imaging; FFA, fusiform face area; OFA, occipital face area; VOT, ventral occipitotemporal cortex; LOC, lateral occipital complex; ROI, region of interest.

Alternative hypotheses challenge the neuromodular view of VOT organization (Gauthier et al., 2000). The process accounts propose that it is the demands of visual processing, rather than stimulus properties that drive activity within the FFA and OFA. Studies focusing on the level of categorical recognition have offered some support for this hypothesis. One specific processing account suggests that FFA activation is associated with subordinate-level processing in general and processing at the individual, identity level in particular (Gauthier et al., 2000; Tarr and Gauthier, 2000). Individuation or identity processing where category membership is limited to a single exemplar is considered the ultimate example of subordinate-level processing (Gauthier et al., 2000).

Behavioral studies have in fact shown that adults tend to use the individual level as the entry point to face processing (Tanaka, 2001). In other words, for adults, processing a face at the individual level is the default mode and seems to be at least as efficient, if not more, than processing that face at the basic level. By contrast, the entry point for nearly all other object classes (e.g., birds, chairs) is at the basic level. Explicit tests of this individuation hypothesis by a small number of studies have yielded inconclusive results. For example, Mason and Macrae (2004) found greater activation for individuating faces compared to face categorization in right-hemisphere brain areas. George et al. (1999) observed that individuating faces produced greater fusiform gyrus activation than detecting faces within a contrast-polarity reversal task, consistent with the individuation hypothesis. Neither of these studies, however, included a condition that required participants to process non-face stimuli at the individual level. Such results are thus inconclusive with regard to whether an activation advantage for individuation over detection is face-specific or a more general phenomenon.

Gauthier et al. trained participants to be experts in identifying highly homogeneous, novel stimuli (Greebles) and showed greater activation in the FFA relative to non-experts when individuating those stimuli (Gauthier and Tarr, 2002). This finding provides the strongest evidence to date that enhanced FFA activation may not be face-specific. However, some researchers have argued that these results do not necessarily pose a true challenge to neuromodular stimulus-specific accounts because the Greeble stimuli contain features that make them very face-like (Kanwisher and Yovel, 2006). In addition, proponents of the neuromodular theory also have pointed out that although individuation of non-face stimuli induces enhanced FFA activation (Gauthier et al., 2000), the magnitude of this activation does not reach the level observed for faces (Kanwisher et al., 1997, 1998; McCarthy et al., 1997; Grill-Spector et al., 2004; Rhodes et al., 2004). Moreover, Kanwisher et al. found that while non-face objects did produce FFA activation during an individuation task, the greatest activation was still produced by faces (Grill-Spector et al., 2004; Yovel and Kanwisher, 2004; Kanwisher and Yovel, 2006).

Rhodes et al. (2004) addressed the individuation hypothesis by contrasting human face processing to the processing of individual Lepidoptera (i.e., moths and butterflies). They found that faces produced significantly greater activation in the FFA than Lepidoptera. However, behavioral performance for the Lepidoptera was substantially lower than that for faces in both Lepidoptera non-experts (66.5% versus 90.5%, respectively) and experts (71.6% for Lepidoptera), raising the possibility that the activation differences arose from differences in task difficulty rather than the nature of the stimuli *per se*. Finally, it is difficult to evaluate the individuation hypothesis from the results of the existing studies because individuation was co-varied with the effects of expertise (Gauthier et al., 1997, 2000) or manipulations of configural and featural processing (Yovel and Kanwisher, 2004). Nevertheless, these studies converge on the fact that while reliable FFA activation can be obtained from non-face stimuli, this activity is invariably quantitatively lower than activation for faces. This fact sustains the proposal that the FFA is specialized for face processing (Kanwisher and Yovel, 2006).

The present study was designed to directly test the specific effects of individuation on FFA activation during face and non-face object processing with the inclusion of controls that we consider crucial for testing the hypothesis. Following the established convention in the field, we used a passive viewing localizer task to identify putative face- and object-selective regions in the VOT. Second, we asked participants to perform an individuation task on which they judged whether a pair of sequentially presented stimuli depicted the same stimulus or different ones. On different trials the pictures were of two different stimuli from the same category. On same trials the two pictures showed slightly different views of the same face, wristwatch, or other common objects. This manipulation ensured that participants were specifically engaged in individuation, and not merely picture matching – a concern that most previous studies failed to address. Also, unlike most prior studies, all stimuli were matched in terms of spectral power, contrast, and brightness to control for possible undue influences of non-critical physical characteristics on face processing. More importantly, because faces are highly homogeneous, we used sets of similarly homogeneous common objects. One set comprised wristwatches that somewhat resembled faces (e.g., circular contours, consistent features contained within the outer counter, hour- and minute-arm configuration, or digital number arrangements). Another set was a collection of diverse objects (e.g., chair, cell phones). Although diverse objects were used, on any given trial participants saw a pair of objects from the same homogenous object category (e.g., a cell phone versus the same cell phone from a different view or a different cell phone) and were asked to determine whether they were of the same identity. We hypothesized that if the FFA and OFA are indeed face-specific, then faces should lead to greater activation compared to objects in both these two areas during this individuation task. However, if the areas are not face-specific but rather are engaged specifically by individuation processing, then FFA and OFA activation to faces and non-face objects should be indistinguishable.

## MATERIALS AND METHODS

### PARTICIPANTS

We tested 17 healthy young adults. Data from two participants were eliminated from the final analyses because one participant had uncorrectable fMRI artifact in one of the individuation tasks, and one participant did not produce a reliable FFA and a lateral occipital complex (LOC) in the localizer task. The final test sample was thus comprised of 15 participants (six females) with a mean age of 24.8 years (SD = 5.7). All participants were right-handed and had normal or corrected-to-normal vision, with no history of psychiatric or neurological disease, significant head trauma, substance abuse, or other known condition that may negatively impact brain function. The Human Research Protections Program

of the University of California, San Diego approved this study. Participants provided informed consent prior to the study and were paid for participation.

## STIMULI

The tasks used images of male and female faces, diverse objects, digital and analog wristwatches, and scrambled stimuli. Intact stimuli were obtained from digital pictures or digitized from photographs, and scrambled stimuli were created by randomly arranging 97% of the pixels within face, watch, and object images. All images were balanced for spectral power, contrast, and brightness to equate lower-level perceptual information using a customized Matlab program (see **Figure 1A**). Non-identical pictures of faces, watches, and objects were used in the matching "same" trials of the individuation task (i.e., "alternate views"). These "same" pairs presented the same stimulus with modest angle and lighting changes, and in the case of faces, slight variations in neutral facial expressions (see **Figure 1C**). Face stimuli (half male and half female) for the individuation task were obtained from the face database of the Psychological Image Collection at Stirling (PICS[1]). Watch stimuli (half analog and half digital) were obtained from various sources on the Internet with all watches set to the same time (i.e., 10:09). Diverse object stimuli were comprised of stimuli selected from 40 different basic-level categories (e.g., baskets, apples, balls, whistles, vases, cameras, umbrellas, jackets, books, lamps, etc.), and were obtained from various sources on the Internet (i.e., retail store sites such as amazon.com) or photographed in the lab. Alternate view stimuli for the individuation task were all obtained from objects photographed in the lab.

## GENERAL TASK DESIGN AND PROCEDURE

The experiment consisted of two tasks conducted during fMRI BOLD acquisition. Participants always received the passive viewing localizer task followed by the individuation task.

[1]http://pics.psych.stir.ac.uk/

### *Passive viewing localizer task*

A blocked fMRI design was used to present images of male and female faces (intermixed), diverse objects, analog and digital watches (intermixed), and scrambled stimuli. Each participant was presented two task runs using unique stimuli in each run (see **Figure 1B**). Each run included eight 20-s stimulus blocks (two for each stimulus category), interleaved with 16-s fixation epochs (crosshair stimulus only), with an additional 8-s fixation block at the beginning of each run. Within each stimulus block, 20 unique images were presented. Stimuli were presented for 300 ms followed by a 700-ms fixation interval. Participants were instructed to simply view the pictures during presentation. To ensure adequate attention during the task, participants were instructed to press a button each time they observed the fixation cross that was presented during the fixation period change from a standard typeface black cross to red bold typeface cross. This occurred within each fixation period 2-s prior to the beginning of the stimulus presentation. Potential effects of this fixation period response were removed during the regression analysis. Each localizer run lasted 4:56, during which 148 fMRI volumes were obtained (TR = 2000 ms).

### *Individuation task*

The individuation task required participants to judge whether pairs of images of faces, watches, objects, and scrambled stimuli were of the same identity (see **Figure 1C**). The task was administered in three experimental runs, one each for faces, watches, and objects. Scrambled stimuli were presented in each of the three task runs. A single task run consisted of 120 trials, including 48 stimulus match trials, 32 scrambled stimulus trials, and 40 fixation-only "null" trials. Individual trials were 2900 ms in duration, including the presentation of an asterisk for 300 ms to signal the start of a trial followed by a 200 ms blank screen, the presentation of the first stimulus (i.e., "target" stimulus) for 500 ms, a 300 ms fixation cross, the presentation of the second stimulus (e.g., "probe" stimulus) for 1200 ms, followed by a 400 ms blank screen. Participants were

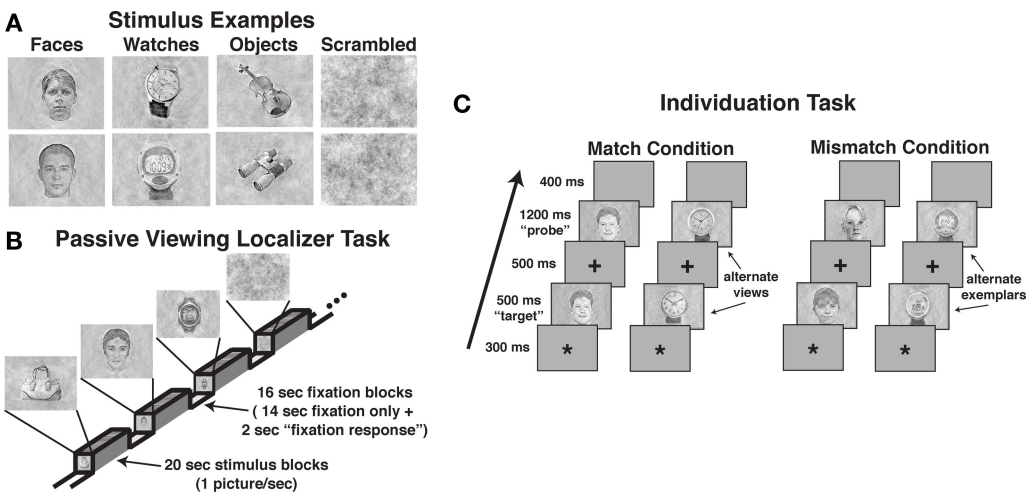



**FIGURE 1 | Examples of stimuli and the task design used in the three tasks.** **(A)** Examples of perceptually balanced stimuli used in the three tasks. Note that all watches were set to the same time. **(B)** The passive viewing localizer task used a blocked design format. **(C)** The design of match and mismatch trials in the individuation task. Alternate views of the same exemplar were used in the match trials (i.e., perceptually non-identical stimuli), whereas different exemplars from the same class (i.e., sex for faces, type for watches) were used in the mismatch trials.

instructed to press one of two buttons to indicate whether the probe stimulus was identical to the target stimulus shown in an alternate view (i.e., match) or was a different stimulus (i.e., mismatch). Half of the trials were match trials, and half were mismatch trials. In each run, 12 unique stimuli together with 12 alternate views of the stimuli were used (24 total pictures). In match trials, the original and alternate view of a stimulus were presented each as the target and probe stimuli during the task in two of the trials. We used alternate views of the same stimuli to prevent participants from using picture-matching strategies rather than individuation.

In the mismatch, the original and alternate view stimuli were paired with similar but different stimuli. For example, in the face runs, female faces were paired with different female faces and male faces were paired with different male faces; in the watch runs, digital watches were paired with different digital watches and analog watches were paired with analog watches; in the diverse object runs, cell phones were paired with different cell phones and chairs were paired with different chairs. The stimuli were paired with unique stimuli across trials (i.e., no pairs of mismatch stimuli were repeated). In addition, eight unique scrambled stimuli were used in each run to create 16 match and 16 mismatch trials. Each run lasted 5:48, during which 120 fMRI volumes were obtained (TR = 2900 ms). The order of face, watch, and object test runs were counterbalanced across participants.

## IMAGE ACQUISITION

Imaging data were obtained at the University of California, San Diego Center for Functional Magnetic Resonance Imaging using a short-bore 3.0-tesla General Electric Signa EXCITE MR scanner (Waubesha, WI, USA) equipped with a parallel-imaging capable GE eight-channel head coil. FMRI data were acquired using a single-shot gradient-recalled echo-planer imaging sequence with BOLD contrast (31 slices; 4-mm slab; TR = 2000 or 2900 ms; TE = 36 ms; flip angle = 90°; FOV = 240 mm; matrix = 64 × 64; in-plane resolution = 3.75 $mm^2$). Two 2D FLASH sequences were collected to estimate magnetic field maps and were used in post-processing to correct for geometric distortions. A high-resolution Fast SPGR scan was acquired for anatomical localization (sagittal acquisition; TR = 8.0 ms; TE = 3.1 ms; TI = 450 ms; NEX = 1; flip angle = 12°; FOV = 250 mm; acquisition matrix = 256 × 192; 172 slices; slice thickness = 1 mm; resolution = 0.98 × 0.98 × 1 mm).

## DATA ANALYSIS

FMRI preprocessing and analyses were performed using the Analysis of Functional NeuroImages (AFNI[2]; Cox and Hyde, 1997) and FSL[3] (Smith et al., 2004) packages. The fMRI BOLD image sequences were corrected for geometric distortion prior to the analyses with a customized script based on the FSL FUGUE program. Motion correction, slice time correction, and three-dimensional registration were done with an automated alignment program that co-registered each volume in the time series to the middle volume of the task run (Cox and Jesmanowicz, 1999). The images in each run were registered into standardized MNI/Talairach space, resampled to 27 $mm^3$ voxels (3 × 3 × 3 mm), and smoothed spatially to a fixed level of FWHM = 8 mm throughout the brain (Friedman et al., 2006).

[2]http://afni.nimh.nih.gov/afni

[3]http://www.fmrib.ox.ac.uk/fsl

### *fMRI analyses of individual participants*

The fMRI data from individual participants were analyzed using a deconvolution approach (AFNI 3dDeconvolve). In the blocked design passive localizer task, the hemodynamic response function (HRF) for each stimulus condition was modeled from a gamma variate function convolved with the stimulus time series (Cohen, 1997). For the event-related individuation tasks, the HRF for each stimulus was estimated from a series of seven spline basis functions (i.e., "tent functions") that modeled the post-trial onset window from 0 to 17.4 s (i.e., the fMRI volume acquisitions that included the stimulus presentation trial and the five subsequent post-stimulus volumes). Each task was analyzed using multiple regression that included the stimulus HRF parameters together with six parameters to account for motion artifacts (three rotation and three displacement variables), and polynomial factors of no interest (i.e., linear (all tasks), quadratic (all tasks), and cubic (individuation task)). The resulting regression weights for the stimuli were converted into percent signal values based on the voxel-wise global mean activation estimated from the regression analysis.

### *Definition of FFA, OFA, and LOC*

The FFA was defined for each participant as the area within the lateral fusiform gyrus where face stimuli produced reliably greater activation than diverse objects and scrambled stimuli in the passive localizer task, a method that has proven in prior studies (Rotshtein et al., 2005, 2007a,b; Yue et al., 2006) to produce FFA ROIs consistent with other methods used to define face superior processing in the lateral fusiform region (for review, see Berman et al., 2010). Similarly, the OFA defined as the area within the inferior occipital gyrus (BA 19) where faces produced reliably greater activation than diverse objects and scrambled stimuli. To correct against Type I errors, a cluster-threshold correction was applied that used a voxel-wise threshold of $P \leq 0.001$ with a minimum cluster volume of 225 μl resulting in an effective alpha level = 0.01 (Forman et al., 1995). The LOC region of interest was defined as the areas within the inferior and middle occipital gyri where diverse object stimuli produced reliably greater activation than scrambled stimuli in the passive localizer task (Malach et al., 1995; Moore and Engel, 2001; Kourtzi et al., 2003; Haushofer et al., 2008a,b). A cluster-threshold correction was applied that used a voxel-wise threshold of $P \leq 0.0001$ with a minimum cluster volume of 225 μl resulting in an effective alpha = 0.001. The location of the FFA, OFA, and LOC regions from each participant were confirmed against their respective high-resolution anatomical scan to insure that voxels were restricted to the FFA, OFA, and LOC, respectively.

### *Analysis of individuation task ROI activity*

Mean activation to each stimulus type was calculated for each participant within the FFA, OFA, and LOC ROIs for each of the six acquisition volumes estimated for the HRF. The data from each of the five ROIs were submitted to separate repeated measures analyses of variance (ANOVA) with stimulus (fixed effect), HRF time sample (fixed effect), and subject (random effect) as factors. Significant ANOVA effects were investigated using Bonferroni-corrected, paired-sample *t* tests (i.e., *post hoc* analyses). Bonferroni-corrected correlation coefficients were used for the comparison of ROI activation and behavioral performance. The effective alpha for the *post hoc* analyses

and correlational analyses following correction was 0.050. Results from male and female faces were combined because each produced comparable results, and likewise for analog and digital watches.

## RESULTS

### PASSIVE VIEWING LOCALIZER TASK

The behavioral task demands during the localizer were minimal and required only that the participants respond to a change in the fixation stimulus that occurred outside the presentation of the stimuli of interest. This procedure documented vigilance during the task while allowing for examination of brain activation to the stimuli of interest under passive viewing conditions. All participants correctly responded to 100% of these stimuli. We removed the brain activation effects of the fixation task manipulation via statistical regression.

The localizer task was administered to define face-selective regions within the VOT, namely, the FFA and OFA, and the object-selective regions within the LOC. Of the 16 participants that produced artifact-free fMRI data, 15 participants produced a reliable right-hemisphere FFA (rFFA), 14 produced a reliable left-hemisphere FFA (lFFA); 12 participants produced a reliable right OFA (rOFA), but only nine participants produced a reliable left OFA (lOFA); 15 participants produced a reliable right LOC (rLOC), and 14 produced a reliable left LOC (lLOC). As noted previously, one participant failed to produce an FFA or LOC and was eliminated from the sample. Due to the low number of participants producing a lOFA, this region was eliminated from the analysis. A composite depiction of the FFA, OFA, and LOC ROIs across participants is shown in **Figures 2A–C**, respectively, with descriptive statistics regarding the location and extent of the ROIs shown in **Table 1**. The location and extent of activation in the ROIs is in accordance with many previous descriptions of the FFA and LOC. Activation differences between hemispheres of homologous regions were not of interest, and due to different numbers of participants providing data in homologous regions, we analyzed the results from each of the five regions separately.

The mean percent BOLD signal activation across the ROI for the four stimulus classes is shown in **Figure 2**. As faces, objects, and scrambled stimuli were used to define the FFA and OFA, direct contrasts between these stimulus types were not warranted (Kriegeskorte et al., 2009). However, watch stimuli were withheld from use in characterizing these ROIs. The contrast of watch stimuli to faces and objects across the ROIs provides two important pieces of information. First, it establishes the reliability of face selectivity in the FFA and OFA ROIs, and it provides a reliability measure of object selectivity within the LOC ROI. Second, because the watch stimuli comprise a homogeneous class of objects relative to the diverse object stimuli, the watch contrasts allow for the evaluation

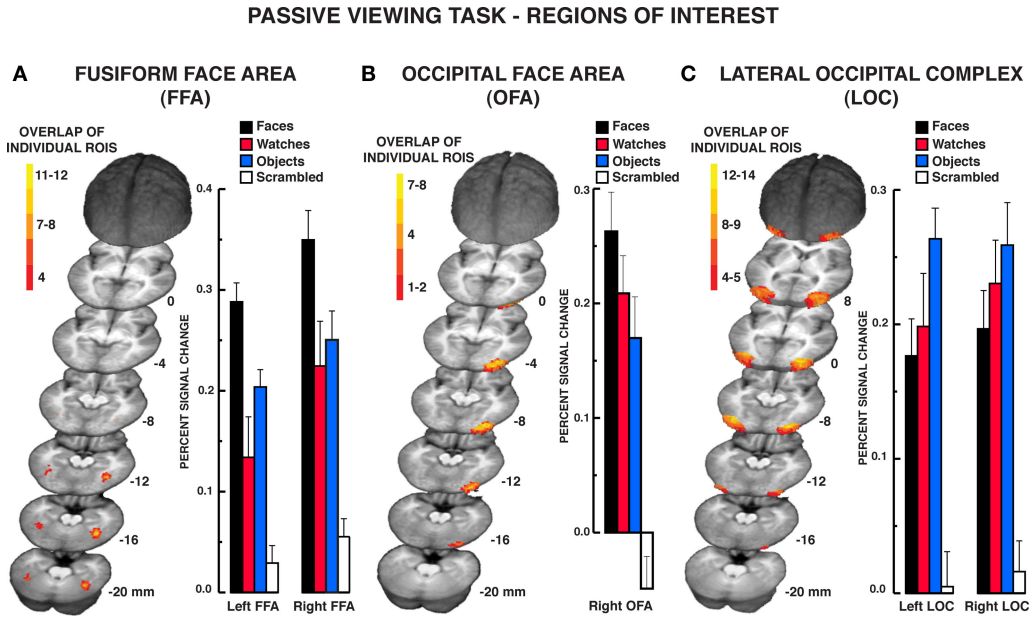



**FIGURE 2 | Results from the passive viewing localizer task.** Color-coded depiction of the location of FFA, OFA, and LOC regions of interest, respectively, from individual participants. The structural MRI underlay is the mean MRI from the 15 participants. **(A)** The FFA was defined as regions within the lateral fusiform gyrus producing significantly greater BOLD activation for faces relative to diverse object and scrambled stimuli. The location of the maximum overlap in the rFFA in standard Talairach coordinates (*x, y, z*; positive values = left, anterior, and superior, respectively) occurred at –40, –50, –16; 12 of 15 participants included this location within their rFFA. The location of the maximum overlap in the lFFA occurred at 38, –44, –16; six of 14 participants included this location within their lFFA. **(B)** The overlap map of individual right-hemisphere OFA ROIs. The left-hemisphere OFA is not shown as only 12 of the 17 participants produced a reliable left OFA. The location of the maximum overlap in the right OFA occurred at –28, 86, –10. Seven of the 15 participants included this location within their right OFA. **(C)** The LOC was defined as regions within the middle and inferior occipital gyri producing significantly greater activation to diverse objects relative to scrambled stimuli. The location of the maximum overlap in the rLOC occurred at –30, –81, –4; 14 of 15 participants included this location within their rLOC. The location of the maximum overlap in the lLOC occurred at 35, –70, –7; 13 of 14 participants included this location within their lLOC. The bar graphs display the mean signal in the respective regions of interest. Error bars = SEM.

**Table 1 | Characterization of FFA, OFA, and LOC regions of interest.**

| Region | *n* | Talairach coordinates center of cluster | | | Talairach coordinates Max. intensity voxel | | | Volume (μl) | Max. intensity difference (%) |
|---|---|---|---|---|---|---|---|---|---|
| | | *x* | *y* | *z* | *x* | *y* | *z* | | |
| Right FFA | 15 | –39 | –49 | –16 | –39 | –50 | –17 | 1179.0 | 0.76 |
| SD | | 2.3 | 5.0 | 2.4 | 4.1 | 7.5 | 3.4 | 825.8 | 0.33 |
| Left FFA | 14 | 37 | –47 | –15 | 37 | –46 | –17 | 754.1 | 0.59 |
| SD | | 3.5 | 6.9 | 2.6 | 4.2 | 8.4 | 3.6 | 488.4 | 0.22 |
| Right OFA | 12 | –34 | –81 | –9 | –37 | –80 | –10 | 1515.2 | 0.83 |
| SD | | 3.8 | 4.0 | 1.8 | 6.3 | 6.2 | 1.8 | 656.8 | 0.41 |
| Right LOC | 15 | –36 | –79 | –1 | –40 | –80 | –3 | 3947.1 | 0.65 |
| SD | | 2.4 | 1.9 | 3.3 | | | | 656.8 | 0.48 |
| Left LOC | 14 | 37 | –78 | –3 | 38 | –81 | –5 | 3997.9 | 0.67 |
| SD | | 3.5 | 3.9 | 3.5 | 5.3 | 8.0 | 7.9 | 2901.1 | 0.28 |

*Notes: Regions of interest were derived from the passive viewing localizer task. n, number of participants producing a statistically reliable region of interest. Talairach coordinates follow the convention of the Talairach and Tournoux atlas (Talairach and Tournoux, 1988). Positive coordinate values indicate left (x), anterior (y), and superior (z). Coordinates are provided for the estimated centroid of the ROI cluster and the location of the voxel showing the greatest intensity difference within the cluster (e.g., "hot spot"). Volume: mean volume of the region of interest. Maximum intensity difference: percent signal difference of faces versus diverse objects and scrambled stimuli (FFA and OFA) or diverse objects versus scrambled stimuli (LOC) at the voxel with the greatest activation difference within the ROI. SD: standard deviation.*

of homogeneity effects under conditions of spontaneous processing during the passive viewing task. Bonferroni-corrected, paired *t* tests were used for the contrast of watches to the other stimuli. Within the FFA, BOLD activation to watches was significantly less than faces bilaterally, rFFA: $t_{15} = 3.80$, $P = 0.002$; lFFA: $t_{13} = 3.59$, $P = 0.003$. However, the BOLD activation for watches and diverse objects did not differ in the FFA bilaterally, rFFA: $t_{15} = 1.01$, $P = 0.329$; lFFA: $t_{13} = 1.74$, $P = 0.106$. Within the rOFA, BOLD activation between watches and objects did not differ significantly, $t_{11} = 1.36$, $P = 0.202$, but there was only a trend for faces to produce greater BOLD activation than watches, $t_{11} = 2.51$, $P = 0.029$. This suggests that the rOFA may be sensitive to the homogeneity of stimuli during spontaneous processing.

The LOC ROIs defined object-preferential processing in the inferior/middle occipital gyrus area using the diverse object and scrambled stimuli. This allowed for the comparison of BOLD activation between objects, watches, and faces using Bonferroni-corrected, paired *t* tests objects versus faces, objects versus watches, and faces versus watches. Diverse objects produced significantly greater BOLD activation than faces within the LOC bilaterally, lLOC: $t_{13} = 3.32$, $P = 0.006$; rLOC: $t_{14} = 3.08$, $P = 0.008$. BOLD activation to watches did not differ significantly from that observed for objects bilaterally, lLOC: $t_{13} = 2.03$, $P = 0.063$; rLOC: $t_{14} = 0.94$, $P = 0.003$, and watches did not differ significantly from faces bilaterally, lLOC: $t_{13} = 0.78$, $P = 0.451$; rLOC: $t_{14} = 1.35$, $P = 0.198$. Thus, when defined in the traditional manner, the LOC produced a reliable object-superior processing effect relative to face stimuli. On the other hand, this effect did not extend to another class of object stimuli, watches that were more homogeneous than the diverse objects used in the localizer.

## INDIVIDUATION TASK

Participants in the individuation task judged whether sequentially presented pairs of faces, watches, objects, or scrambled stimuli were of the same identity. The stimulus pairs, with the exception of scrambled stimuli, were alternate views of the same exemplar or different exemplars from the same object class (e.g., male faces, digital watches, cell phones). A d' statistic was calculated to describe behavioral performance accuracy in discriminating same and different stimulus pairs as this measure considers both sensitivity to item similarity and response bias. Mean behavioral performance indicated that participants were highly accurate in their judgments across all stimulus types (d' mean ± SEM for faces = 4.58 ± 0.17, watches = 3.71 ± 0.23, objects = 4.44 ± 0.13, scrambled [collapsed across three runs] = 3.95 ± 0.14). A repeated measures one-way ANOVA indicated a significant main effect of stimulus, $F_{3,42} = 6.07$, $P = 0.002$. *Post hoc* analyses indicated that performance was lower for watches relative to objects, $t_{14} = 3.32$, $P = 0.005$. However, watch and object accuracies were not different from accuracy for faces, $P$s > 0.05. These relatively small differences in otherwise highly accurate behavioral performance did not appreciably impact the essential fMRI BOLD findings. There was no statistically significant correlation between behavioral performance (d') and activation within any of the five regions of interest in the individuation task (peak HRF response), Pearson correlation coefficient $P$s > 0.050.

The design of the individuation task tested face, diverse object, and watch stimuli in three sequential test blocks, counter-balanced between subjects, rather than intermixed within test blocks. The major drawback to the use of an intermixed design is that participants would not be able to anticipate which type of stimuli they were going to view and they would likely have to both categorize the stimuli and individuate them. Thus, BOLD activation associated with each stimulus trial might be attributed to both categorical and individuation processing, not individuation exclusively. In this stimulus blocked, event-related design, participants knew the type of stimuli that they were going to view. Their task was purely to individuate them. Moreover, there is considerable precedence in the face-processing literature for investigating different stimuli

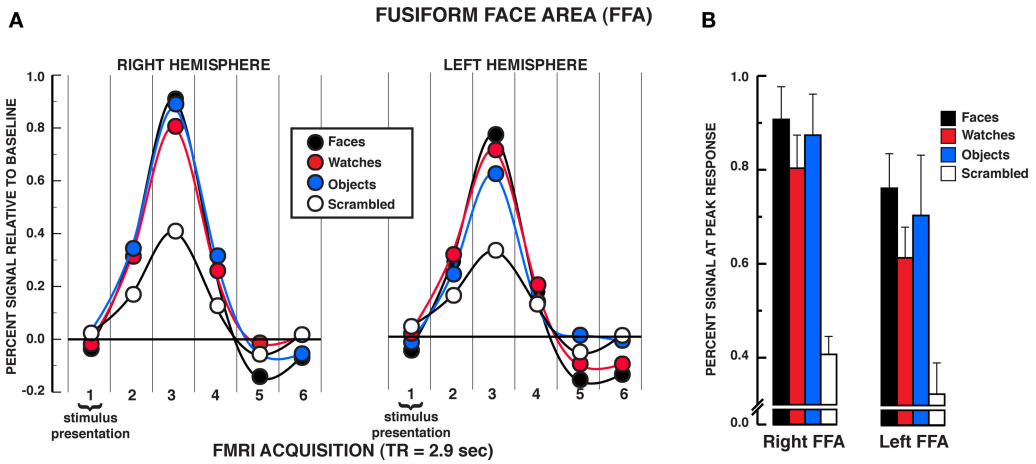



**FIGURE 3 | FMRI BOLD results from the individuation task in the FFA regions of interest. (A)** Estimated hemodynamic response function (HRF) for the faces, watches, diverse objects, and scrambled stimuli (mean from three individuation task runs) in the FFA. BOLD signal expressed in percent signal difference from null trial fixation baseline. The six acquisition volumes from the initiation of stimulus presentation were modeled. Each volume was acquired over 2.9 s (TR) that translates to the first volume acquired from 0 to 2.9 s, the second volume from 2.9 to 5.8 s, etc., and the entire HRF covered the 0–17.4 s window from the beginning of the trial. With slice-time resampling to the middle of the acquisition period, the peak HRF occurred in the third acquisition for each stimulus type in each ROI that translates to the peak response occurring at approximately 7.25 s post-trial onset. The HRF is shown with a spline interpolation for display purposes. **(B)** Mean peak response (third HRF acquisition volume) for the four stimuli in the FFA regions of interest. In both the rFFA and lFFA, no statistically significant differences were found between faces, watches, and diverse objects in the rFFA and lFFA. Error bars = SEM.

across separate experimental runs within the same scanning session (Downing et al., 2001, 2006; Grill-Spector et al., 2004). Nevertheless, acquiring data from different stimuli in different runs raises the potential that fMRI BOLD signal differences may arise due to baseline fluctuations or context effects. To evaluate this possibility, our design included similar scrambled stimuli tested in each of the three runs. A repeated measures one-way ANOVA of the peak BOLD signal response in the estimated HRF to scrambled stimuli in each task run within each of the five ROIs revealed no significant task run main effects, rFFA: $F_{2,28} = 2.33$, $P = 0.116$; lFFA: $F_{2,26} = 1.46$, $P = 0.251$; rOFA: $F_{2,22} = 1.55$, $P = 0.235$; rLOC: $F_{2,28} = 2.51$, $P = 0.099$; lLOC: $F_{2,26} = 1.77$, $P = 0.191$. This similarity in scrambled stimuli demonstrates that baseline or context effects arising from separate stimulus task runs did not influence the findings described below. Based on the similarity in BOLD activation for scrambled stimuli across tasks, the following analyses use the mean scrambled stimulus HRF from the three test runs.

The results from the individuation task offer compelling evidence that the FFA responds significantly to individuation processing factors rather than face-specific stimulus factors exclusively. The mean BOLD activations across the five regions of interest over the estimated HRF are shown in **Figures 3–5** (see Supplementary material for whole brain analysis of the individuation task). In the FFA bilaterally, a significant stimulus × time interaction was found in the two-way repeated measures ANOVA, rFFA: $F_{15,210} = 8.33$, $P < 0.001$; lFFA: $F_{15,195} = 4.88$, $P < 0.001$ (see **Figure 3A**). Focusing on the peak response in the HRF shown in **Figure 3B** (acquisition volume 3 in **Figure 3A**), *post hoc* analyses revealed that faces, diverse objects, and watches produced similar activation bilaterally, rFFA: *P*s > 0.104, lFFA: *P*s > 0.050. All three stimuli produced significantly greater activation than scrambled stimuli at the peak response, rFFA: *P*s < 0.001; lFFA: *P*s < 0.003. In addition, no differences were found

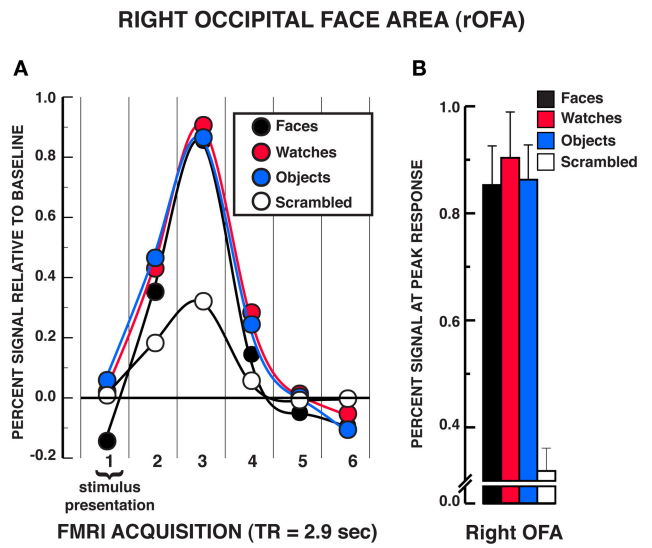



**FIGURE 4 | FMRI BOLD results from the individuation task in the right OFA region of interest. (A)** Estimated hemodynamic response function (HRF) for the faces, watches, diverse objects, and scrambled stimuli (mean from three individuation task runs) in the rOFA. See **Figure 3** for detailed description of HRF function. **(B)** Mean peak response (third HRF acquisition volume) for the four stimuli in the rOFA region of interest. The results are virtually identical to those observed in the rFFA in that no statistically significant differences were found between faces, watches, and diverse objects in the rOFA. Error bars = SEM.

in the activation levels for faces, diverse objects, and watches at any of the other HRF time points, rFFA and lFFA: *P*s > 0.050. In the rFFA, the stimulus main effect was reliable, $F_{3,42} = 7.24$, $P = 0.001$, due to faces, watches, and diverse objects producing overall greater

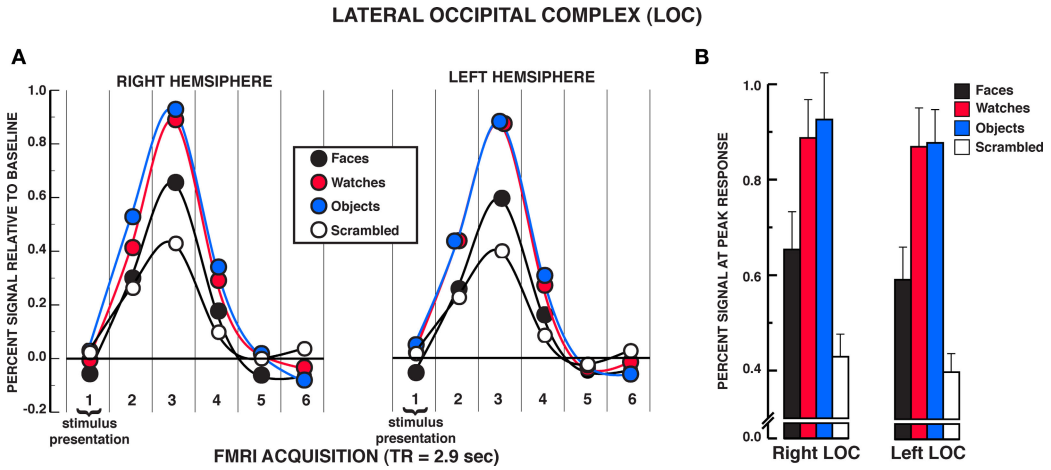



**FIGURE 5 | FMRI BOLD results from the individuation task in the LOC regions of interest. (A)** Estimated hemodynamic response function (HRF) for the faces, watches, diverse objects, and scrambled stimuli (mean from three individuation task runs) in the LOC. See **Figure 3** for detailed description of HRF function. **(B)** Mean peak response (third HRF acquisition volume) for the four stimuli in the LOC regions of interest. Activation to diverse objects and watches was significantly greater than faces in both the rLOC and lLOC. Error bars: SEM.

activation than scrambled stimuli. The stimulus main effect was not significant in the lFFA, $F_{3,39} = 1.08$, $P = 0.369$. The time main effect was reliable in both the rFFA and lFFA, rFFA: $F_{5,70} = 72.32$, $P < 0.001$; lFFA: $F_{5,65} = 36.88$, $P < 0.001$. In summary, there was no difference in activation between faces, diverse objects, and watches within the FFA. All three of these categories produced reliably greater activation than the scrambled stimulus control.

The findings in the rOFA substantially replicated those in the rFFA (see **Figure 4**). A significant stimulus × time interaction was found, rOFA: $F_{15,165} = 11.01$, $P < 0.001$. Focusing on the peak response in the HRF shown in **Figure 4B**, *post hoc* analyses revealed that faces, diverse objects, and watches produced similar activation in the rOFA, *P*s> 0.438. All three stimuli produced significantly greater activation than scrambled stimuli at the peak response, *P*s < 0.001. In addition, no differences were found in the activation levels for faces, diverse objects, and watches at any of the other HRF time points, *P*s > 0.050, with the exception that watches produced modestly greater activation than faces at the fourth acquisition, $t_{11} = 2.75$, $P = 0.019$. The stimulus main effect was reliable, $F_{3,33} = 7.65$, $P = 0.001$, due to faces, watches, and diverse objects producing overall greater activation than scrambled stimuli. The time main effect was reliable $F_{5,55} = 40.88$, $P < 0.001$. In summary, there was virtually no difference in activation between faces, diverse objects, and watches within the rOFA. All three of these categories produced reliably greater activation than the scrambled stimulus control.

Unlike findings within the FFA and OFA for equivalent activation of face and non-face stimuli under individuation instructions, the findings from the LOC suggested reliable differences between face and object processing (see **Figure 5A**). A significant stimulus × time interaction was observed in the two-way repeated measures ANOVA in both the rLOC and lLOC, rLOC: $F_{15,210} = 10.06$, $P < 0.001$; lLOC: $F_{15,195} = 8.78$, $P < 0.001$. Within the peak response of the HRF shown in **Figure 5B**, *post hoc* analyses revealed that diverse objects and watches produced similar activation, rLOC, $t_{14} = 0.59$, $P = 0.567$; lLOC, $t_{13} = 0.10$, $P = 0.919$, and both produced significantly greater activation than faces, rLOC: *P*s < 0.005; lLOC, *P*s < 0.002. All three stimuli produced greater activation within the LOC than scrambled stimuli, rLOC: *P*s < 0.005; lLOC, *P*s < 0.007. No significant differences were found between objects and watches at any of the other HRF time points, rLOC and lLOC: *P*s > 0.050. In the rLOC, diverse objects produced significantly greater activation than faces at the second and fourth time points in the HRF as well as at the peak response, HRF time 2, $t_{14} = 3.22$, $P = 0.006$; HRF time 4, $t_{14} = 3.35$, $P = 0.005$. In the lLOC, diverse objects and watches produced greater activation than faces in the second HRF time point, faces versus objects, $t_{13} = 3.64$, $P = 0.003$; faces versus watches, $t_{13} = 4.16$, $P = 0.001$. The stimulus main effect was reliable in both the rLOC and lLOC, rLOC: $F_{3,42} = 12.52$, $P = 0.001$; lLOC: $F_{3,39} = 13.35$, $P < 0.001$. In both LOC ROIs, watches and diverse objects produced greater overall activation than faces and scrambled stimuli, *P*s < 0.050, but did not differ from one another, *P*s > 0.050. In addition, overall activation to faces did not differ from that of scrambled stimuli in either LOC ROI, *P*s > 0.050. The time main effect was reliable in both the rLOC and lLOC, rLOC: $F_{5,70} = 36.97$, $P < 0.001$; lLOC: $F_{5,65} = 13.35$, $P < 0.001$. In summary, in both LOC ROIs, faces produced significantly less activation than diverse objects and watches in the individuation task.

To summarize, the FFA and OFA results suggest that there is no difference in FFA and OFA activation to faces and non-face objects under individuation instructions. This occurred in a region defined by a standard localizer task of passive viewing as having produced reliably greater activation to faces than to objects and watches. In a region commonly associated with object processing, the LOC produced significantly greater activation for diverse objects and watches than faces. The processing differences between stimuli in the LOC suggest that individuation processing *per se* is not the dominant factor driving BOLD activation in this region.

## DISCUSSION

The present study was designed to determine whether the putative face-selective brain regions within the ventral occipitotemporal cortex (VOT) are indeed selectively responsive to faces. We find that this is not the case. More specifically, we find that although these areas appear to be face-specific when the contrast is between the passive viewing of faces versus other non-face object stimuli, this apparent specificity vanishes when the task requires object individuation. In other words, under task conditions calling for processing at the level of individual objects (i.e., are these the exact same or different object?), these putative face-selective brain regions within the VOT (FFA and OFA) are equally activated for carefully selected non-face stimuli as for faces.

Two findings from this study are paramount to understanding the functional role of the VOT in visual face and object processing. First, we showed that under task conditions that require explicit individuation of visual stimuli, both face and non-face objects produced equivalent levels of BOLD activation within the functionally defined FFA and OFA regions of interest that produced a face-superiority effect in a traditional form of a passive viewing localizer task. By contrast, this individuation task did not eliminate the object-superior processing effect (preferential processing for objects relative to faces) within the LOC, especially the left-hemisphere LOC. This pattern of results indicates that individuation processing does not indiscriminately raise face and non-face object processing to equal levels across all regions within the VOT extrastriate brain regions; it does so only in the putative face areas – FFA and OFA.

Taken together, these findings have significant implications for current models of the role of the FFA specifically in face processing, and more generally in visual object recognition. Most notably, our findings do not support the hypothesis that the FFA and OFA are specialized for face processing (Yovel and Kanwisher, 2004; Kanwisher and Yovel, 2006) because we find that this is not the case. Indeed, we found that at least under individuation task conditions, non-face objects such as watches and diverse objects generate the same level of FFA and OFA activation as faces. To our knowledge, this is the first report of face-equivalent activation of the FFA for non-face objects, despite many such attempts.

Naturally, the key question is why our task induces these levels of activation to non-face objects when previous studies have not. Proponents of face-specific processing accounts have dismissed individuation as a key function of the FFA based primarily on the failed attempts to produce face-level activation for non-face objects within the FFA in presumed individuation processing tasks (c.f., Kanwisher and Yovel, 2006). One of the early studies establishing the role of the FFA in face processing, for example, demonstrated face-superior processing in a passive viewing task similar to that presented in this study and in a task that used a sparse one-back matching task (Kanwisher et al., 1997). In fact, the sparse one-back matching task is one of the most commonly used face-processing tasks used today; in this task participants are instructed to identify infrequent repetitions of stimuli within a stimulus block. The original finding is taken as evidence against an individuation account of the FFA because face stimuli produced greater activation within the FFA (across five subjects) in both the passive viewing and sparse one-back tasks, thereby suggesting that individual item matching did not meaningfully alter FFA responsiveness. This argument is based on the assumption that the sparse one-back matching task requires item-by-item individual discrimination decisions, even though behavioral responses occur only to the infrequent mismatches. We question this assumption.

The typical sparse one-back matching task, including Kanwisher and colleagues' original instantiation of it, does not adequately assess individuation processing because the match decisions always refer to the exact same stimulus whereas the mismatching decisions always refer to physically different stimuli. In these paradigms, then, match and mismatch decisions can be made purely on the basis of simple visual perceptual differences and do not by necessity entail substantive individuation discriminations. In contrast, our individuation task required item-by-item matching decisions on alternative views of the identical items; mismatching decisions were based on visually similar items in the same category (e.g., male versus another male face, analog watch versus another analog watch). We maintain that using alternative views of the same object for matching decisions and highly homogeneous though different stimuli for mismatching decision is crucial for minimizing simple perceptual matching and truly evaluating individuation processing without confounding factors.

Results from the Yovel and Kanwisher (2004) study evaluating the role of the FFA in configural and featural processing might, at first blush, appear to challenge a role for individuation in the FFA. They presented participants with pairs of faces or houses and asked them to judge whether the probe stimulus was identical to the target stimulus. Stimuli were either identical or varied in the spacing of the features (e.g., configural differences) or in the features themselves (e.g., featural differences). Their key finding was that faces produced significantly greater FFA activation than did houses regardless of the type of stimulus change, which they interpreted as evidence for the face-specificity of the FFA. However, the task manipulation they used to insure task compliance may have altered the tasks demands so as to direct participant judgments toward categorical level judgments. Specifically, a cue at the beginning of each stimulus block instructed participants as to whether the stimuli would differ in their features or in the configuration of features. This cue essentially altered the task to a decision from one requiring individuation to one in which participants decided whether or not a change took place in features or in feature spacing. This, in essence, converted the task into a form of categorization or subordinate processing task. While subordinate processing task can modulate FFA activity, past studies of subordinate processing have failed to eliminate face superior processing in the FFA (see Kanwisher and Yovel, 2006). The results of Rhodes and colleagues contrasting brain activity for judgments about faces versus Lepidoptera also might seem to contradict the FFA individuation hypothesis (Rhodes et al., 2004). In addition to the problematic difference in behavioral performance between the conditions, as already noted, the decision task in this study also precluded a direct test of the individuation hypothesis. Specifically, they had participants study individual exemplars of faces and Lepidoptera over several days, and then perform old/new judgments. Studies of declarative memory using tasks such as this have shown clearly that familiarity may provide the sole basis for any given recognition judgment, rather than explicit recollection of definitive features of items or context (Yonelinas, 2002). The difference in recognition levels for Lepidoptera versus faces in

Rhodes et al. raises the valid concern that participants made judgments for the Lepidoptera with lower confidence. It is reasonable to suggest that familiarity-based judgments do not engage individuation processing that requires explicit recollection of specific features, configurations, or holistic information. As discussed above, the entry point for nearly all non-face object classes, which would include Lepidoptera, is at the basic level (Tanaka, 2001). Thus, the failure to modulate FFA activation for non-face objects in a task using difficult to discriminate objects with which one has minimal recollection of specific details is not strong evidence against the role of the FFA in individuation processing.

Grill-Spector et al. (2004) investigated the role of the FFA in face and non-face object detection and identification, concluding that the FFA had no major role in within-category identification of non-face objects, even in people with specialized expertise with those particular objects (e.g., car experts). They asked participants to identify a specific target picture appearing in a stream of non-target pictures drawn from the same category (e.g., other faces), stimuli from a different category (e.g., birds), or textured (non-sense) pictures. Identification was defined as a correct judgment of the target stimulus, whereas detection was defined either as judging the target as a category member without indicating it was the target or by correctly identifying a non-target stimulus as a member of a particular category (e.g., judging a car as a car). They found that faces produced greater activation of the FFA in both detection and identification. In contrast, detection of non-face objects was not correlated with FFA activation. Again, an analysis of Grill-Spector et al.'s methods leads us to conclude that they were not assessing the process of individuating non-face objects. To assess face processing, participants were instructed to identify a particular face, such as Harrison Ford, in a series of other famous faces; this is a direct test of individuation processing. Their assessment of non-face objects, by contrast, required only subordinate, categorical level processing. Specifically, their participants were instructed to identify pigeons from among a series of other type of birds. Their participants were not asked to identify any individual or particular pigeon in a series of pigeons. As a consequence, this experiment compared individuation processing in faces to subordinate processing at the categorical, not the individual, level in non-face objects. In sum, ours is the first neuroimaging study to directly compare the process of *individuation* for faces and non-face objects without any confounding concerns, and to find that the level of face and non-face object activations in the FFA and OFA are the same.

Our results offer new insights into accounts of FFA and OFA functions that have emphasized various processing factors. For example, the homogeneity of stimuli to be discriminated is proposed to be one factor that modulates FFA and OFA activation (Gauthier et al., 2000; Tarr and Gauthier, 2000). Stimulus homogeneity alone, however, cannot account for the present findings. All of the stimuli in our individuation task were indeed highly homogeneous. That is, we paired faces to faces, watches to watches, and similar basic-level objects to each other (e.g., cell phone to cell phone). Moreover, the results from the passive viewing task suggest that homogeneity of stimuli *per se* did not significantly modulate spontaneous object processing in the FFA. Watches represented a more homogeneous category of stimuli relative to diverse objects, but watches and diverse objects produced equivalent activation in the FFA during passive viewing, which were both significantly less than the activation to faces. We note, however, that watches produced only a marginal difference in activation within the OFA compared to faces during the localizer task. This may suggest that the OFA may be slightly more sensitive to homogeneity factors than the FFA. Perhaps more notable than stimulus homogeneity effects, visual processing expertise also cannot explain the present findings. Our individuation task results suggest that substantial expertise is not a necessary precondition for FFA activation modulation by non-face objects as observers (presumably non-experts with either watches or our other diverse objects) produced equivalent activations to faces and non-face object stimuli when assessed under strict instructions to individuate them.

Whereas visual processing expertise does not seem to play a necessary role in engendering a heightened level of activation in the FFA, our data suggest another role for expertise. Specifically, expertise with a particular stimulus class seems to influence the entry level or default processing strategy adopted by an individual. The vast expertise for faces acquired by typical adults establishes individuation as the entry level of processing regardless of external task demands (Tanaka, 2001). Expertise with faces appears to lead to the automatic (default) engagement of individuation processing, thereby activating the FFA regardless of the externally specified task demands. In the current study, the default engagement of individuation during face processing produced apparent face-preferential FFA and OFA activation during the passive localizer task. In contrast, given a lower level of visual processing expertise with a particular stimulus class, such as watches or diverse objects, differing tasks demands modulate the level of processing, and the associated FFA and OFA activation. Because individuation is not the entry level of processing for watches and diverse objects, only the explicit instructions to individuate items in these categories produced robust FFA and OFA activation. We observed reduced FFA and OFA activation for watches and other common objects when the task was unspecified as in the localizer task.

Data accumulating from a number of different fronts are beginning to inform the functional architecture of individuation processing for faces. An important source of such data is functional imaging studies of patients with the severe facial recognition deficits of prosopagnosia. Several reports based on a limited number of well-characterized patients with acquired prosopagnosia question a crucial role for the FFA in individuation processing as we suggest based on the present findings, and instead suggest a larger role in individuation processing for the OFA. For example, the severe facial recognition deficits of two often-studied acquired prosopagnosia patients, P.S. and D.F., are linked to lesions that include the right inferior occipital gyrus region that encompasses the territory of the OFA, but the right middle fusiform gyrus region that includes the FFA is spared in both patients (Rossion et al., 2003; Schiltz et al., 2006; Steeves et al., 2006, 2009; Dricot et al., 2008a,b). A normal FFA has been observed in both patients obtained across multiple localizer tasks. Perhaps the most striking finding is that both patients fail to show the typical face repetition suppression effect when identical face stimuli are repeated across multiple presentations (Schiltz et al., 2006; Dricot et al., 2008a,b; Steeves et al., 2009). Moreover, it has been observed that regions adjacent to the FFA that are not typically observed as showing face-specific processing can produce the face repetition suppresssion effect (Dricot et al., 2008a,b).

Overall, findings from acquired prosopagnosia suggest that individual identity processing, particularly for faces, is not the sole domain of middle fusiform gyrus region and emphasizes an important role for the OFA, a finding consistent with the present study. However, a recent study indicates that the OFA alone may not be the exclusive locus for face individuation processing because it may not be the locus for the face repetition suppression effect. Gilaie-Dotan et al. (2010) applied transcranial magnetic stimulation (TMS) to the OFA region during the period between the first and second repetition of famous faces, thus disrupting potential stimulus repetition effects in the OFA during this period, and found a repetition suppression effect. This is not inconsistent with data from congenital prosopagnosia that has found relatively normal fMRI activation to faces and objects within the FFA and OFA regions despite severe facial recognition deficits, and decreased activation within the face-sensitive areas in anterior brain areas, the so-called "extended" face-processing network (Avidan et al., 2005; Avidan and Behrmann, 2009). One region within this extended network garnering attention for a specific role in individuation of faces beyond that of simple detection is located on the extreme inferior aspects of the temporal lobe including the inferior temporal gyrus (BA 20) (Nestor et al., 2008) and/or middle temporal gyrus (BA 38) (Kriegeskorte et al., 2007). We did not find face-specific individuation processing activation in this region (see Supplementary material). Due to its location in a region well known to be susceptible to MR drop-out effects, this region may be difficult to ascertain in group studies (Kriegeskorte et al., 2007). In addition, those other studies used face individuation tasks that significantly differed in methodology from that used here; Kriegeskorte et al. (2007) used an anomaly-detection task with infrequent deviations in sequential image presentation, while Nestor et al. (2008) used a challenging face individuation task based on face fragment stimuli. We did, however, find regions that produced face-greater-than-object activation during individuation that are considered part of the extended face-processing network, including posterior cingulate gyrus and middle temporal gyrus (see Supplementary material). In contrast, a larger number of regions were identified that produced object-greater-than-face activation during individuation. This may be due to the simple fact that individuation processing for non-face objects is more difficult than faces because individuation processing is not the default or typical entry point of processing for non-face objects (Tanaka, 2001). Taken together, these findings suggest that the FFA or OFA are not the sole regions participating in individuation processing; other brain areas certainly participate in the individuation for faces and non-face objects. These brain areas may be part of a network that is enlisted as individuals process any visual stimuli at the subordinate levels with individuation engaging the network to the greatest extent.

In summary, we find that fMRI activations in the FFA and the OFA show no face specificity when faces and non-face objects are both processed at an individual (as opposed to categorical) level. Rather than being specific to faces, these brain areas play a critical role in individuation – the process of distinguishing one individual exemplar (face or otherwise) from another. It has taken this long to come to this realization because our natural tendency when we look at faces (relative to other objects) is to individuate.

## ACKNOWLEDGMENTS

We thank Janet Shin for assistance in the conduct of this study, Deborah Weiss and Sarah Noonan for assistance with preliminary studies, Flora Suh for assistance with manuscript preparation, Dirk Beer for the development of the balanced stimuli, and Maha Adamo and Marta Kutas for comments on the manuscript. The National Institute of Child Health and Human Development Grants R01-HD041581, R01-HD046526, and R01-HD060595, Natural Sciences and Engineering Research Council of Canada Grant, and National Science Foundation of China Grants 60910006 and 31028010 supported this work.

## REFERENCES

Aguirre, G. K., Zarahn, E., and D'Esposito, M. (1998). An area within human ventral cortex sensitive to "building" stimuli: evidence and implications. *Neuron* 21, 373–383.

Avidan, G., and Behrmann, M. (2009). Functional MRI reveals compromised neural integrity of the face processing network in congenital prosopagnosia. *Curr. Biol.* 19, 1146–1150.

Avidan, G., Hasson, U., Malach, R., and Behrmann, M. (2005). Detailed exploration of face-related processing in congenital prosopagnosia: 2. Functional neuroimaging findings. *J. Cogn. Neurosci.* 17, 1150–1167.

Berman, M. G., Park, J., Gonzalez, R., Polk, T. A., Gehrke, A., Knaffla, S., and Jonides, J. (2010). Evaluating functional localizers: the case of the FFA. *Neuroimage* 50, 56–71.

Clark, V. P., Keil, K., Maisog, J. M., Courtney, S., Ungerleider, L. G., and Haxby, J. V. (1996). Functional magnetic resonance imaging of human visual cortex during face matching: a comparison with positron emission tomography. *Neuroimage* 4, 1–15.

Cohen, M. S. (1997). Parametric analysis of fMRI data using linear systems methods. *Neuroimage* 6, 93–103.

Cox, R. W., and Hyde, J. S. (1997). Software tools for analysis and visualization of fMRI data. *NMR Biomed.* 10, 171–178.

Cox, R. W., and Jesmanowicz, A. (1999). Real-time 3D image registration for functional MRI. *Magn. Reson. Med.* 42, 1014–1018.

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.

Downing, P. E., Jiang, Y., Shuman, M., and Kanwisher, N. (2001). A cortical area selective for visual processing of the human body. *Science* 293, 2470–2473.

Dricot, L., Sorger, B., Schiltz, C., Goebel, R., and Rossion, B. (2008a). Evidence for individual face discrimination in non-face selective areas of the visual cortex in acquired prosopagnosia. *Behav. Neurol.* 19, 75–79.

Dricot, L., Sorger, B., Schiltz, C., Goebel, R., and Rossion, B. (2008b). The roles of "face" and "non-face" areas during individual face perception: evidence by fMRI adaptation in a brain-damaged prosopagnosic patient. *Neuroimage* 40, 318–332.

Forman, S. D., Cohen, J. D., Fitzgerald, M., Eddy, W. F., Mintun, M. A., and Noll, D. C. (1995). Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. *Magn. Reson. Med.* 33, 636–647.

Friedman, L., Glover, G. H., Krenz, D., and Magnotta, V. (2006). Reducing inter-scanner variability of activation in a multicenter fMRI study: role of smoothness equalization. *Neuroimage* 32, 1656–1668.

Gauthier, I., Anderson, A. W., Tarr, M. J., Skudlarski, P., and Gore, J. C. (1997). Levels of categorization in visual recognition studied using functional magnetic resonance imaging. *Curr. Biol.* 7, 645–651.

Gauthier, I., Skudlarski, P., Gore, J. C., and Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nat. Neurosci.* 3, 191–197.

Gauthier, I., and Tarr, M. J. (2002). Unraveling mechanisms for expert object recognition: bridging brain activity and behavior. *J. Exp. Psychol. Hum. Percept. Perform.* 28, 431–446.

Gauthier, I., Tarr, M. J., Moylan, J., Anderson, A. W., Skudlarski, P., and Gore, J. C. (2000). Does visual subordinate-level categorisation engage the functionally defined fusiform face area? *Cogn. Neuropsychol. Special Issue: Cogn. Neurosci. Face Process.* 17, 143–163.

Gauthier, I., Tarr, M. J., Moylan, J., Skudlarski, P., Gore, J. C., and Anderson, A. W. (2000). The fusiform "face area" is part of a network that processes faces at the individual level. *J. Cogn. Neurosci.* 12, 495–504.

George, N., Dolan, R. J., Fink, G. R., Baylis, G. C., Russell, C., and Driver, J. (1999). Contrast polarity and face recognition in the human fusiform gyrus. *Nat. Neurosci.* 2, 574–580.

Gilaie-Dotan, S., Silvanto, J., Schwarzkopf, D. S., and Rees, G. (2010). Investigating representations of facial identity in human ventral visual cortex with transcranial magnetic stimulation. *Front. Hum. Neurosci.* 4:50. doi: 10.3389/fnhum.2010.00050.

Grill-Spector, K., Knouf, N., and Kanwisher, N. (2004). The fusiform face area subserves face perception, not generic within-category identification. *Nat. Neurosci.* 7, 555–562.

Hadjikhani, N., and Gelder, B. D. (2002). Neural basis of prosopagnosia: an fMRI study. *Hum. Brain Mapp.* 16, 176–182.

Hanson, S. J., Matsuka, T., and Haxby, J. V. (2004). Combinatorial codes in ventral temporal lobe for object recognition: Haxby (2001) revisited: is there a "face" area? *Neuroimage* 23, 156–166.

Haushofer, J., Baker, C. I., Livingstone, M. S., and Kanwisher, N. (2008a). Privileged coding of convex shapes in human object-selective cortex. *J. Neurophysiol.* 100, 753–762.

Haushofer, J., Livingstone, M. S., and Kanwisher, N. (2008b). Multivariate patterns in object-selective cortex dissociate perceptual and physical shape similarity. *PLoS Biol.* 6:7, e187. doi:10.1371/journal.pbio.0060187.

Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., and Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 293, 2425–2430.

Haxby, J. V., Hoffman, E. A., and Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends Cogn. Sci.* 4, 223–233.

Kanwisher, N., McDermott, J., and Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J. Neurosci.* 17, 4302–4311.

Kanwisher, N., Tong, F., and Nakayama, K. (1998). The effect of face inversion on the human fusiform face area. *Cognition* 68, B1–B11.

Kanwisher, N., and Yovel, G. (2006). The fusiform face area: a cortical region specialized for the perception of faces. *Philos. Trans. R. Soc. Lond. B, Biol. Sci.* 361, 2109–2128.

Kourtzi, Z., Erb, M., Grodd, W., and Bulthoff, H. H. (2003). Representation of the perceived 3-D object shape in the human lateral occipital complex. *Cereb. Cortex* 13, 911–920.

Kriegeskorte, N., Formisano, E., Sorger, B., and Goebel, R. (2007). Individual faces elicit distinct response patterns in human anterior temporal cortex. *Proc. Natl. Acad. Sci. U.S.A.* 104, 20600–20605.

Kriegeskorte, N., Simmons, W. K., Bellgowan, P. S. F., and Baker, C. I. (2009). Circular analysis in systems neuroscience: the dangers of double dipping. *Nat. Neurosci.* 12, 535–540.

Malach, R., Reppas, J. B., Benson, R. R., Kwong, K. K., Jiang, H., Kennedy, W. A., Ledden, P. J., Brady, T. J., Rosen, B. R., and Tootell, R. B. (1995). Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proc. Natl. Acad. Sci. U.S.A.* 92, 8135–8139.

Mason, M. F., and Macrae, C. N. (2004). Categorizing and individuating others: the neural substrates of person perception. *J. Cogn. Neurosci.* 16, 1785–1795.

McCarthy, G., Puce, A., Gore, J. C., and Allison, T. (1997). Face-specific processing in the human fusiform gyrus. *J. Cogn. Neurosci.* 9, 605–610.

Moore, C., and Engel, S. A. (2001). Neural response to perception of volume in the lateral occipital complex. *Neuron* 29, 277–286.

Nestor, A., Vettel, J. M., and Tarr, M. J. (2008). Task-specific codes for face recognition: how they shape the neural representation of features for detection and individuation. *PLoS One* 3: 12, e3978. doi:10.1371/journal.pone.0003978.

O'Toole, A. J., Jiang, F., Abdi, H., and Haxby, J. V. (2005). Partially distributed representations of objects and faces in ventral temporal cortex. *J. Cogn. Neurosci.* 17, 580–590.

Puce, A., Allison, T., Gore, J. C., and McCarthy, G. (1995). Face-sensitive regions in human extrastriate cortex studied by functional MRI. *J. Neurophysiol.* 74, 1192–1199.

Rhodes, G., Byatt, G., Michie, P. T., and Puce, A. (2004). Is the fusiform face area specialized for faces, individuation, or expert individuation? *J. Cogn. Neurosci.* 16, 189–203.

Rossion, B., Caldara, R., Seghier, M., Schuller, A. M., Lazeyras, F., and Mayer, E. (2003). A network of occipito-temporal face-sensitive areas besides the right middle fusiform gyrus is necessary for normal face processing. *Brain* 126(Pt. 11), 2381–2395.

Rotshtein, P., Geng, J. J., Driver, J., and Dolan, R. J. (2007a). Role of features and second-order spatial relations in face discrimination, face recognition, and individual face skills: behavioral and functional magnetic resonance imaging data. *J. Cogn. Neurosci.* 19, 1435–1452.

Rotshtein, P., Vuilleumier, P., Winston, J., Driver, J., and Dolan, R. (2007b). Distinct and convergent visual processing of high and low spatial frequency information in faces. *Cereb. Cortex* 17, 2713–2724.

Rotshtein, P., Henson, R. N., Treves, A., Driver, J., and Dolan, R. J. (2005). Morphing Marilyn into Maggie dissociates physical and identity face representations in the brain. *Nat. Neurosci.* 8, 107–113.

Schiltz, C., Sorger, B., Caldara, R., Ahmed, F., Mayer, E., Goebel, R., and Rossion, B. (2006). Impaired face discrimination in acquired prosopagnosia is associated with abnormal response to individual faces in the right middle fusiform gyrus. *Cereb. Cortex* 16, 574–586.

Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E., Johansen-Berg, H., Bannister, P. R., De Luca, M., Drobnjak, I., Flitney, D. E., Niazy, R. K., Saunders, J., Vickers, J., Zhang, Y., De Stefano, N., Brady, J. M., and Matthews, P. M. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* 23(Suppl. 1), S208–S219.

Steeves, J., Dricot, L., Goltz, H. C., Sorger, B., Peters, J., Milner, A. D., Goodale, M. A., Goebel, R., and Rossion, B. (2009). Abnormal face identity coding in the middle fusiform gyrus of two brain-damaged prosopagnosic patients. *Neuropsychologia* 47, 2584–2592.

Steeves, J. K., Culham, J. C., Duchaine, B. C., Pratesi, C. C., Valyear, K. F., Schindler, I., Humphrey, G. K., Milner, A. D., and Goodale, M. A. (2006). The fusiform face area is not sufficient for face recognition: evidence from a patient with dense prosopagnosia and no occipital face area. *Neuropsychologia* 44, 594–609.

Talairach, J., and Tournoux, P. (1988). *Co-Planar Stereotaxic Atlas of the Human Brain* (M. Rayport, Trans.). New York: Thieme Medical.

Tanaka, J. W. (2001). The entry point of face recognition: evidence for face expertise. *J. Exp. Psychol. Gen.* 130, 534–543.

Tarr, M. J., and Gauthier, I. (2000). FFA: a flexible fusiform area for subordinate-level visual processing automatized by expertise. *Nat. Neurosci.* 3, 764–769.

Yonelinas, A. P. (2002). The nature of recollection and familiarity: a review of 30 years of research. *J. Mem. Lang.* 46, 441–517.

Yovel, G., and Kanwisher, N. (2004). Face perception: domain specific, not process specific. *Neuron* 44, 889–898.

Yue, X., Tjan, B. S., and Biederman, I. (2006). What makes faces special? *Vision Res.* 46, 3802–3811.

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

*Received: 24 February 2010; paper pending published: 05 May 2010; accepted: 07 September 2010; published online: 08 October 2010.*

*Citation: Haist F, Lee K and Stiles J (2010) Individuating faces and common objects produces equal responses in putative face-processing areas in the ventral occipitotemporal cortex. Front. Hum. Neurosci.* ***4****:181. doi:10.3389/fnhum.2010.00181*

## SUPPLEMENTARY MATERIAL

The main study investigated BOLD activation in three ROIs, the FFA, OFA, and LOC, determined from the passive viewing localizer task. Here, we present the results from the whole brain analysis of the individuation task.

## DATA ANALYSIS

### ANALYSIS OF INDIVIDUATION TASK WHOLE BRAIN ACTIVITY

The data from the peak response of the HRF for faces, watches, and objects from each of the 15 participants were analyzed using a repeated measures analysis of variance (ANOVA) with stimulus (fixed effect) and subject (random effect) as factors. We restricted analysis to voxels that produced a significant main effect of stimulus ($P \leq 0.010$ uncorrected). To correct that statistics for multiple comparisons, a voxel-cluster-threshold correction was used with parameters based on a Monte Carlo simulation (Forman et al., 1995). The cluster-threshold correction required a voxelwise threshold of $P \leq 0.005$ within a volume of at least 351 μl (13 contiguous voxels) to yield an effective alpha ≤ 0.050.

## RESULTS

### INDIVIDUATION TASK

The central ROI-based analysis revealed that faces, watches, and diverse object produced virtually identical activation within the FFA in the individuation task. Here, we describe differences in activation observed during the individuation task throughout the rest of the brain for the contrast between the faces, watches, and diverse object stimuli. These results are shown in **Figure S1**. **Table S1** describes the regional activation observed in the individuation task. Faces produced greater activation than watches during the individuation task within a limited number of posterior regions located primarily in the right hemisphere. However, faces did not produce any regions of greater activity relative to objects. We observed many more regions active for watches and objects relative to faces in the individuation task. These were distributed primarily as bilateral activation of posterior temporal, parietal, and occipital areas, but also included left hemisphere superior dorsolateral prefrontal cortical areas. Note the substantial watch and object greater than face activation located bilaterally in fusiform and occipital regions surrounding the FFA (see **Figure 2** in main text for FFA). These regions are commonly activated in object processing studies (c.f. Haxby et al., 2001; Downing et al., 2006; Reddy and Kanwisher, 2006; Spiridon et al., 2006). The only noteworthy differences seen between objects and watches relative to faces was that watches produced greater activation than faces bilaterally in dorsolateral frontal cortex, whereas no differences were observed in frontal areas between objects and faces. In the direct contrast between watches and objects, the only significant difference that was observed was objects greater than watches activation in the right angular gyrus/inferior parietal lobule region.

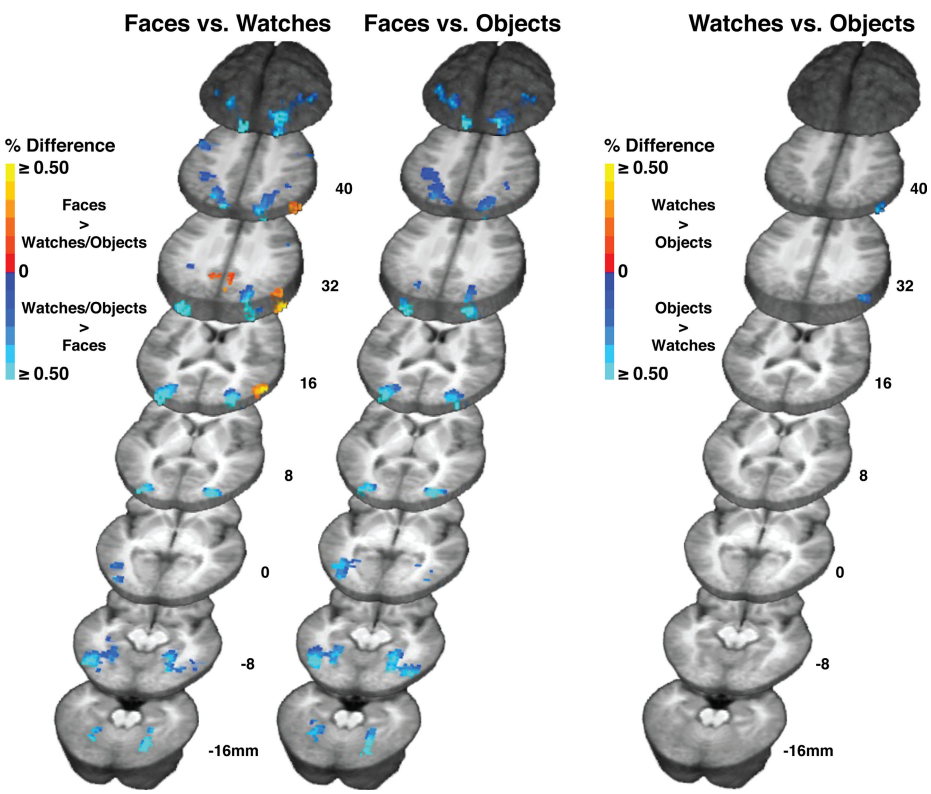



**FIGURE S1 | Whole brain activation results for faces and watches in the individuation task.** Color bars indicate percent signal change differences. Red to yellow colors indicates regions where faces produced greater activation than watches. Blue colors indicate regions where watches produced greater activation than faces. All activation differences corrected to alpha ≤ 0.05. Activation is shown on averaged structural anatomical scans from all 15 participants.

**Table S1 | Brain activation differences between face, watch, and object stimuli from the individuation task.**

| | | | Maximum intensity voxel in region | | | | |
|---|---|---|---|---|---|---|---|
| | | | Talairach coordinates | | | | |
| **Region** | **BA** | **Hemi** | ***x*** | ***y*** | ***z*** | **%** | ***t*** |
| ***FACES > WATCHES*** | | | | | | | |
| Middle temporal gyrus/Middle occipital gyrus | 39/19 | R | –55 | –66 | 18 | 0.53 | 4.94 |
| Precuneus | 31 | R | –2 | –68 | 29 | 0.29 | 4.46 |
| Posterior cingulate gyrus | 31/23 | R | –4 | –53 | 23 | 0.27 | 3.41 |
| Inferior parietal lobule/Angular gyrus | 39 | R | –49 | –67 | 38 | 0.28 | 5.00 |
| Precuneus | 31 | L | 6 | –53 | 32 | 0.21 | 3.96 |
| ***FACES > OBJECTS*** | | | | | | | |
| No regions found | | | | | | | |
| ***WATCHES > FACES*** | | | | | | | |
| Fusiform gyrus | 19 | R | –25 | –52 | –13 | 0.53 | 5.97 |
| Middle occipital gyrus | 19 | R | –31 | –77 | 12 | 0.61 | 6.05 |
| Parahippocampal gyrus | 36 | R | –28 | –38 | –10 | 0.40 | 6.01 |
| Inferior parietal lobule | 40 | R | –46 | –47 | 52 | 0.24 | 3.71 |
| Middle frontal gyrus | 9 | R | –37 | 5 | 36 | 0.26 | 4.25 |
| Superior occipital gyrus | 19 | R | –25 | –77 | 34 | 0.34 | 6.01 |
| Superior parietal lobule | 7 | R | –26 | –62 | 44 | 0.22 | 5.33 |
| Precuneus/Superior occipital gyrus | 19 | R | –22 | –80 | 42 | 0.38 | 5.47 |
| Fusiform gyrus | 37 | L | 26 | –46 | –12 | 0.37 | 5.58 |
| Parahippocampal gyrus | 36 | L | 26 | –38 | –10 | 0.27 | 4.24 |
| Middle occipital gyrus | 19 | L | 34 | –86 | 12 | 0.68 | 6.45 |
| Middle occipital gyrus/Inferior temporal gyrus | 37 | L | 44 | –58 | –6 | 0.51 | 10.24 |
| Middle frontal gyrus | 10 | L | 32 | 50 | 20 | 0.23 | 4.30 |
| Middle frontal gyrus | 6 | L | 46 | 2 | 38 | 0.29 | 6.09 |
| Inferior temporal gyrus | 19/37 | L | 44 | 58 | –6 | 0.51 | 10.24 |
| Inferior parietal lobule | 40 | L | 34 | –44 | 38 | 0.19 | 4.74 |
| Inferior parietal lobule | 40 | L | 40 | –52 | 54 | 0.33 | 3.46 |
| Superior parietal lobule | 7 | L | 22 | –68 | 44 | 0.37 | 4.04 |
| Precuneus | 7 | L | 10 | –76 | 42 | 0.30 | 4.39 |
| ***OBJECTS > FACES*** | | | | | | | |
| Fusiform gyrus | 19 | R | –28 | –52 | –12 | 0.60 | 6.88 |
| Middle occipital gyrus | 19 | R | –32 | –80 | 12 | 0.58 | 5.49 |
| Parahippocampal gyrus | 36 | R | –26 | –38 | –12 | 0.39 | 6.38 |
| Middle occipital gyrus | 37 | R | –50 | –64 | –10 | 0.49 | 3.66 |
| Middle temporal gyrus | 37 | R | –52 | –58 | –10 | 0.45 | 3.71 |
| Middle occipital gyrus | 19 | R | –32 | –80 | 14 | 0.46 | 5.95 |
| Inferior parietal lobule | 40 | R | –40 | –50 | 54 | 0.24 | 4.66 |
| Precuneus/Superior occipital gyrus | 19 | R | –28 | –70 | 36 | 0.28 | 4.56 |
| Superior parietal lobule | 7 | R | –28 | –68 | 48 | 0.26 | 4.90 |
| Fusiform gyrus | 19 | L | 26 | –58 | –12 | 0.44 | 5.64 |
| Fusiform gyrus | 37 | L | –44 | –64 | –10 | 0.46 | 4.84 |
| Parahippocampal gyrus | 36 | L | 26 | –44 | –12 | 0.50 | 6.01 |
| Middle occipital gyrus | 18 | L | 28 | –82 | 8 | 0.47 | 5.22 |
| Middle occipital gyrus/Inferior temporal gyrus | 37 | L | 44 | –62 | –4 | 0.52 | 4.69 |
| Middle occipital gyrus | 19 | L | 34 | –82 | 12 | 0.40 | 5.31 |
| Superior occipital gyrus/Cuneus | 19 | L | 32 | –86 | 24 | 0.41 | 4.85 |
| Precuneus | 7 | L | 26 | –70 | 38 | 0.28 | 4.50 |
| Superior parietal lobule | 7 | L | 26 | –58 | 44 | 0.16 | 6.02 |
| Superior parietal lobule | 7 | L | 26 | –64 | 54 | 0.26 | 5.78 |
| Inferior parietal lobule | 40 | L | 34 | –50 | 44 | 0.12 | 5.77 |
| Inferior parietal lobule | 40 | L | 38 | –52 | 54 | 0.24 | 6.32 |
| ***WATCHES > OBJECTS*** | | | | | | | |
| No regions observed | | | | | | | |
| ***OBJECTS > WATCHES*** | | | | | | | |
| Angular gyrus/Inferior parietal lobule | 39/40 | R | –50 | –68 | 36 | 0.31 | 3.91 |

*Notes: Talairach coordinates follow the convention of the Talairach and Tournoux atlas (Talairach and Tournoux, 1988). Positive coordinate values indicate left (x), anterior (y), and superior (z). BA, Brodmann's area. %, percent signal difference of contrast. t = t value of contrast.*

## SUMMARY OF WHOLE BRAIN ACTIVATION

In the main text, we argue that individuation may be a default processing procedure for faces in typically developed adults and that the equivalency of FFA activation during individuation processing for faces and non-face objects is consistent with this suggestion. The whole brain results generally suggest that this proposition extends beyond processing in the FFA. Specifically, relatively few brain regions produced greater activation to faces than watches during the individuation task. Moreover, most of the regions that did produce greater face compared to watch activation during the individuation task form what is sometimes termed an extended face-processing network, including the posterior cingulate gyrus, middle temporal gyrus, and middle occipital gyrus (Clark et al., 1996; Haxby et al., 2000; Ishai et al., 2005; Avidan and Behrmann, 2009). Therefore, the results may be interpreted to suggest that individuation processing may be a component of the extended face system, and that the extended system may have greater face specificity during individuation processing than the FFA.

## REFERENCES

Ishai, A., Schmidt, C. F., and Boesiger, P. (2005). Face perception is mediated by a distributed cortical network. *Brain Res. Bull.* 67, 87–93.

Reddy, L., and Kanwisher, N. (2006). Coding of visual objects in the ventral stream. *Curr. Opin. Neurobiol.* 16, 408–414.

Spiridon, M., Fischl, B., and Kanwisher, N. (2006). Location and spatial profile of category-specific regions in human extrastriate cortex. *Hum. Brain Mapp.* 27, 77–89.

frontiers in
HUMAN NEUROSCIENCE

REVIEW ARTICLE
published: 08 October 2010
doi: 10.3389/fnhum.2010.00166

# Worth a glance: using eye movements to investigate the cognitive neuroscience of memory

***Deborah E. Hannula[1]*, Robert R. Althoff[2,3,4], David E. Warren[5], Lily Riggs[6,7], Neal J. Cohen[8] and Jennifer D. Ryan[6,7,9]***

[1] *Department of Psychology, University of Wisconsin, Milwaukee, WI, USA*
[2] *Department of Psychiatry, University of Vermont, Burlington, VT, USA*
[3] *Department of Psychology, University of Vermont, Burlington, VT, USA*
[4] *Department of Pediatrics, University of Vermont, Burlington, VT, USA*
[5] *Department of Neurology, University of Iowa, Iowa City, IA, USA*
[6] *Rotman Research Institute, Baycrest, Toronto, ON, Canada*
[7] *Department of Psychology, University of Toronto, Toronto, ON, Canada*
[8] *Beckman Institute and Department of Psychology, University of Illinois, Urbana-Champaign, IL, USA*
[9] *Department of Psychiatry, Division of Geriatric Psychiatry, University of Toronto, Toronto, ON, Canada*

***Edited by:***
*Michael X. Cohen, University of Amsterdam, Netherlands*

***Reviewed by:***
*Walter Boot, Florida State University, USA*
*Kelly Giovanello, Cognitive Neuroscience of Memory Laboratory, USA*

****Correspondence:***
*Deborah E. Hannula, Department of Psychology, University of Wisconsin-Milwaukee, 224 Garland Hall, 2441 E. Hartford Ave., Milwaukee, WI 53211, USA.*
*e-mail: hannula@uwm.edu*

Results of several investigations indicate that eye movements can reveal memory for elements of previous experience. These effects of memory on eye movement behavior can emerge very rapidly, changing the efficiency and even the nature of visual processing without appealing to verbal reports and without requiring conscious recollection. This aspect of eye movement based memory investigations is particularly useful when eye movement methods are used with special populations (e.g., young children, elderly individuals, and patients with severe amnesia), and also permits use of comparable paradigms in animals and humans, helping to bridge different memory literatures and permitting cross-species generalizations. Unique characteristics of eye movement methods have produced findings that challenge long-held views about the nature of memory, its organization in the brain, and its failures in special populations. Recently, eye movement methods have been successfully combined with neuroimaging techniques such as fMRI, single-unit recording, and magnetoencephalography, permitting more sophisticated investigations of memory. Ultimately, combined use of eye-tracking with neuropsychological and neuroimaging methods promises to provide a more comprehensive account of brain–behavior relationships and adheres to the "converging evidence" approach to cognitive neuroscience.

**Keywords: eye movements, fMRI, MEG, memory, hippocampus, amnesia**

In this review we advocate for the use of eye movement monitoring as a powerful tool that can advance the field of cognitive neuroscience. Because eye movement based investigations of attention and language have been described comprehensively in several excellent reviews (cf. Van der Stigchel et al., 2006; Rayner, 2009; Theeuwes et al., 2009), emphasis is placed on investigations of memory; however, the general approach is expected to have merit for other domains of investigation as well. In the sections that follow we will show that viewers' eye movements can reveal memory for elements of previous experience without appealing to verbal reports and without requiring conscious recollection. We also point to research which has shown that the effects of memory on eye movement behavior can emerge very rapidly, changing the efficiency and even the nature of visual processing. These effects have been characterized as obligatory (and perhaps automatic) and are consistently reported independent of task instructions or viewers' intentions. Taking advantage of these properties, several important questions have been addressed and certain long-held views about the nature of memory, its organization in the brain, and its failures in certain special populations have been challenged. In general, a converging methods approach is highlighted, with special emphasis placed on the utility of eye movement monitoring when it is used in conjunction with other cognitive neuroscience methods.

Measurement of memory with eye movements is powerful because the nature of visual processing requires that we sample one region of the world at a time by directing the high-acuity portion of our retinas onto successively sampled regions. This movement of the eyes across the visual world is not random, but rather is dictated by two factors. The first concerns stimulus characteristics – the physical properties (e.g., luminance, hue) of the elements embedded in visual arrays (e.g., Buswell, 1935; Mackworth and Morandi, 1967; Antes, 1974). The second concerns our previous experience (i.e., episodic memory) and the knowledge we bring to a particular viewing situation (i.e., semantic memory). We look more at objects that are the targets of volitional search, but also dwell on objects that are discrepant with expectations, based, for example, on general knowledge about the context in which they are embedded (e.g., an octopus in a barnyard scene; Loftus and Mackworth, 1978), or on previous viewing history in a laboratory setting (see Revealing Distinct Types of Mnemonic Information).

When control of basic physical properties has been achieved it becomes possible to discern effects of memory on eye movement behavior. An early example of this strategy comes from a report by Yarbus (1967). In this work, eye movements were recorded as a viewer examined a picture by the Russian artist Repin depicting the unexpected return of a man, who is shown entering a room

filled with various people and objects. Eye movements elicited by this scene under free-viewing conditions were largely to prominent visual elements – the man and objects in the room. The same viewer was then asked to examine the scene again while providing answers to several different questions (e.g., how old are the people in the scene, what might they have been doing before the arrival of the unexpected visitor). Different patterns of viewing were observed across these instructional conditions, presumably reflecting (pre-experimental) knowledge the viewer brought to bear about where in the scene relevant information would be located.

In what follows, we illustrate just how powerful eye movement methods can be for addressing important questions about memory. In short, by using eye movements to assess memory, rather than merely relying on verbal reports or introspective judgments, we gain the ability to test memory under circumstances in which behavioral reports may not (or cannot) be reliably obtained. This pays particular dividends when eye movement methods are used with special populations (e.g., young children, elderly individuals, and patients with severe amnesia), and also permits use of comparable paradigms in animals and humans, helping to bridge different memory literatures and permitting cross-species generalizations.

To best illustrate the use of eye movements to advance the cognitive neuroscience of memory, we will provide an extended example of how eye movement studies offered a way to resolve a major disagreement in the field about the nature of amnesia and the role of the hippocampus in memory (see Revealing a Critical Role for the Hippocampus in Memory Without Awareness). In this work (Althoff, 1998; Ryan et al., 2000; Hannula et al., 2007), eye movement methods provided critical data about the role of hippocampus in relational memory, as distinct from the role it may play in explicit memory or conscious recollection. Accordingly, we will discuss and attempt to resolve a number of controversies about memory, awareness, behavioral and eye movement measures, and the hippocampus. Promising results from some recent efforts to apply these methods to other patients and special populations will also be presented.

Finally, and as alluded to above, we will discuss the particular strengths of eye movement measures as part of a larger converging methods approach to the study of memory, outlining some promising early steps in relating behavioral, eye movement, and neuroimaging measures (Herdman and Ryan, 2007; Ryan et al., 2007b; Hannula and Ranganath, 2009; Riggs et al., 2009; see Integration of Eye Movement Monitoring With Neuroimaging). Including eye movement monitoring in neuropsychological and neuroimaging investigations provides a comprehensive account of the brain-behavior relationship and adheres to the "converging evidence" approach to cognitive neuroscience.

## REVEALING DISTINCT TYPES OF MNEMONIC INFORMATION

Before the field of cognitive neuroscience could use eye movement methods to link the function of specific neural regions to particular aspects of memory, it was necessary to produce evidence showing that memory does indeed influence eye movement behavior. This section provides a very brief, and necessarily selective, discussion of studies aimed at documenting eye movement based memory effects.

An early line of work on eye movements and memory was guided by the "scanpath hypothesis" (Noton and Stark, 1971), which postulated that recognition occurs when the same scanning pattern enacted during initial viewings is recapitulated during subsequent viewing and recognition of a stimulus. Consistent with this hypothesis, striking similarity in scanning patterns between initial inspection and subsequent recognition was indeed observed by some investigators for a variety of materials and tasks (Walker-Smith et al., 1977; Fisher et al., 1978; Parker, 1978). However, these investigators interpreted their results as suggesting that during recognition (e.g., for a face), a participant would move to features *more salient to him or her* after first registering the presence of the object. They proposed a filtering system that would interpret incoming data from the periphery, identify the most salient feature, and then direct a saccade towards it (Didday and Arbib, 1975; see also Carpenter and Just, 1977). Starting with this work and continuing through ever more sophisticated studies, multiple measures were developed to characterize the types of changes that might occur in eye movement behavior that reflect – and, thus, that reveal – the operation of mnemonic processes (see **Boxes 1 and 2**).

Robust effects of semantic memory on the way in which we evaluate or extract information from a visual stimulus have been clearly demonstrated (cf. Henderson, 2003). For example, general world knowledge about the context in which objects are typically found (e.g., a tractor is usually found in a barnyard scene) and their relative positions within the environment (e.g., a toaster is typically on a counter) has been found to facilitate the speed with which the eyes detect a visual target (e.g., Loftus and Mackworth, 1978; Henderson et al., 1999; Brockmole and Henderson, 2008; Hollingworth, 2009; see **Figure 1**). Target detection is also facilitated by repeated exposure to a specific scene context in the lab (Brockmole et al., 2006), and by brief exposure to, or preview of, that scene (Castelhano and Henderson, 2007; Hollingworth, 2009).

In our work, we have examined how eye movements reveal memory for pre-experimentally familiar materials and specific prior experiences in a series of studies conducted with images of famous and non-famous faces (Althoff and Cohen, 1999), familiar and unfamiliar buildings (Althoff et al., 1998), and novel scenes (Ryan et al., 2000). In these and subsequent studies, eye movement behavior revealed memory for previous occurrence of these various types of items. Pre-experimentally familiar items were viewed with fewer fixations and with lower constraint on the location of successive fixations than were novel items. Furthermore, with repeated exposure to pre-experimentally unfamiliar items, the amount of sampling decreased systematically (see **Figure 2**; Althoff, 1998; Althoff et al., 1998; Heisz and Shore, 2008; Heisz and Ryan, submitted). This repetition effect was observed whether participants performed a recognition task or an emotion labeling task (i.e., regardless of task demands; Althoff et al., 1998).

Subsequent experiments, which used pre-experimentally unfamiliar materials, demonstrated that eye movements could also reveal relational memory for the spatial positions of elements within scenes (e.g., Ryan et al., 2000; Ryan and Cohen, 2004a,b), arbitrary pairings of faces and scenes (Hannula et al., 2007; Hannula and Ranganath, 2009), and temporal sequences (Ryan and Villate, 2009). In the work with scenes (Ryan et al., 2000), participants viewed several pictures, each accompanied by a question regarding the relative relationships

**BOX 1 | Eye movement measures of memory.**

Eye movement data can be compiled and analyzed in several different ways. Sampling of visual materials can be characterized either at the level of an entire experimental display (i.e., overall viewing), or at the level of regions, objects, or stimuli within that display (i.e., directed viewing). For example, eye movements to a display containing several faces could be summarized as the total number of fixations made to the display. Alternatively, fixations to the display could be categorized according to their targets (e.g., novel and studied faces), and summarized separately. Measures of directed viewing divide a display into more than one region of interest (ROI) in order to permit evaluation of effects related to independent variables of interest. Directed viewing measures may also be combined with temporal indices in order to gauge changes in a particular measure over time (i.e., time-course analyses), or may be time-locked to a particular event, such as an overt behavioral response (i.e., response-locked analyses). Various commonly used characterizations of overall and directed viewing are defined below.

**Measures of overall viewing**

- **Number of fixations[1]:** the number of discrete pauses of the eyes for a display.
- **Fixation duration[1]:** the length of time in which the eye pauses on a display, typically between 200–300 ms long. Median or mean fixation duration to a display can be calculated.
- **Saccade amplitude:** the distance traversed between successive fixations, reported in degrees per second.
- **Number of regions fixated[2]:** the number of discrete regions sampled within a display.
- **Number of transitions between regions:** the number of transitions made by the eyes between discrete regions.
- **First return fixation:** the number of fixations made before returning to a previously sampled region.
- **First-order entropy[3]:** the predictability of the transitions between the locations of a given fixation and the preceding fixation.
- **Second-order entropy[3]:** the predictability of the transitions to a given fixation location based on the location of the two immediately preceding fixations.
- **Chi-square, Asymmetric lambda[2]:** other measures used to quantify the randomness of an eye movement transition table.

**Measures of directed viewing**

- **Proportion of fixations:** the proportion of total fixations that are directed to an experimenter-defined ROI.
- **Proportion of time[2]:** the proportion of total viewing time that is directed to an experimenter-defined ROI.
- **Number of transitions into/out of a critical region:** the number of gaze transitions into/out of an experimenter-defined ROI.
- **Duration of the first gaze:** total viewing time to an experimenter-defined ROI on the first gaze that is directed into that ROI.
- **Number of fixations in the first gaze:** number of fixations to an experimenter-defined ROI during the first gaze that is directed into that ROI.

**Considerations for Analysis**

While the above outlines general definitions for the predominant measures derived from eye movement monitoring, there are differences in how such measures are calculated. For instance, the definition of a *fixation* may vary between research groups and/or eyetracking platforms. Successive recording samples of eye position may be considered as a single fixation if changes in gaze position across samples are less than 1° of visual angle and, when combined, have a minimum duration of 100 ms (e.g., Hannula et al., 2007). Alternatively (e.g., Ryan et al., 2007), a fixation can be defined as the absence of any *saccade* (e.g., the velocity of two successive eye movement samples exceeds 22° per second over a distance of 0.1°), or blink (e.g., pupil is missing for three or more samples) activity. In turn, definitions for saccades and blinks may vary across labs/platforms.

Regarding the *number of regions fixated,* in our work (cf. Althoff and Cohen, 1999) this number is defined based on the pattern by which each participant's fixations are clustered; that is, fixations that fall within an *a priori* specified distance from one another are considered to be within the same region, while fixations outside of this distance are considered to belong to a separate region. This approach takes into account individual differences in viewing behavior as well as differences in particular stimulus characteristics, such as the size or prominence of distinct features on an object. An alternative approach uses a fixed grid with equally-sized sections for all of the stimuli and all of the participants indiscriminately (e.g., Smith and Squire, 2008). This latter approach has an advantage of standardizing the number of possible regions that could be fixated across stimuli and participants, but this approach may lack precision and/or sensitivity. For example, if an object attracts disproportionate viewing, and occupies two or more sections of the grid, then the measure of sampling behavior may be artificially inflated; by contrast, if two smaller objects occupy the same section, and each is distinctly fixated by the viewer, the number of unique regions fixated is under-sampled.

At least two different approaches have also been used to calculate *the proportion of time* that is directed to a ROI. In our work, the proportion is considered in reference to a baseline in which only the amount of viewing time that is directed to the stimulus (e.g., a scene) is counted, excluding any time spent looking outside of the display, blinking, or making saccadic movements of the eyes (e.g., Ryan et al., 2000; Hannula et al., 2007). The alternative approach (e.g., Smith and Squire, 2008) involves calculating the proportion of viewing time using the entire trial period as the baseline, regardless of how much time was spent not actually viewing the stimulus. The number of saccades, as well as the number of blinks, varies from image to image, both within and across participants; therefore, the actual time that is spent inspecting a stimulus is never constant. Thus, the use of total trial duration in the denominator will either artificially inflate or reduce the proportion of viewing time measure differentially across trials, participants, and experimental conditions. As such, we do not advocate this latter approach as it introduces noise into the data.

[1]also measures of directed viewing

[2]please refer to Considerations for Analysis (this box) for more information about calculating this measure

[3]see Box 2 and Althoff and Cohen (1999)

among scene elements (e.g., is there a cat to the left of the boy?). During a final, critical block, eye movements were monitored as participants viewed scenes that were either novel (i.e., not studied), repeated (i.e., unchanged from previous exposures), or manipulated (i.e., with a change in spatial relationships among elements). Critically, a manipulated scene for one participant was repeated for

**BOX 2 | Markovian measures.**

To determine whether there is particular top-down influence of cognition on eye movement patterns, several investigators turned to the examination of scanpaths. To describe scanpaths, contingency tables are created; these tables examine the probability of moving from one ROI to another within a display and can be probed for their structure. The idea that eye movements followed a first-order Markovian process (i.e., a stochastic process in which any transition between states depends only the current state) had been hypothesized as early as 1966 by Molnar (as described in Molnar and Ratsikas, 1987). The simplest examination of this structure is a chi-square test which examines the probability of seeing a particular pattern in the contingency table over and above what would be expected by chance. Stark and Ellis (1981) applied a chi-square goodness of fit calculation to determine the difference between observed and expected matrices $\mathbf{M}_{(1)}$ where the expected matrix was based on the stratified random situation (i.e., the situation where a particular saccade could be predicted entirely by the marginals of the contingency table). The results from this experiment yielded the expected finding that there were significant differences between the observed and expected first-order dependencies as measured by chi-square. To calculate the change in the direction of dependency in those data, he further calculated the conditional *information* $H_c$ in both the observed contingency table $\mathbf{M}_{(1)}$ and the expected contingency table $\mathbf{M}_{(1)}$ (see Stark and Ellis, 1981; Althoff and Cohen, 1999 or Althoff, 1998 for calculation details). This provided a measure of statistical dependency in the matrix (Brillouin, 1962); the larger the $H_c$, the less statistical dependency there was in the matrix, and the more random the eye movements were. One could, therefore, simply take this information as a measure of the amount of randomness in the signal, with a high degree of entropy ($\boldsymbol{H}_c$) indicating a high degree of randomness. Further refining these methods, Ellis and Stark (1986) modeled the behavior of pilots in viewing a cockpit display of traffic information by examining whether sampling of the regions of the a display was "random", "stratified random" (i.e., dependent only on the information on the marginals of a contingency table) or "statistically dependent" (i.e., with dependencies in the data that linked movements from one region of the display to another). In our work, we have termed these variables $H_1$ for the first-order transitions and $H_2$ for the second-order transitions.

Since different participants have different path lengths (i.e., number of fixations), Tole et al. (1982) normalized these entropy measures for the number of fixations to show a decrease in the entropy of a scanning pattern when pilots were under an increased workload. In other words, as pilots became stressed, their scanning patterns became less random and more structured, revealing the top-down influence of cognition and task demands on eye movement behavior. These types of entropy measures have been used by our group (termed $S_1$ and $S_2$) and others to demonstrate that viewing of an unfamiliar face or scene is typically highly constrained (i.e., has low entropy) with viewers tending to follow an idiosyncratic pattern while scanning novel stimuli (Althoff and Cohen, 1999) which becomes less so as faces or scenes have been viewed repeatedly.

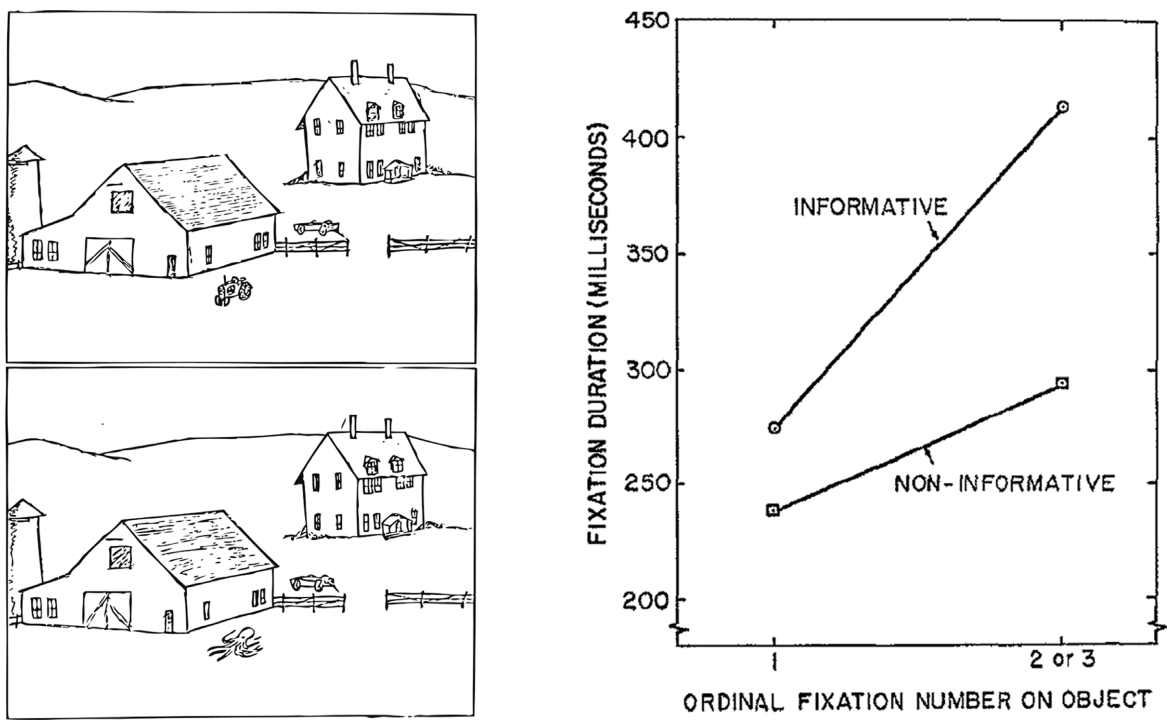



**FIGURE 1 | Illustration of the experimental methods and results from Loftus and Mackworth (1978).** Participants viewed scenes in which an object was either consistent (i.e., a tractor; top), or inconsistent (i.e., octopus; bottom) with the semantic context of the scene (i.e., barnyard). Objects that were inconsistent with the scene context were viewed with longer fixation durations than non-informative objects, and this increase in viewing occurred within the first few fixations, revealing early and obligatory influences of prior knowledge on eye movements.

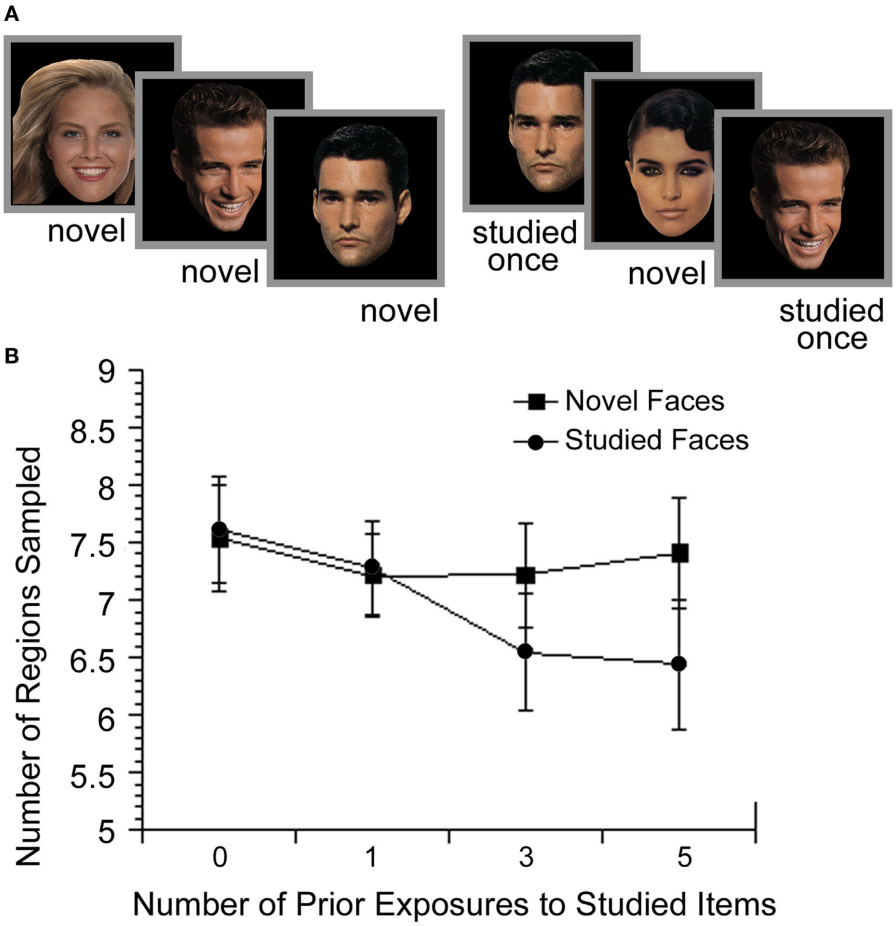



**FIGURE 2 | Illustration of the experimental methods and results from Althoff (1998). (A)** Participants viewed non-famous faces either 0 (novel), 1, 3 or 5 times throughout the experiment while their eye movements were monitored. **(B)** With increasing exposure, participants sampled fewer regions of the studied faces, compared to novel (non-studied) faces that were introduced during the same phase of the experiment. These findings illustrate the influence of memory for the previously viewed faces on subsequent processing.

another, permitting comparison of viewing directed to the exact same regions across participants for whom the only difference was in viewing history. Results showed more fixations to, and transitions into and out of, the critical region(s) when scenes were manipulated versus repeated. This relational manipulation effect was documented in four separate experiments, and was evident whether participants were instructed to identify changes or were merely instructed to view the scenes (i.e., free-viewing). Furthermore, it did not depend upon use of orienting questions meant to direct a viewer's attention and gaze to regions of scenes that might ultimately be manipulated, as it was documented even when these questions were not used (see Ryan et al., 2000 for details; see also Ryan and Cohen, 2004a). Particularly compelling were findings of greater viewing directed to "now-empty" manipulated regions as compared to the exact same regions when they were "always empty" (see **Figure 3**), revealing the effects of relational memory on current processing.

Relational memory for arbitrary scene-face pairings has also been revealed in viewers' eye movement behavior (Hannula et al., 2007; Hannula and Ranganath, 2009). In this work, participants studied several scene-face pairs and were tested with 3-face displays superimposed on previously studied scenes. Participants showed disproportionate viewing of the face that matched (i.e., had previously been paired with) the scene, even among equally familiar faces (i.e., equated in terms of previous viewing history) and in the absence of any reliable spatial cues to guide choices (i.e., the matching face could be in any of three spatial locations, none of which matched the original presentation location; see **Figure 4**).

Finally, memory for temporal relations was shown in eye movement behavior in an experiment in which participants were presented with three objects each shown one at a time in different spatial locations during the study phase and then presented simultaneously after a short delay (Ryan and Villate, 2009). While all three objects were shifted in their absolute position, the relative positions of objects in the display with respect to one another were either intact, or were manipulated by displacing one object with respect to the others. Despite simultaneous presentation during the test phase, participants tended to inspect the objects in the order that matched the originally experienced temporal sequence; this tendency decreased when spatial relationships were manipulated.

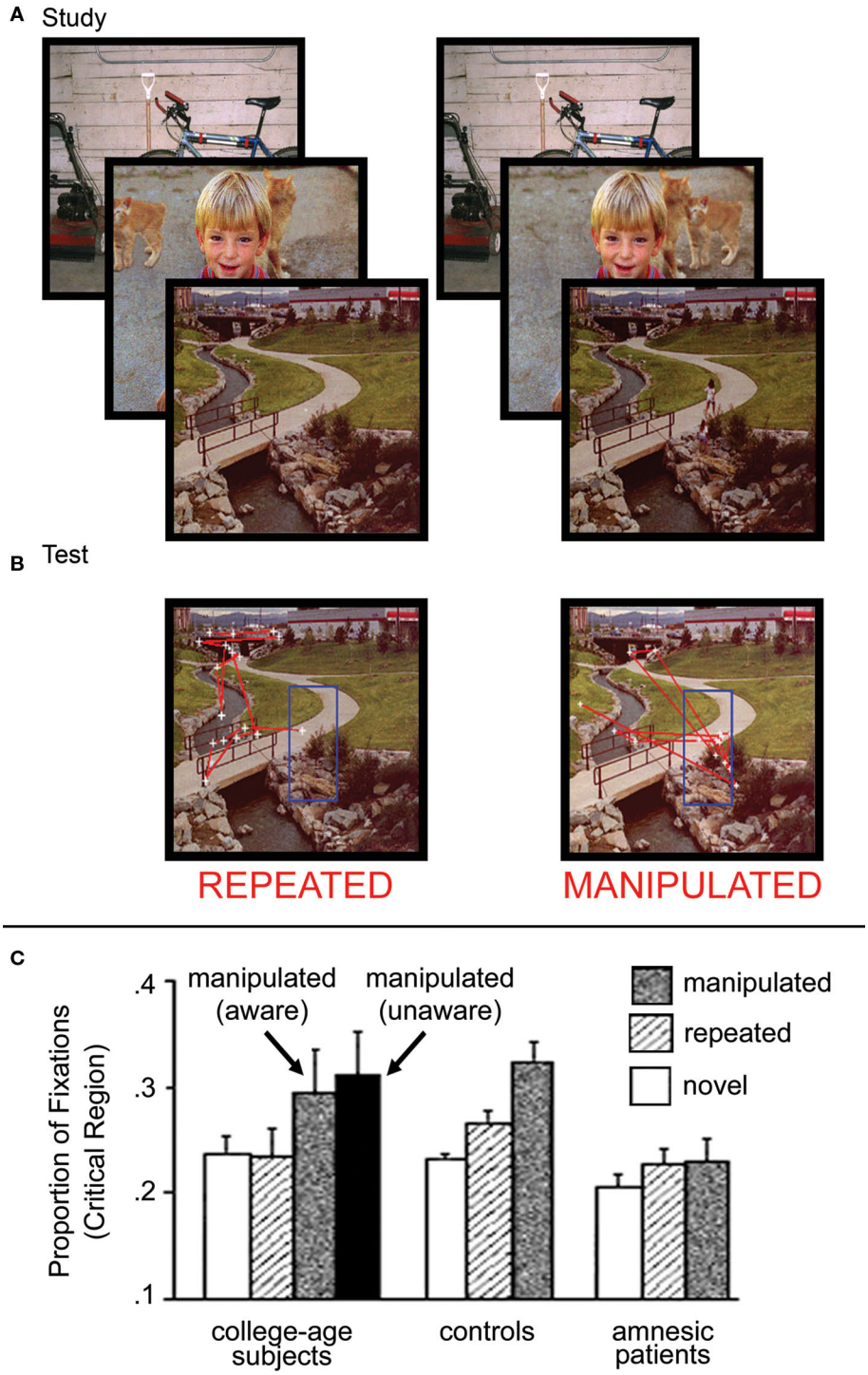



**FIGURE 3 | Illustration of the experimental methods and results from Ryan et al. (2000).** **(A)** Scenes presented during the study phase and **(B)** during the test phase, as seen by two different participants. In this case, the scenes in the test phase are exactly the same, and differ only with respect to the previous viewing histories of the two participants. For one participant, the scene is repeated (i.e., unchanged) from the study phase, whereas for the other participant, the scene has been manipulated (i.e., the girls near the bridge have been removed from the scene). Eye movements, superimposed on each scene, illustrate the eye movement based relational memory effect. More fixations (i.e., white cross hairs) are directed to, and more saccades (i.e., red lines) are made into and out of the critical region for manipulated scenes. **(C)** Graph illustrating the proportion of total viewing time directed to the critical regions of novel, repeated and manipulated scenes for college-age participants, neurologically intact controls, and amnesic patients. College-age participants look disproportionately at manipulated regions with or without concomitant awareness of the change. Neurologically intact controls also show this disproportionate viewing effect, which is completely absent from the viewing behavior of amnesic patients.

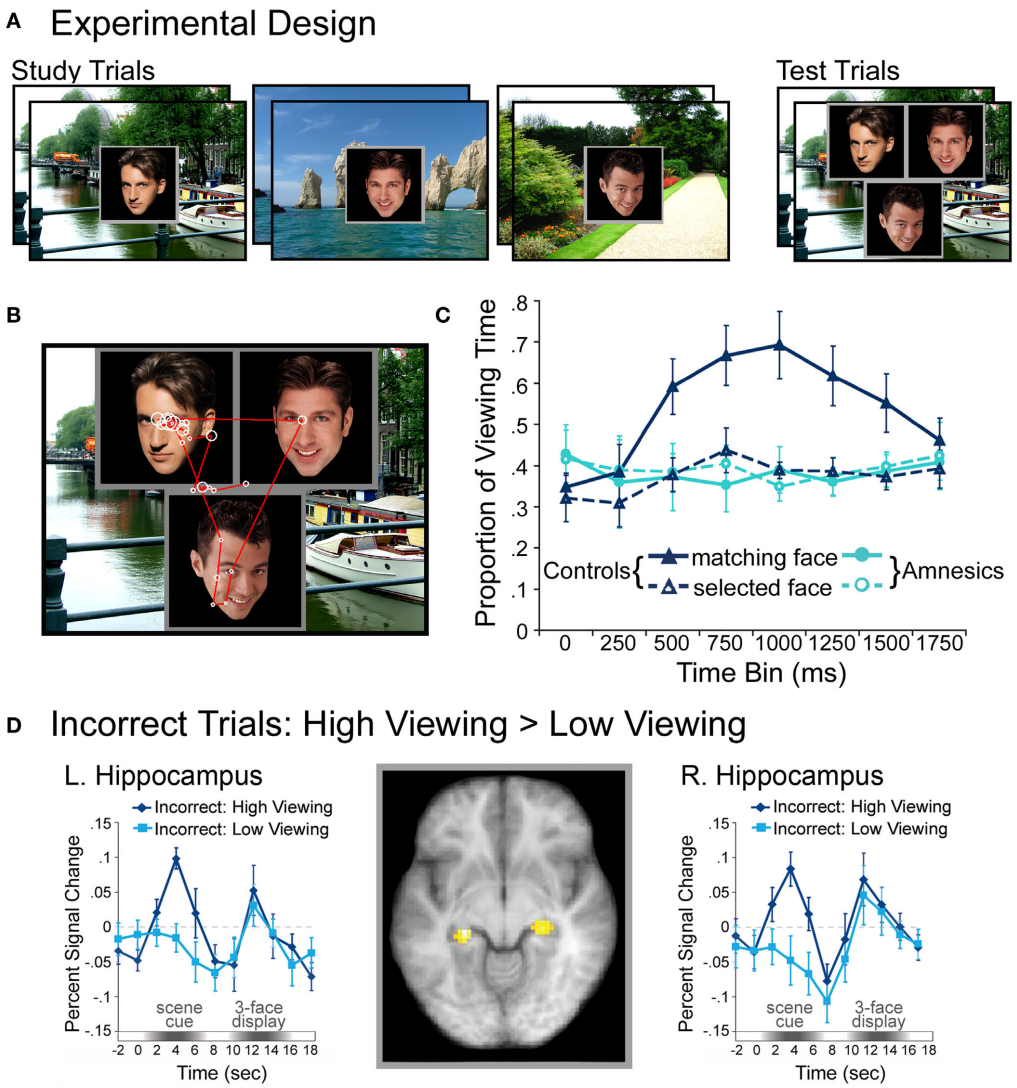



**FIGURE 4 | Illustration of the basic paradigm and results from Hannula et al. (2007) and Hannula and Ranganath (2009).** **(A)** Examples of scene-face pairs presented during the study trials, along with a single, associated test display in which the face on the left was the match (i.e., was the associate of the scene). Each test trial began with the presentation of a scene cue meant to prompt retrieval of the associated face. **(B)** Eye movements from a representative participant superimposed on the test display shown above. Fixations are indicated by white circles and the size of the circle was proportionate to the amount of viewing time directed to the fixated region. Transitions from one fixation to the next are indicated by red lines. **(C)** Proportion of total viewing time directed to correctly identified matching faces vs. faces that were merely selected from displays that did not contain a match broken down into 250 ms time bins following the onset of the three-face test display. Neurologically intact control participants showed disproportionate viewing of the matching face just 500–750 ms after the faces were presented; no evidence of relational memory was evident in the eye movement behavior of amnesic patients. **(D)** Bilateral regions of the hippocampus for which BOLD signal was greater for incorrect trials during presentation of the scene cue when subsequent viewing of the match was high vs. when subsequent viewing of the match was low in college-age participants. Trial-averaged time courses extracted from the left and right hippocampal regions, respectively, show differences in BOLD signal between incorrect high and incorrect low viewing trials during presentation of the scene cue. This result illustrates hippocampal recruitment, even when explicit memory has failed.

Taken together, these results indicate that memory for different aspects of experience (e.g., individual items, spatial and non-spatial relationships, temporal order) guide eye movement behavior. In the sections to follow, we show that the effects of memory on eye movement behavior can occur very rapidly and obligatorily, even in the absence of conscious awareness, which affords us a powerful tool with which to examine memory in a variety of special populations and in conjunction with various other cognitive neuroscience methods.

## REVEALING MEMORY PRIOR TO CONSCIOUS AWARENESS

The observation that effects of memory on eye movement behavior develop quickly was initially reported by Parker (1978) who showed that manipulated regions of previously viewed scenes were fixated earlier in scanning than unchanged regions of the same scenes. This rapid acquisition of remembered content with the eyes is accompanied by extended fixation duration (e.g., Ryan and Cohen, 2004a; Ryan et al., 2007a) suggesting

that eye movements might be used to gather evidence that can inform subsequent behavioral responses. Consistent with this idea, recent work has shown that participants look disproportionately at fragmented target objects embedded in distractor-filled displays as many as 25 fixations before explicit object identification; systematic and nearly exclusive evaluation of the target region was evident four fixations prior to naming (Holm et al., 2008).

Extending the above observations, time-course and response-locked measures (see Box 1) commonly used in event-related potential investigations have recently been adapted to examine when, with respect to stimulus onset and response execution, eye movement based memory effects emerge (Hannula et al., 2007; Ryan et al., 2007a; Hannula and Ranganath, 2009). Using these measures, relational memory effects were documented in the face-scene experiment described earlier just 500–750 ms after the onset of the 3-face display (see **Figure 4**), and as much as 1000 ms prior to explicit behavioral responses. The time-course of this eye movement based relational memory effect was impervious to manipulations of task instructions, emerging within 500–750 ms when participants were explicitly instructed to identify the matching face, and emerging within the exact same time frame even when viewing of the match was counterproductive to the task at hand. That these rapid disproportionate viewing effects were uninfluenced by task demands and develop so far in advance of behavioral responding provides strong evidence for the obligatory nature of memory on eye movement behavior, and suggests that such effects may precede and contribute to conscious recollection of the previously learned association (see Moscovitch, 2008; Hannula and Ranganath, 2009). Collectively, these results indicate that eye movements provide a "temporally precise measure that indexed the evolution of … memory expression from perception to action" (Kumaran and Wagner, 2009, p. 563), and hint at the possibility that eye movements might reflect remembered content even when explicit (conscious) recognition has failed.

## REVEALING MEMORY IN THE ABSENCE OF AWARENESS

A suggestion that eye movements may be decoupled from conscious awareness comes from the attention literature. It has been shown, for example, that task-irrelevant abrupt onsets (e.g., visual stimuli that suddenly appear in a display) capture viewing even when there is volitional effort to avoid them, and even when there is no conscious awareness for the action itself (Kramer et al., 2000; Belopolsky et al., 2008). Converging evidence from several investigations indicates that eye movements can also be influenced by higher-level cognitive processes, like memory, without concomitant awareness for prior learning episodes.

As described above, when the relationships among elements in a scene have changed, eye movements are drawn disproportionately to the manipulated region (e.g., Ryan et al., 2000). Results from several experiments have confirmed that these changes (Ryan and Cohen, 2004a; Beck et al., 2007) and others like them (Hayhoe et al., 1997; Hollingworth et al., 2001, 2008; Hollingworth and Henderson, 2002; Henderson and Hollingworth, 2003) are reflected in modulations of eye movements, even when they go unreported by participants. For example, gaze duration to an object embedded in a scene is significantly longer when it is replaced with a different exemplar, even when participants fail to explicitly detect the change (e.g., Hollingworth et al., 2001). Along similar lines, recent work using the visual paired comparison task, in which a novel object and a studied object are presented simultaneously as eye movements are recorded, has shown that viewing is modified based on the type of associate a learned object was paired with during a study exposure. Novelty preferences in patterns of eye movements were disrupted when learned objects had been paired with positively or negatively valenced scenes, but were intact when learned objects had been paired with a neutral gray image. Differences across these conditions were not evident in explicit recognition responses, suggesting that different memory processes support eye movement based expressions of memory and explicit behavioral responses in this task (Snyder et al., 2008). Collectively, these results suggest that eye movement based memory effects can be expressed even when conscious recollection has failed or is non-diagnostic, which, as we will see in Section "Revealing a Critical Role for the Hippocampus in Memory Without Awareness", proves important in supporting conclusions from eye movement investigations of the nature of amnesia and the functional role of the hippocampus.

The purpose of these memory-guided eye movements is worth considering, especially since evidence suggests that they may precede or be unaccompanied by conscious awareness. In general, eye movements are used to extract information from the environment; they permit us to continually compare prior experience with current perceptual input in order to detect novelty/change and to guide our subsequent behavior (e.g., make decisions, navigate; Ryan and Cohen, 2004a). Irrespective of whether conscious awareness plays a role, the intersection of current perception with remembered content ensures that perceptual processing proceeds efficiently and rapidly, e.g., by allowing attention to be biased towards particular regions of interest as a function of past experience without the need for supervisory control (Chun and Nakayama, 2000; Maljkovic and Nakayama, 2000).

As time-course analyses have suggested (e.g., Hannula et al., 2007), the rapid influence of memory on eye movements may permit eventual conscious appreciation for items and/or relations that had been previously viewed. That is, conscious awareness may be derived from changes in our eye movements, rather than eye movement based memory effects resulting from conscious awareness for prior learning episodes (Ryan and Villate, 2009). Under most circumstances, eye movements may ultimately be correlated with conscious recognition (e.g., later in processing), but converging evidence from several investigations (see above) indicates that awareness is not a requirement for the expression of memory in eye movement behavior. Precise factors that determine whether eye movements will be correlated with, and give rise to, overt expressions of memory or, alternatively, will be expressed in the absence of explicit awareness, have yet to be identified. However, existing evidence does confirm that eye movements provide a sensitive measure with which to investigate the manner in which multiple memory systems create, access and update mnemonic representations (Ryan et al., 2000; Brockmole et al., 2002, 2003; Brockmole and Irwin, 2005; Ryan and Villate, 2009).

## REVEALING A CRITICAL ROLE FOR THE HIPPOCAMPUS IN MEMORY WITHOUT AWARENESS

Nowhere are the effects of multiple memory systems more evident than in the study of special populations, especially those involving patients with amnesia following damage to the hippocampus and related medial temporal lobe (MTL) structures. We will develop this example in some detail, as it makes a strong case for the promise of eye movements as part of a converging methods approach. Not only does this example provide data critical for resolving debates about the nature of the memory deficit in hippocampal amnesia and the fundamental role of the hippocampus in memory; it also affords an important illustration of the power of eye movement data when combined with other cognitive neuroscience methods.

The nature of the impairment in amnesia and the functional role of the hippocampal system in memory have been the subject of intense study ever since the report of profound and pervasive memory impairment following bilateral removal of the hippocampus in patient H.M. (Scoville and Milner, 1957). A major advance came from findings of preserved learning abilities in amnesia – while some aspects of learning and memory were profoundly impaired, others were left fully intact (e.g., Milner, 1962; Milner et al., 1968; Cohen and Squire, 1980; Graf and Schacter, 1985). Such findings, and many that followed, led to the emergence of theories about multiple memory systems of the brain (Cohen, 1984; Tulving, 1985; Schacter, 1987; Squire, 1992; Cohen and Eichenbaum, 1993; Gabrieli, 1998; Eichenbaum and Cohen, 2001).

One multiple memory systems view that received considerable support distinguished explicit from implicit memory, emphasizing the role of the hippocampus in conscious recollection or conscious awareness of prior experiences (Graf and Schacter, 1985; Schacter, 1987; Squire, 1992). According to this view conscious awareness for prior learning episodes is a fundamental property of the memory representations that are formed by the hippocampus, with hippocampal representations necessarily available to conscious introspection (c.f., Manns and Squire, 2001; Smith and Squire, 2008). Preserved vs. impaired performances in hippocampal amnesia would therefore revolve around the necessity of conscious access to memories for successful performance.

An alternative view (Cohen and Eichenbaum, 1993; Eichenbaum and Cohen, 2001) has emphasized the role of the hippocampus in relational memory binding (i.e., the formation of representations that consist of relationships among constituent elements of scenes or events). On this view, although the ability for conscious recollection of prior episodes requires relational memory for the items that comprise those prior episodes and/or the contexts in which they were encountered, preserved vs. impaired performance in hippocampal amnesia instead revolves around the demand for relational memory (i.e., leaving memory for individual items intact). By this account, relational representations may be expressed without concomitant conscious awareness.

The challenge in adjudicating between these theories is to vary relational memory demands independently of conscious access demands. That is, because tests involving conscious recollection as the measure of performance also involve relational memory (Ryan and Cohen, 2003), it was necessary to step outside of standard recognition memory testing paradigms. Based on our findings that eye movements provide a means for indirect testing and expression of both item memory and relational memory, we applied eye movement studies to answer this fundamental question about memory and amnesia. Findings revealed that patients with hippocampal amnesia had selective relational memory impairments, sparing item memory, independent of issues of conscious access.

Despite amnesia, patients exhibited normal eye movement based memory effects for individual items, showing decreased sampling in the viewing of repeated, as compared to novel, faces (Althoff et al., 1993; Althoff, 1998) and scenes (Ryan et al., 2000). In the case of faces, one patient with severe amnesia presented with famous and non-famous faces showed intact repetition effects despite impaired recognition of the viewed faces (Althoff et al., 1993). Similarly, seven amnesic patients tested in a different experiment with sets of non-famous faces viewed either 0, 1, 3, or 5 times across a study session showed significant changes in viewing (on measures of constraint and the number of regions sampled) with increased repetitions. These changes in viewing were evident despite recognition rates that remained at chance and did not improve with repeated exposure to the faces (Althoff, 1998).

By contrast, eye movement measures of relational memory were found to be impaired in hippocampal amnesia (Ryan et al., 2000; Hannula et al., 2007). In the Ryan et al. (2000) study of novel, repeated, and relationally manipulated scenes, neurologically intact controls looked disproportionately at regions of scenes that had undergone a change in spatial relationships, even when they were unaware of the nature of the changes that had occurred; such eye movement behavior was absent in amnesic patients (see **Figure 4**). However, as noted above, the same patients showed normal repetition effects for the scenes, revealing normal memory for items. In the Hannula et al. (2007) study of arbitrary face-scene pairings, neurologically intact control subjects presented with test displays composed of three equally familiar faces superimposed on a familiar scene showed early preferential viewing of the one face that had actually been paired with the scene during the study phase; amnesic patients failed to show the normal effect of relational memory in their eye movement behavior. In a follow-up investigation that combined eye movement monitoring with fMRI (Hannula and Ranganath, 2009), hippocampal activity during the scene cue was shown to predict preferential viewing of the matching face, even when participants failed to explicitly identify the associated face (see **Figure 4**).

Taken together, the use of eye movement paradigms for indirect testing of item and relational memory provided the field with critical evidence in favor of a relational memory account over an explicit memory account with respect to the impairment in hippocampal amnesia, and the functional role of hippocampus in normal memory. Memory for single items was intact while memory for relations among the items was impaired, despite using methods by which both types of memory could be observed without requiring conscious recollection. This example also illustrates the great promise of converging cognitive neuroscience methods that include eye movement measures. Using the same paradigm with combined eye-tracking and fMRI methods, Hannula and Ranganath (2009) linked the deficit seen in hippocampal amnesic patients to hippocampal activity in normal individuals, and linked hippocampal activity to the eye movement based relational memory effect independent of explicit remembering.

Some of these findings, and the corresponding conclusions that were drawn with regard to the hippocampus and memory, have been challenged. Based on their failure to replicate some of the results reported by Ryan et al. (2000), Smith and Squire (2008) concluded that eye movement based memory effects were only expressed with accompanying conscious awareness for prior learning episodes, and that both item and relational memory effects were critically dependent on hippocampal integrity.

Consistent with previous work (e.g., Althoff and Cohen, 1999; Ryan et al., 2000), Smith and Squire (2008) reported that neurologically intact participants sampled fewer regions and made fewer fixations to repeated (vs. novel) images. However, these effects of repetition on eye movement behavior were not statistically reliable when amnesic patients were tested, a finding that contradicts past work (Althoff et al., 1993; Althoff, 1998; Ryan et al., 2000). There are important methodological differences between the experiments, including issues concerning design and measurement that bear importantly on the ability to elicit and detect the eye movement effects of interest. Such concerns emphasize how critical it is that careful consideration be given to the manner in which eye movement measures are defined and calculated (see Box 1).

Methodological issues aside, claims that eye movement based memory effects require conscious recollection (Smith and Squire, 2008) are at odds not only with studies from our group (Althoff et al., 1993; Althoff, 1998; Ryan et al., 2000), but also with several other eye movement investigations including those by Hayhoe et al. (1997), Hollingworth and colleagues (Hollingworth et al., 2001, 2008; Hollingworth and Henderson, 2002; Henderson and Hollingworth, 2003), and Beck et al. (2007; see Revealing Memory in the Absence of Awareness). When these results are considered together with more recent work (e.g., Hannula and Ranganath, 2009), the findings provide a powerful demonstration that the crucial mnemonic contribution of the hippocampus is relational and distinct from any potential role in conscious awareness. To date, there have been no demonstrations of *intact* eye movement based relational memory effects in the face of hippocampal damage. Moreover, our conclusion about the central role of relational memory in any account of hippocampal function receives overwhelming support from other lines of cognitive neuroscience research (for review see Davachi, 2006; Konkel and Cohen, 2009).

## REVEALING SPARED AND IMPAIRED MEMORY IN OTHER SPECIAL POPULATIONS

Cognitive tasks developed for use with college-age participants may not always lend themselves easily to translation for work with special populations. Traditional behavioral measures can be confounded by issues of task comprehension, complex decision making requirements, and behavioral response mapping difficulties in these individuals (e.g., Luck and Gold, 2008). Accordingly, sensitive, indirect methods, especially those that have the potential to be translated for use with animals or for use with neuroimaging methods, aid in the development of new investigations of cognition with special populations. A notable strength of eye movement methods is that they can be used to assess memory with or without concomitant collection of behavioral responses. As such, eye movement methods are particularly useful as a memory assessment tool for animals and infants (see **Boxes 3 and 4,** and **Figure 5**), and for investigations of the memory impairments associated with normal aging and psychiatric disorders (e.g., schizophrenia). Better characterization of memory performance may ultimately permit remediation when performance is impaired (see **Box 5**), and may contribute to identification of neural substrates that can be targeted for pharmaceutical intervention (see Carter and Barch, 2007).

The promise of *indirect* eye movement methods for providing insight into the nature of reported memory deficits associated with normal aging (see Light, 1996 for a review) has been realized in recent work (Firestone et al., 2007; Ryan et al., 2007c; Heisz and Ryan, submitted). Older individuals expect to perform poorly if they know their memory is being tested, thereby unwittingly undermining their performance (e.g., Rahhal et al., 2001; Chasteen et al., 2005). As eye movement methods do not require participants to explicitly comment on the contents of their memory (i.e., in free-viewing paradigms), they can be used to eliminate confounding effects of anticipated failure. Here, we highlight one such investigation (Ryan et al., 2007c), which shows the power of eye movements to test the predictions of more than one theoretical account simultaneously, very much like the approach taken in investigations of amnesia.

Several competing theories have been proposed to explain memory decline associated with normal aging (see Light, 1996; Balota et al., 2000 for reviews). Two influential theories attribute age-related memory deficits to impaired inhibition of irrelevant information (e.g., Zacks et al., 2000) and impaired binding of relational representations (Naveh-Benjamin, 2000). In an eye movement experiment designed to test predictions consistent with these theoretical perspectives, we used an experimental manipulation in which deficits in inhibition should benefit relational memory binding: objects that were to be ignored within scenes ultimately underwent a change in their spatial relations (Ryan et al., 2007c). Older adults displayed a concurrent deficit in both inhibition and binding. Relative to younger adults, they were more likely to fixate abrupt onsets (i.e., the objects that were subsequently manipulated). Despite this lack of inhibition, and in contrast to younger adults, older adults did not look disproportionately at manipulated objects. Additional work from our group has shown that age-related deficits in binding can be ameliorated when older adults can appeal to information that is already in memory (Firestone et al., 2007; Heisz and Ryan, submitted). Together, these findings suggest an age-related deficit in relational binding, akin to amnesic patients, which is consistent with the finding that aging is associated with disproportionate atrophy in the hippocampus (e.g., Driscoll et al., 2003; Erickson et al., 2010).

Recent work suggests that eye movements may also have some diagnostic value in the early identification of mild cognitive impairment (MCI; Crutcher et al., 2009) and may provide some insight into the nature of the deficit in Alzheimer's disease (Daffner et al., 1992, 1999). For instance, Crutcher et al. (2009) have shown that, unlike matched controls, patients with MCI fail to look preferentially at a novel picture over one presented 2 min earlier in the context of the visual paired comparison task. Instead, patients dis-

**BOX 3 | Eye movement monitoring of non-human primates.**

Non-human primates provide a means to pose scientific questions that cannot be addressed ethically in human participants, and both the lesion method and single-cell recording have been fruitfully combined with eye movement methods in this population. In some cases eye movement data have settled longstanding questions within the memory literature, and a notable example is the performance on certain memory tasks by animals assumed to model anterograde amnesia.

The delayed-non-match-to-sample (DMNS) task was a mainstay of animal research into memory systems for decades, but the results from lesion studies of the task were difficult to interpret. For example, lesions of the hippocampus did not impair DNMS performance, while lesions of the perirhinal cortex did, even at fairly short delays (see Suzuki and Amaral, 2004). This suggested that the perirhinal cortex must support knowledge about single objects across delays, but that functionality could reasonably have been assigned to late visual areas. By combining lesions with eye-tracking, Buffalo et al. (1999) demonstrated a dissociation between the contributions of late visual areas and perirhinal cortex in a preferential viewing task. Lesions of late visual areas produced at-chance viewing, while lesions of perirhinal cortex produced impaired-but-above-chance performance at most delays. This showed that the DNMS and visual preferential viewing tasks were differentially sensitive to MTL damage.

The study of non-human primates also allows for observation of the normal operation of MTL systems using single-unit recording in conjunction with eye movement monitoring. Jutras and Buffalo (2010) recently investigated the firing of hippocampal neurons in just this fashion. Using a set of thousands of images novel to the monkeys, they employed the visual paired comparison task (also termed the visual preferential looking task) and observed single units in hippocampus that were reliably modulated by the novelty or familiarity of a stimulus. Further, firing rates of hippocampal neurons sensitive to novelty were reliably correlated with an eye movement index of memory strength, indicating that the hippocampus plays a direct and crucial role in recognition memory.

The finding that hippocampal cell firing rates correspond to eye movement based measures of memory prompts another question, i.e., whether saccadic eye movements directly affect brain systems crucial for memory. Given that the visual experience of many organisms is frequently and unconsciously interrupted by abrupt changes in our point of gaze, it is reasonable to hypothesize that the memory systems that evolved alongside these visual systems might make use of saccades as boundaries between memoranda. Single-unit recording studies indicate that the activity of MTL structures is influenced by eye movements in intriguing ways: transmission of electrical impulses within the MTL is uniquely enhanced in the 100 ms following a saccade, and saccades elicit activity in those same structures (Ringo et al., 1994; Sobotka et al., 1997, 2002). Critically, these effects are extraretinal, all having been observed under both light and dark conditions. These findings indicate that individual eye movements are constantly modulating the activity of the same MTL structures known to be necessary for normal declarative (relational) memory.

tributed viewing almost equally between the two pictures. Because novelty preferences were intact when the delay between the familiarization phase and the test phase was shorter (i.e., 2 s), between-group differences in patterns of viewing were unlikely to have been a consequence of perceptual, motivational, or attentional deficits. The implication is that eye movement based memory assessment may be of use in the identification and characterization of dementia (Crutcher et al., 2009), but because this research is just getting underway, future investigations are needed to confirm the potential utility of eye movement methods in this arena.

Studies conducted with psychiatric populations (e.g., schizophrenia patients) may also benefit from the use of *indirect* eye movement methods. Schizophrenia is a psychiatric disease that affects performance on several cognitive and information-processing tasks, but disproportionate impairments have been observed on tests of episodic memory (e.g., McKenna et al., 1990; Saykin et al., 1991), and memory is a strong predictor of functional outcome in these individuals (Green et al., 2000). Recent results suggest that patients with schizophrenia may have relatively spared memory for items and disproportionately impaired memory for item-context relationships, although others maintain that the deficit is generalized (see Ranganath et al., 2008). Eye movement based memory experiments, like those described in the preceding sections, may help resolve this debate. In contrast to behavioral measures, eye movement data can be acquired after providing participants with relatively simple instructions, and need not invoke behavioral responses with complex decision making requirements; therefore, this technique may be particularly suitable for investigating the nature of the memory impairment in schizophrenia.

Recently, results from two independent investigations have shown that relational memory effects are considerably reduced in the eye movement behavior of schizophrenic patients (Hannula et al., 2010; Williams et al., 2010), while other measures of viewing related to specific task demands (e.g., viewing a face that will be selected) are similar to those of controls. Results like these are tantalizing, and help to validate use of eye movement monitoring as a tool for studying memory in these individuals. They also suggest that eye movement measures may be useful in future attempts to relate impaired cognitive performance to underlying neural mechanisms.

## INTEGRATION OF EYE MOVEMENT MONITORING WITH NEUROIMAGING

The preceding sections have provided examples of the ability of eye movement methods to address questions about memory that were not necessarily possible with more traditional behavioral methods. However, such methods are particularly powerful for advancing the cognitive neuroscience of memory when combined with other methods and approaches. Eye movement data acquired in a scanning environment is clearly useful for observing participants' compliance if the task requires fixation on a central location, but more nuanced data can also be used to categorize trials according to some variable of interest (e.g., viewing time directed to particular regions of a display). The use of eye movement monitoring in conjunction with neuroimaging techniques can begin to dissociate different stages of stimulus processing (e.g., early obligatory effects of memory versus later explicit retrieval), and to determine whether these stages are supported by different underlying neural networks.

**BOX 4 | Eye movement monitoring of infants.**

Characterizing the development of human memory systems is necessary for a thorough understanding of how memory operates in adults. Representing the beginning of the development spectrum, infants are a challenging population to study owing to their lack of language, short attention span, and poor motor control. Eye movements are one index of infant behavior that can be compared to adults, and with an appropriate testing apparatus, short tasks, and many participants, eye movement data can be gathered from even very young infants. Importantly, early investigations established that infants can discriminate relatively complex visual stimuli, paving the way for sophisticated investigations of infant memory that followed (see **Figure 5**). Fantz (1964) repeatedly exposed the same stimulus to infants between 1 and 6 months old, observing that the stimulus was fixated less and less frequently with increasing repetition. Extending from this, Fagan (1970) simultaneously presented two stimuli with different viewing histories, one novel and one familiar, and found that infants as young as 2–6 months showed preferential viewing for novel items. Further evidence for memory in infants emerged fortuitously; Fagan re-used stimuli on three consecutive days and on the second and third days of testing, there was no preferential viewing of stimuli seen the day before. This demonstrated that even before 6 months of age infants could form memories of visual stimuli that persisted for days. Preferential viewing indicates that humans have at least basic mnemonic functions soon after birth. However, researchers have often argued that there are aspects of adult memory that may not be present early in life. For instance, findings from behavioral studies have suggested that the flexible, relational, and hippocampally dependent memories that typify declarative memory may not be available to young children, perhaps due to a medial temporal lobe system that has not been fully developed (e.g., Sluzenski et al., 2006). However, a fascinating replication of an experiment first used to investigate relational memory in healthy and amnesic adults calls these ideas into question. Hannula et al. (2007) used a passive viewing paradigm to test for knowledge of previously established face-scene relations expressed in eye movements (see **Figure 4**). Richmond and Nelson (2009) adapted the same experiment to the testing of infants, and observed patterns of viewing that demonstrated memory for face-scene pairings in infants with the same time-course reported in work with adults. This startling result indicates that relational memories are being formed as early as 9 months of age, a major departure from existing hypotheses.

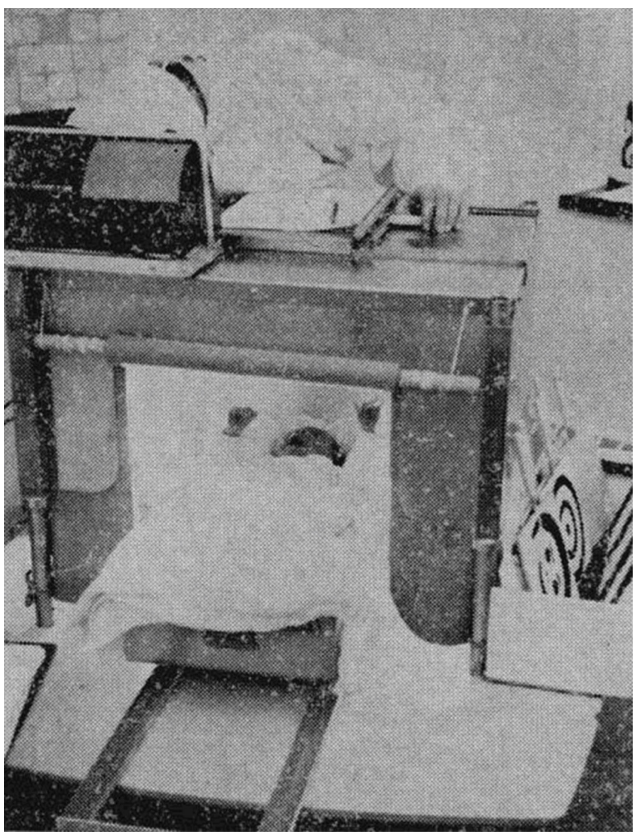


**FIGURE 5 | Illustration of early methods used to examine eye movement behavior in infants.** This infant "looking chamber" was used by Fantz (1963) to examine the length of gaze directed to visual targets. Fixation duration was recorded by the experimenter who monitored the infant's gaze through a small hole in the ceiling of the chamber. Current methods are akin to those used with adult participants (cf. Richmond and Nelson, 2009)

A compelling example of the integration of fMRI with eye movement methods was discussed briefly above (Hannula and Ranganath, 2009; see Revealing a Critical Role for the Hippocampus in Memory Without Awareness). Adapting our previous paradigm in which eye movement effects of relational memory were observed in healthy participants but not in patients with hippocampal damage (Hannula et al., 2007; see **Figure 4**), the use of converging methods revealed that hippocampal activity during presentation of the scene cue predicted viewing of the matching face, even when participants failed to identify that face correctly via behavioral response. Further, when participants made a correct behavioral response, activity in lateral prefrontal areas, and functional connectivity between these areas and the hippocampus, was increased, suggesting that while the hippocampus may be the critical area for retrieval of the relational information, additional recruitment of extrahippocampal brain regions may be required for its explicit expression (Hannula and Ranganath, 2009).

The concurrent collection of eye movements and magnetoencephalography (MEG) is also expected to provide insight into mnemonic processes. A non-invasive neuroimaging technique that estimates neuronal activity based on recordings of the magnetic flux outside of the head (Hämäläinen et al., 1993; Hari et al., 2000), MEG allows recording of neural activity with temporal resolution on the order of milliseconds and spatial resolution comparable to that of fMRI (Miller et al., 2007), making it an ideal tool for studying the dynamics of brain function.

Exploration of memory processes with combined MEG and eye movement techniques has just begun, but work has already shown that MEG can localize signals from the hippocampus (Riggs et al., 2009; see also Breier et al., 1998, 1999; 2000; Tesche and Karhu, 2000; Hanlon et al., 2003, 2005; Gonsalves et al., 2005; Moses et al., 2009), and that hippocampal responses are evident as early as 120–130 ms after stimulus onset during a recognition task (Riggs et al., 2009).

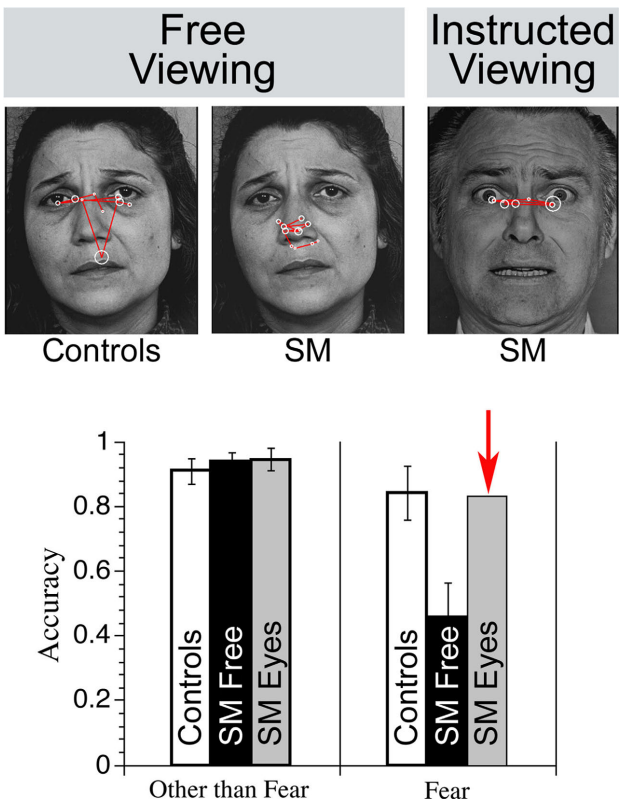
**BOX 5 | Eye movements as a diagnostic tool and key to remediation following brain damage or dysfunction.**

Recognition of facial identity and facial expression of emotion are fundamental to human interaction, but these capacities can be disrupted by damage to or dysfunction of specific brain regions. Measurement of eye movements can both enhance understanding of these disruptions and suggest strategies for remediation and rehabilitation.

Prosopagnosia refers to a severe deficit in recognizing familiar faces either subsequent to brain damage (acquired prosopagnosia, AP) or present since birth (congenital prosopagnosia, CP). Eye movement monitoring has revealed that individuals with prosopagnosia, in particular CP, show abnormal scanning patterns of faces: fewer fixations to central features like the eyes, nose, and mouth; enhanced viewing of the mouth over the eyes; and more fixations to peripheral features including hair and hairline (Le et al., 2003; Schwarzer et al., 2007; de Xivry et al., 2008; Schmalzl et al., 2008; Stephan and Caine, 2009; but see Bate et al., 2008). If abnormal scanning patterns underlie CP, then it stands to reason that rehabilitative techniques that target eye movements may be beneficial for patients with prosopagnosia. In independent efforts (De Gutis et al. 2007; Schmalzl et al., 2008), two patients with well-characterized CP received specific instructions regarding attention to internal features of faces. Both patients benefitted from this remedial technique, and were better able to recognize faces afterward. Further, one of these investigations (De Gutis et al., 2007) used neuroimaging techniques to determine whether this new ability was reflected in neural activity. In that patient, an N170 ERP response to faces was evident only after training, and increased functional connectivity was observed between face-selective regions of the brain. These neural changes presumably mirrored and supported the new abilities that the patient developed following conscious changes to their eye movement behavior.

Problems with recognition of facial expression of emotion can also benefit from the study of eye movements. Research conducted with an individual who has nearly complete bilateral amygdala damage due to Urbach–Weithe disease has provided new insight into the mechanism that underlies her inability to identify/recognize fearful faces (Adolphs et al., 2005). In contrast to intact recognition of other facial expressions, S.M. showed severely impaired recognition of faces that bear fearful expressions (Adolphs et al. 1994). Perhaps surprisingly, S.M. could recognize fearful prosody (Adolphs and Tranel, 1999) and could describe fearful situations (Adolphs et al. 1995). Eye movement monitoring has been used to examine this patient's visual recognition problem in detail; in contrast to controls, S.M. spent very little time looking spontaneously at the eye region of faces under free-viewing conditions (Adolphs et al., 2005). Having identified this abnormal viewing pattern, investigators instructed S.M. specifically to look at the eyes of the faces she was shown (i.e., instructed viewing). Under these circumstances her ability to identify fearful facial expressions improved and equaled that of healthy comparisons (see associated Figure).

Jointly, these promising results from eye movement monitoring of two patient populations suggest that remediation of deficient recognition can sometimes be accomplished with a simple, conscious behavioral modification.




Critically, recent work has also shown that MEG can be successfully combined with eye movement monitoring (Herdman and Ryan, 2007; see also Hirvenkari et al., 2010). In this work, MEG data were time-locked to eye movements, permitting examination of neural activity immediately preceding and following saccades. The next step in the combined use of eye movement monitoring and MEG is to reveal the neural networks that drive eye movement based memory effects.

## CONCLUDING STATEMENTS

In this review, we have attempted to illustrate the utility and promise of eye movement methods for advancing cognitive neuroscience investigations, with a focus on investigations of memory. Eye movements have been shown to reveal the influence of different types of memory (e.g., item memory and relational memory) using a variety of materials (e.g., faces, building, scenes, arbitrary pairings and sequences), permitting us to address questions about distinct memory systems. The influence of memory on eye movements can be observed obligatorily, soon after stimulus onset, and may occur long before, or in the absence of, conscious awareness for remembered content. Given that eye movements can be acquired without explicit reports or other overt responses, this method is an ideal tool for indexing memory function in special populations (i.e., non-human primates, infants, and patients with neuropsychological, psychiatric or neurodegenerative conditions). We have exploited this technique to observe changes in memory function that are associated with healthy aging and amnesia, and in so doing have reinforced the critical role of the hippocampal system in memory. Beyond memory, findings from eye movement investigations have influenced, or promise to influence, our

understanding of disordered cognition exhibited by patients who have prosopagnosia, schizophrenia, and dementia. Indeed, when eye movement methods are combined with patient populations, there is even promising translational potential. This is exemplified by work conducted by Adolphs et al. (2005; see **Box 5**), in which observation of abnormal viewing patterns prompted remedial strategies that restored recognition of fearful faces in the laboratory.

Studies that combine eye movements with neuroimaging techniques have the potential to provide unparalleled insights into the brain networks that support various memory abilities and answer questions about the role of conscious awareness in the use of memory. Eye movement methods also hold great promise for relating the activity of particular brain regions and systems to the time at which the various influences of memory emerge. Given that eye movement findings reveal the early and obligatory influences of memory in online processing (Hannula et al., 2007; Ryan et al., 2007a; Warren et al., 2010), such findings challenge our traditional notions of "perception" and "memory" and suggest that the very nature of perceptual processing is altered as a result of our prior experiences (Hannula et al., 2007; Ryan et al., 2007, 2008). Furthermore, eye movement methods provide a powerful tool for revealing the influences of memory across cognitive domains (e.g., in language: Rubin et al., 2009), and may reveal the influence of other cognitive processes on memory itself (e.g., emotional valence: Riggs et al., inpress).

Cognitive neuroscience as a discipline has benefited greatly from the converging methods approach, using multiple methods to provide comprehensive answers to difficult questions. This review was intended to illustrate the promise of eye movement monitoring as one of the methods that should be considered by cognitive neuroscientists, and to demonstrate the advances that have been made in the cognitive neuroscience of memory as a result of combining eye movement methods with neuropsychological and neuroimaging approaches.

## ACKNOWLEDGMENTS

We wish to thank Douglas McQuiggan for his assistance with the manuscript,Ralph Adolphs for providing us with a high resolution version of the Figure in Box 5, and Tanya Kirouac for improving the quality of the drawing in Figure 1. Funding from NIMH grant MH062500 was provided to Neal J. Cohen. Funding from the Natural Sciences and Engineering Research Council of Canada (NSERC), Canadian Institutes of Health Research (CIHR), Canada Research Chairs Program and Canadian Foundation for Innovation (CRC/CFI) was provided to Jennifer D. Ryan.

## REFERENCES

Adolphs, R., Gosselin, F., Buchanan, T. W., Tranel, D., Schyns, P., and Damasio, A. R. (2005). A mechanism for impaired fear recognition after amygdala damage. *Nature* 433, 68–72 (Figure reprinted by permission from Macmillan Publishers Ltd: Nature).

Adolphs, R., and Tranel, D. (1999). Intact recognition of emotional prosody following amygdale damage. *Neuropsychologia* 37, 1285–1292.

Adolphs, R., Tranel, D., Damasio, H., and Damasio, A. (1994). Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. *Nature* 372, 669–672.

Adolphs, R., Trauel, D., Damasio, H., and Damasio, A. (1995). Fear and the human amygdala. *J. Neurosci.* 15, 5879–5892.

Althoff, R. R. (1998). *Eye-Movement-Based Memory Assessment: The Use of Eye Movement Monitoring as an Indirect Measure of Memory*. Unpublished doctoral dissertation, University of Illinois at Urbana-Champaign, Urbana.

Althoff, R. R., and Cohen, N. J. (1999). Eye-movement based memory effect: a reprocessing effect in face perception. *J. Exp. Psychol.* 25, 997–1010.

Althoff, R. R., Cohen, N. J., McConkie, G., Wasserman, S., Maciukenas, M., Azen, R., and Romine, L. (1998). "Eye-movement based memory assessment," in *Current oculomotor research: Physiological and Psychological Aspects*, eds W. Becker, H. Deubel, and T. Mergner (New York: Kluwer Academic/Plenum Press), 293–302.

Althoff, R. R., Maciukenas, M., and Cohen, N. J. (1993). Indirect assessment of memory using eye movement monitoring. *Abstr. - Soc. Neurosci.* 19, 439.

Antes, J. R. (1974). The time course of picture viewing. *J. Exp. Psychol.* 103, 62–70.

Balota, D. A., Dolan, P. O., and Duchek, J. (2000). "Memory changes in healthy older adults," in *The Oxford Handbook of Memory*, eds E. Tulving and F. I. M. Craik (New York: Oxford University Press), 395–409.

Bate, S., Haslam, C., Tree, J. J., and Hodgson, T. L. (2008). Evidence of an eye-movement based memory effect in congenital prosopagnosia. *Cortex* 44, 806–819.

Beck, M. R., Peterson, M. S., and Angelone, B. L. (2007). The roles of encoding, retrieval, and awareness in change detection. *Mem. Cognit.* 35, 610–620.

Belopolsky, A. V., Kramer, A. F., and Theeuwes, J. (2008). The role of awareness in processing of oculomotor capture: evidence from event-related potentials. *J. Cogn. Neurosci.* 20, 2285–2297.

Breier, J. I., Simos, P. G., Zouridakis, G., and Papanicolaou, A. C. (1998). Relative timing of neuronal activity in distinct temporal lobe areas during a recognition memory task for words. *J. Clin. Exp. Neuropsychol.* 20, 782–790.

Breier, J. I., Simos, P. G., Zouridakis, G., and Papanicolaou, A. C. (1999). Lateralization of cerebral activation in auditory verbal and non-verbal memory tasks using magnetoencephalography. *Brain Topogr.* 12, 89–97.

Breier, J. I., Simos, P. G., Zouridakis, G., and Papanicolaou, A. C. (2000). Lateralization of activity associated with language function using magnetoencephalography. *J. Clin. Neurophysiol.* 17, 503–510.

Brillouin, L. (1962). *Science and Information Theory*, 2nd Edn. New York: Academic Press, Inc.

Brockmole, J. R., Castelhano, M. S., and Henderson, J. M. (2006). Contextual cueing in naturalistic scenes: global and local contexts. *J. Exp. Psychol. Learn. Mem. Cogn.* 32, 699–706.

Brockmole, J. R., and Henderson, J. M. (2008). Prioritizing new objects for eye fixation in real-world scenes: effects of object-scene consistency. *Vis. Cogn.* 16, 375–390.

Brockmole, J. R., and Irwin, D. E. (2005). Eye movements and the integration of visual memory and visual perception. *Percept. Psychophys.* 67, 495–512.

Brockmole, J. R., Irwin, D. E., and Wang, R. F. (2003). The locus of spatial attention during the temporal integration of visual memories and visual percepts. *Psychon. Bull. Rev.* 10, 510–515.

Brockmole, J. R., Wang, R. F., and Irwin, D. E. (2002). Temporal integration between visual images and visual percepts. *J. Exp. Psychol. Hum. Percept. Perform.* 28, 315–334.

Buffalo, E. A., Ramus, S. J., Clark, R. E., Teng, E., Squire, L. R., and Zola, S. M. (1999). Dissociation between the effects of damage to perirhinal cortex and area TE. *Learn. Mem.* 6, 572–599.

Buswell, G. T. (1935). *How People Look at Pictures: A study of the Psychology of Perception in Art*. Chicago, IL: University of Chicago Press.

Carpenter, P. A., and Just, M. A. (1977). "Reading comprehension as eyes see it," in *Cognitive processes in comprehension*, eds M. A. Just and P. A. Carpenter (Hillsdale, NJ: Erlbaum), 109–139.

Carter, C. S., and Barch, D. M. (2007). Cognitive neuroscience-based approaches to measuring and improving treatment effects on cognition in schizophrenia: the CNTRICS initiative. *Schizophr. Bull.* 33, 1131–1137.

Castelhano, M. S., and Henderson, J. M. (2007). Initial scene representations facilitate eye movement guidance in visual search. *J. Exp. Psychol. Hum. Percept. Perform.* 33, 753–763.

Chasteen, A. L., Bhattacharyya, S., Horhota, M., Tam, R., and Hasher, L. (2005). How feelings of stereotype threat influence older adults' memory performance. *Exp. Aging Res.* 31, 235–260.

Chun, M. M., and Nakayama, K. (2000). On the functional role of implicit visual memory for the adaptive deployment of attention across scenes. *Vis Cogn.* Special Issue: Change Blindness and Visual Memory, 7, 65–81.

Cohen, N. J. (1984). "Preserved learning capacity in amnesia: evidence for multiple memory systems," in *Neuropsychology of Memory*, eds L. R. Squire and N. Butters (New York: Guilford), 83–103.

Cohen, N. J., and Eichenbaum, H. (1993). *Memory, Amnesia, and the Hippocampal System*. Cambridge, MA: The MIT Press.

Cohen, N. J., and Squire, L. R. (1980). Preserved learning and retention of pattern-analyzing skill in amnesia: Dissociation of knowing how and knowing that. *Science* 210, 207–210.

Crutcher, M. D., Calhoun-Haney, R., Manzanares, C. M., Lah, J. J., Levey, A. I., and Zola, S. M. (2009). Eye tracking during a visual paired comparison task as a predictor of early dementia. *Am. J. Alzheimers Dis. Other Demen.* 24, 258–266.

Daffner, K. R., Mesulam., M. M., Cohen, L. G., and Scinto, L. F. M. (1999). Mechanisms underlying diminished novelty-seeking behavior in patients with probably Alzheimer's disease. *Neuropsychiatry Neuropsychol. Behav. Neurol.* 12, 58–66.

Daffner, K. R., Scinto, L. F. M., Weintraub, S., Guinessey, J., and Mesulam, M. M. (1992). diminished curiosity in patients with probable Alzheimer's disease as measured by exploratory eye movements. *Neurology* 42, 320–328.

Davachi, L. (2006). Item, context and relational episodic encoding in humans. *Curr. Opin. Neurobiol.* 16, 693–700.

De Gutis, J. M., Bentin, S., Robertson, L. C., and D'Esposito, M. (2007). Functional plasticity in ventral temporal cortex following cognitive rehabilitation of a congenital prosopagnosic. *J. Cogn. Neurosci.* 19, 1790–1802.

de Xivry, J. O., Ramon, M., Lefèvre, P., and Rossion, B. (2008). Reduced fixation on the upper area of personally familiar faces following acquired prosopagnosia. *J. Neuropsychol.* 2, 245–268.

Didday, R. L., and Arbib, M. A. (1975). Eye movements and visual perception: A "two visual system" model. *Int. J. Man. Mach. Stud.* 7, 547–569.

Driscoll, I., Hamilton, D. A., Petropoulos, H., Yeo, R. A., Brooks, W. M., Baumgartner, R. N., and Sutherland, R. J. (2003). The aging hippocampus: Cognitive, biochemical and structural findings. *Cereb. Cortex* 13, 1344–1351.

Eichenbaum, H., and Cohen, N. J. (2001). *From Conditioning to Conscious Recollection*. New York: Oxford University Press.

Ellis, S. R., and Stark, L. (1986). Statistical dependency in visual scanning. *Hum. Factors* 28, 421–438.

Erickson, K. I., Prakash, R. S., Voss, M. W., Chaddock, L., Heo, S., McLaren, M., Pence, B. D., Martin, S. A., Vieira, V. J., McAuley, E., and Kramer, A. F. (2010). Brain-derived neurotrophic factor is associated with age-related decline in hippocampal volume. *J. Neurosci.* 30, 5368–5375.

Fagan, J. F. (1970). Memory in the infant. *J. Exp. Child. Psychol.* 9, 217–226.

Fantz, R. L. (1963). Pattern vision in newborn infants. *Science* 140, 296–297 (Figure reprinted with permission from AAAS).

Fantz, R. L. (1964). Visual experience in infants: Decreased attention to familiar patterns relative to novel ones. *Science* 146, 668–670.

Firestone, A., Turk-Browne, N., and Ryan, J. D. (2007). Age-related deficits in face recognition are related to underlying changes in scanning behavior. *Neuropsychol. Dev. Cogn. B Aging Neuropsychol. Cogn.* 14, 594–607.

Fisher, D. F., Monty, R. A., and Perlmuter, L. C. (1978). Visual recognition memory for binary pictures: another look. *J. Exp. Psychol. Hum. Learn.* 4, 158–164.

Gabrieli, J. D. E. (1998). Cognitive neuroscience of human memory. *Annu. Rev. Psychol.* 49, 87–115.

Graf, P., and Schacter, D. L. (1985). Implicit and explicit memory for new associations in normal and amnesic patients. *J. Exp. Psychol. Hum. Learn.* 11, 501–518.

Gonsalves, B. D., Kahn, I., Curran, T., Norman, K. A., and Wagner, A. D. (2005). Memory strength and repetition suppression: multimodal imaging of medial temporal cortical contributions to recognition. *Neuron* 47, 751–761.

Green, M. F., Kern, R. S., Braff, D. L., and Mintz, J. (2000). Neurocognitive deficits and functional outcome in schizophrenia: Are we measuring the "right stuff"? *Schizophr. Bull.* 26, 119–136.

Hämäläinen, M., Hari, R., Ilmoniemi, R. J., Knuutila, J., and Lounasmaa, O. V. (1993). Magnetoencephalography – theory, instrumentation, and applications to noninvasive studies of the working human brain. *Rev. Modern Phys.* 65, 413–496.

Hanlon, F. M., Weisend, M. P., Huang, M., Lee, R. R., Moses, S. N., Paulson, K. M., Thoma, R. J., Miller, G. A., and Cañive, J. M. (2003). A non-invasive method for observing hippocampal function. *Neuroreport* 14, 1957–1960.

Hanlon, F. M., Weisend, M. P., Yeo, R. A., Huang, M., Lee, R. R., Thoma, R. J., Moses, S. N., Paulson, K. M., Miller, G. A., and Cañive, J. M. (2005). A specific test of hippocampal deficit in schizophrenia. *Behav. Neurosci.* 119, 863–875.

Hannula, D. E., and Ranganath, C. (2009). The eyes have it: hippocampal activity predicts expression of memory in eye movements. *Neuron* 63, 592–599.

Hannula, D. E., Ranganath, C., Ramsay, I. S., Solomon, M., Yoon, J., Niendam, T. A., Carter, C. S., and Ragland, J. D. (2010). Use of eye movement monitoring to examine item and relational memory in schizophrenia. *Biol. Psychiatry*. doi:10.1016/j. biopsych.2010.06-001. [Epub ahead of print].

Hannula, D. E., Ryan, J. D., Tranel, D., and Cohen, N. J. (2007). Rapid onset relational memory effects are evident in eye movement behavior, but not in hippocampal amnesia. *J. Cogn. Neurosci.* 19, 1690–1705.

Hari, R., Levänen, S., and Raij, T. (2000). Timing of human cortical functions during cognition: role of MEG. *Trends Cogn. Sci.* 4, 455–462.

Hayhoe, M. M., Bensinger, D. G., and Ballard, D. H. (1997). Task constraints in visual working memory. *Vision Res.* 38, 125–137.

Heisz, J. J., and Shore, D. I. (2008). More efficient scanning for familiar faces. *J. Vis.* 8, 1–10.

Henderson, J. M. (2003). Human gaze control during real-world scene perception. *Trends Cogn. Sci.* 7, 498–504.

Henderson, J. M., and Hollingworth, A. (2003). Eye movements and visual memory: detecting changes to saccade targets in scenes. *Percept. Psychophys.* 65, 58–71.

Henderson, J. M., Weeks, P. A., and Hollingworth, A. (1999). The effects of semantic consistency on eye movements during complex scene viewing. *J. Exp. Psychol. Hum. Percept. Perform.* 25, 210–228.

Herdman, A. T., and Ryan, J. D. (2007). Spatiotemporal brain dynamics underlying saccade execution, suppression and error-related feedback. *J. Cogn. Neurosci.* 19, 420–432.

Hirvenkari, L. Jousmäki, V., Lamminmäki, S., Saarinen, V., Sams, M. E., and Hari, R. (2010). Gaze-direction-based MEG averaging during audiovisual speech perception. *Front. Hum. Neurosci.* 4:17. doi: 10.3389/fnhum.2010.00017.

Hollingworth, A. (2009). Two forms of scene memory guide visual search: Memory for scene context and memory for the binding of target object to scene location. *Vis. Cogn.* 17, 273–291.

Hollingworth, A., and Henderson, J. M. (2002). Accurate visual memory for previously attended objects in natural scenes. *J. Exp. Psychol. Hum. Percept. Perform.* 28, 113–136.

Hollingworth, A., Richard, A. M., and Luck, S. J. (2008). Understanding the function of visual short-term memory: Transsaccadic memory, object correspondence, and gaze correction. *J. Exp. Psychol. Gen.* 137, 163–181.

Hollingworth, A., Williams, C. C., and Henderson, J. M. (2001). To see and remember: Visually specific information is retained in memory from previously attended objects in natural scenes. *Psychon. Bull. Rev.* 8, 761–768.

Holm, L., Eriksson, J., and Andersson, L. (2008). Looking as if you know: systematic object inspection precedes object recognition. *J. Vis.* 8, 1–7.

Jutras, M. J., and Buffalo, E. A. (2010). Recognition memory signals in the macaque hippocampus. *Proc. Natl. Acad. Sci. U.S.A.* 107, 401–406.

Konkel, A., and Cohen, N. J. (2009). Relational memory and the hippocampus: representation and methods. *Front. Neurosci.* 3:23. doi: 10.3389/neuro.01.023.2009.

Kramer, A. F., Hahn, S., Irwin, D. E., and Theeuwes, J. (2000). Age differences in the control of looking behavior: do you know where your eyes have been? *Psychol. Sci.* 11, 210–217.

Kumaran, D., and Wagner, A. D. (2009). It's in my eyes, but it doesn't look that way to me. *Neuron* 63, 561–563.

Le, S., Raufaste, E., and Demonet, K. (2003). Processing of normal, inverted, and scrambled faces in a patient with prosopagnosia: behavioural and eye tracking data. *Cogn. Brain Res.* 17, 26–35.

Light, L. L. (1996). "Memory and aging," in *Memory*, eds E. L. Bjork and R. A. Bjork (San Diego, CA: Academic Press), 443–490.

Loftus, G. R., and Mackworth, N. H. (1978). Cognitive determinants of fixation location during picture viewing. *J. Exp. Psychol. Hum. Percept. Perform.* 4, 565–572.

Luck, S. J., and Gold, J. M. (2008). The translation of cognitive paradigms for patient research. *Schizophr. Bull.* 34, 629–644.

Mackworth, N. H., and Morandi, A. J. (1967). The gaze selects informative details within pictures. *Percept. Psychophys.* 2, 547–552.

Maljkovic, V., and Nakayama, K. (2000). Priming of popout: III. A short-term implicit memory system beneficial for rapid target selection. *Vis. Cogn.* 7, 571–595.

Manns, J. R., and Squire, L. R. (2001). Perceptual learning, awareness, and the hippocampus. *Hippocampus* 11, 776–782.

McKenna, P. J., Tamlyn, D., Lund, C. E., and Mortimer, A. M. (1990). Amnesic syndrome in schizophrenia. *Psychol. Med. A Journal of Research in Psychiatry and the Allied Sciences,* 20, 967–972.

Miller, G. A., Elbert, T., Sutton, B. P., and Heller, W. (2007). Innovative clinical assessment technologies: Challenges

and opportunities in neuroimaging. *Psychol. Assess.* 19, 58–73.

Milner, B. (1962). "Les troubles de la mémoire accompagnant des lesions hippocampiques bilatérales," in *Physiologie de l'hippocampe*, ed. P. Passouant (Paris: Centre de la Recherche Scientifique), 257–262.

Milner, B., Corkin, S., and Teuber, H. L. (1968). Further analysis of the hippocampal amnesic syndrome: 14-year follow-up study of H.M. *Neuropsychologia* 6, 215–234.

Molnar, F., and Ratsikas, D. (1987). "Some aesthetical aspects of visual exploration," in *Eye Movements: From Physiology to Cognition*, eds J. K. O'Regan and A. Levy-Schoen (Amsterdam: North Holland), 363–374.

Moscovitch, M. (2008). The hippocampus as a "stupid" domain-specific module: Implications for theories of recent and remote memory, and of imagination. *Can. J. Exp. Psychol.* 62, 62–79.

Moses, S. N., Ryan, J. D., Bardouille, T., Kivacevic, N., Hanlon, F. M., and McIntosh, A. R. (2009). Semantic information alters neural activation during transverse patterning performance. *Neuroimage* 46, 863–873.

Naveh-Benjamin, M. (2000). Adult age differences in memory performance: Tests of an associative deficit hypothesis. *J. Exp. Psychol. Learn. Mem. Cogn.* 26, 1170–1187.

Noton, D., and Stark. L. (1971). Scanpaths in eye movements during pattern perception. *Science* 171, 308–311.

Parker, R. E. (1978). Picture processing during recognition. *J. Exp. Psychol. Hum. Percept. Perform.* 4, 284–293.

Rahhal, T. A., Hasher, L., and Colcombe, S. J. (2001). Instructional manipulations and age differences in memory: now you seem, now you don't. *Psychol. Aging* 16, 697–706.

Ranganath, C., Minzenberg, M. J., and Ragland, J. D. (2008). The cognitive neuroscience of memory function and dysfunction in schizophrenia. *Biol. Psychiatry* 64, 18–25.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *Q. J. Exp. Psychol.* 62, 1457–1506.

Richmond, J., and Nelson, C. A. (2009). Relational memory during infancy: Evidence from eye tracking. *Dev. Sci.* 12, 549–556.

Riggs, L., McQuiggan, D., Anderson, A. K., and Ryan, J. D. (in press). Eye movement monitoring reveals differential influences of emotion on memory. *Front. Emo. Sci.*

Riggs, L., Moses, S. N., Bardouille, T., Herdman, A. T., Ross, B., and Ryan, J. D. (2009). A converging analytic approach to examining medial temporal lobe sources using magnetoencephalography. *Neuroimage* 45, 627–642.

Ringo, J. L., Sobotka, S., Diltz, M. D., and Bunce, C. M. (1994). Eye movements modulate activity in hippocampal, parahippocampal, and inferotemporal neurons. *J. Neurophysiol.* 71, 1285–1288.

Rubin, R. D., Brown-Schmidt, S., Duff, M. C., Tranel, D., and Cohen, N. J. (2009). "Common ground representations in hippocampal amnesia," In *Society Neuroscience, 39th Annual Meeting*, Chicago, IL.

Ryan, J. D., Althoff, R. R., Whitlow, S., and Cohen, N. J. (2000). Amnesia is a deficit in relational memory. *Psychol. Sci.* 11, 454–461.

Ryan, J. D., and Cohen, N. J. (2003). Evaluating the neuropsychological dissociation evidence for multiple memory systems. *Cogn. Affect. Behav. Neurosci.* 3, 168–185.

Ryan, J. D., and Cohen, N. J. (2004a). The nature of change detection and on-line representations of scenes. *J. Exp. Psychol. Hum. Percept. Perform.* 30, 988–1015.

Ryan, J. D., and Cohen, N. J. (2004b). Processing and short-term retention of relational information in amnesia. *Neuropsychologia* 42, 497–511.

Ryan, J. D., Hannula, D. E., and Cohen, N. J. (2007a). The obligatory effects of memory on eye movements. *Memory* 15, 508–525.

Ryan, J. D., Herdman, A. T., Riggs, L., and Moses, S. N. (2007b). "Eye movement and spatiotemporal brain dynamics underlying processing of relations among objects," In *Cognitive Neuroscience Society Annual Meeting*, New York, NY.

Ryan, J. D., Leung, G., Turk-Browne, N. B., and Hasher, L. (2007c). Assessment of age-related changes in inhibition and binding using eye movement monitoring. *Psychol. Aging* 22, 239–250.

Ryan, J. D., Moses, S. N., Ostreicher, M., Bardouille, T., Herdman, A. T., Riggs, L., and Tulving, E. (2008). Seeing sounds and hearing sights: The influence of prior learning on current processing. *J. Cogn. Neurosci.* 20, 1030–1042.

Ryan, J. D., and Villate, C. (2009). Building visual representations: the binding of relative spatial relations across time. *Vis. Cogn.* 17, 254–272.

Saykin, A. J., Gur, R. C., Gur, R. E., Mozley, P. D., Mozley, L. H., Resnick, S. M., Kester, D. B., and Stafiniak, P. (1991). Neuropsychological function in schizophrenia: Selective impairment in memory and learning. *Arch. Gen. Psychiatry* 48, 618–624.

Schacter, D. L. (1987). Implicit memory: history and current status. *J. Exp. Psychol. Learn. Mem. Cogn.* 13, 501–508.

Schmalzl, L., Palermo, R., Green, M., Brunsdon, R., and Coltheart, M. (2008). Training of familiar face recognition and visual scan paths for faces in a child with congenital prosopagnosia. *Cogn. Neuropsychiatry* 25, 704–729.

Schwarzer, G., Huber, S., Grueter, M., Grueter, T., Grob, C., Hipfel, M., and Kennerknecht, I. (2007). Gaze behaviour in hereditary prosopagnosia. *Psychol. Res.* 71, 583–590.

Scoville, W. B., and Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *J. Neuropsychiatry Clin. Neurosci.* 20, 11–21.

Sluzenski, J., Newcombe, N., and Kovacs, S. (2006). Binding, relational memory, and recall of naturalistic events: a developmental perspective. *J. Exp. Psychol. Learn. Mem. Cogn.* 32, 89–100.

Smith, C. N., and Squire, L. R. (2008). Experience-dependent eye movements reflect hippocampus-dependent (aware) memory. *J. Neurosci.* 28, 12833–12825.

Snyder, K. A., Blank, M. P., and Marsolek, C. J. (2008). What form of memory underlies novelty preferences? *Psychon. Bull. Rev.*, 15, 315–321.

Sobotka, S., Nowicka, a., and Ringo, J. L. (1997). Activity linked to externally cued saccades in single units recorded from hippocampal, parahippocampal, and inferotemporal areas of macaques. *J. Neurophysiol.* 78, 2156–2163.

Sobotka, S., Zuo, W., and Ringo, J. L. (2002). Is the functional connectivity within temporal lobe influenced by saccadic eye movements? *J. Neurophysiol.* 88, 1675–1684.

Squire L. R. (1992). Memory and the hippocampus: a synthesis from findings with rats, monkeys, and humans. *Psychol. Rev.* 99, 195–231.

Stark, L. W., and Ellis, S. R. (1981). "Scanpaths revisited: cognitive models direct active looking," in *Eye Movements: Cognition and Visual Perception*, eds D. F. Fisher, R. A. Monty, and J. W. Senders. (Hillsdale, NJ: Lawrence Erlbaum Associates), 193–226.

Stephan, B. C. M., and Caine, D. (2009). Aberrant pattern of scanning in prosopagnosia reflects impaired face processing. *Brain Cogn.* 69, 262–268.

Suzuki, W. A., and Amaral, D. G. (2004). Functional neuroanatomy of the medial temporal lobe memory system. *Cortex* 40, 220–222.

Tesche, C. D., and Karhu, J. (2000). Theta oscillations index human hippocampal activation during a working memory task. *Proc. Natl. Acad. Sci. U.S.A.* 97, 919–924.

Theeuwes, J., Belopolsky, A., and Olivers, C. N. L. (2009). Interactions between working memory, attention and eye movements. *Acta Psychologica* 132, 106–114.

Tole, J. R., Stephens, A. T., Vivaudou, M., Harris, R. L. and Ephrath, A. R. (1982). "Entropy, instrument scan, and pilot workload. In *Proceedings of the International Conference on Cybernetics and Society* (New York: IEEE), 588–592.

Tulving, E. (1985). How many memory systems are there? *Am. Psychol.* 40, 385–398.

Van der Stigchel, S., Meeter, M., and Theeuwes, J. (2006). Eye movement trajectories and what they tell us. *Neurosci. Biobehav. Rev.* 30, 666–679.

Walker-Smith, G. J., Gale, A. G., and Findlay, J. M. (1977). Eye movement strategies involved in face perception. *Perception* 6, 313–326.

Warren, D. E., Duff, M. C., Tranel, D., and Cohen, N. J. (2010). Medial temporal lobe damage impairs representation of simple stimuli. *Front. Hum. Neurosci*. 4:35. doi: 10.3389/fnhum.2010.00035.

Williams, L. E., Must, A., Avery, S., Woodward, N. D., and Cohen. N. J., and Heckers, S. (2010). Eye movements reveal disrupted relational memory in schizophrenia. *Biol. Psychiatry*.

Yarbus, A. L. (1967). *Eye Movements and Vision*. New York: Plenum Press.

Zacks, R. T., Hasher, L., and Li, K. Z. (2000). "Human memory," in *The Handbook of Aging and Cognition*, 2nd Edn, eds F. I. M. Craik and T. Salthouse (Mahwah, NJ: Erlbaum), 293–357.

**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 2010; paper pending published: 13 June 2010; accepted: 26 July 2010; published online: 08 October 2010.*
*Citation: Hannula DE, Althoff RR, Warren DE, Riggs L, Cohen NJ and Ryan JD (2010). Worth a glance: Using eye movements to investigate the cognitive neuroscience of memory. Front. Hum. Neurosci.* ***4****:166. doi:10.3389/fnhum.2010.00166*

frontiers in
HUMAN NEUROSCIENCE

GENERAL COMMENTARY
published: 23 December 2010
doi: 10.3389/fnhum.2010.00228

# Quantitative neuroimaging and the prediction of rehabilitation outcome following traumatic brain injury

***Erin D. Bigler[1,2]* and Elisabeth A. Wilde[3]***

[1] *Department of Psychology and Neuroscience Center, Brigham Young University, Provo, UT, USA*
[2] *Department of Psychiatry, University of Utah, Salt Lake City, UT, USA*
[3] *Cognitive Neuroscience Laboratory, Departments of Physical Medicine and Rehabilitation, Neurology, and Radiology, Baylor College of Medicine, Houston, TX, USA*
**Correspondence: erin_bigler@byu.edu*

**A commentary on**

**Regional brain morphometry predicts memory rehabilitation outcome after traumatic brain injury.**
*by Strangman, G. E., O'Neil-Pirozzi, T. M., Supelana, C., Goldstein, R., Katz, D. I., and Glenn, M. B. (2010). Front. Hum. Neurosci. 4:182. doi:10.3389/fnhum.2010.00182.*

Diverse motor, sensory, cognitive, and emotional disabilities may be the aftermath of any brain injury (Svestkova et al., 2010). While some systematic interventions exist to treat TBI-related disabilities, there are only a limited number of evidence-based therapies to treat cognitive impairments, which some have argued are the most prevalent of TBI disabilities (Cicerone et al., 2005; Stuss et al., 2008). Potential reasons for the limited number of therapeutic options in treating cognitive deficits following TBI are numerous, but heterogeneity of the injury is at the top of the list (Skandsen et al., 2010). Each brain injury produces a unique set of pathologies not only in the underlying pathophysiology of the injury itself, but also the injury occurs to a brain with its own unique, and individualized organization. Admittedly, there are some universalities that apply to overall brain organization such as general cognitive functions of temporofrontal areas involved in language processing and the medial temporal lobe in memory, but substantial individual differences exist in how even these general regions may be linked to one another and how their underlying circuitry relates to function (Brown et al., 2010). Because of these individual differences, sampling one type of lesion or region of damage in TBI will likely provide limited understanding of the relationship between brain injury, underlying brain pathology, and rehabilitation outcome especially when TBI produces a complex set of diffuse and focal injuries, where multiple brain regions may exhibit atrophic changes (Warner et al., 2010a,b).

Past cognitive rehabilitation TBI research often used neuroimaging findings to define a location of injury or the amount of structural damage, frequently as a singular measure of regional, or whole brain atrophy. ***All*** quantitative neuroimaging analysis methods of the past, however required time-consuming operator-controlled methods that had to be done by hand. Having lesion metrics that not only assess the diffuseness of injury but also location and/or the degree of where atrophic changes have occurred provides the rehabilitation clinician with additional information that may be useful in predicting outcome or even guiding therapies, because it taps the multifaceted manner in which the brain may be injured. However, as already stated, each brain has its own unique organization which means that singular measures such as just lesion volume, localized, or whole brain measures of atrophy fall short of capturing the complexity of the injury, but that is all that could reasonably be done with past quantitative neuroimaging methods. Furthermore, because any quantification in the past required time-consuming methods of analysis, it was not practical for the rehabilitation clinician to even use such information because it could never be reliably generated in a timely fashion. Fortunately, advances in volumetric measurement have become more automated, and as shown by Strangman et al. (2010) in this issue of Frontiers of Human Neuroscience using FreeSurfer quantification techniques, numerous brain regions from conventional T1-weighted volume-acquisition MRI studies can be calculated in the TBI patient needing rehabilitation intervention. Strangman et al. capitalize on this technology and demonstrate that in TBI of all severities, the key to understanding memory outcome following internal memory training strategies is by performing *multiple* brain morphology measurements (i.e., hippocampal, cingulate, and prefrontal cortical). A singular metric insufficiently captures the complexity of pathological changes in TBI, but multiple metrics do. The internal memory strategies method combined with neuroimaging findings may provide a neurobiological rationale for how to treat memory deficits following TBI (O'Neil-Pirozzi et al., 2010).

Strangman et al. (2010) discuss the implications of their findings with a view toward the future. Evaluating the patency of regions of interest that contribute to memory networks before embarking on cognitive rehabilitation treatment programs may have major implications for TBI rehabilitation. Figure 3 in Strangman et al. shows that if reduced hippocampal and posterior dorsal cingulate volumes are present in TBI patients, then memory deficits are likely to remain substantially impaired, even after the structured memory therapies that were applied in this study.

There are other more futuristic implications of this research. Diffusion tensor imaging (DTI) provides methods to examine the health of white matter connections between regions of interest, and may be particularly sensitive in assessing affected pathways in TBI (Niogi and Mukherjee, 2010). Functional MRI studies are also showing the utility of fMRI activation patterns in assessing the effects of brain injury (Laatsch, 2007). Combining fMRI and DTI with the type of morphometric structural analysis done by Strangman et al. could be important in assessing not only the effectiveness of a treatment regimen, but in determining which rehabilitation treatment modalities are likely to be most useful to an individual patient. Strangman et al. discuss some of these potentialities and this line of research

will hopefully lead to new innovative findings that blends neuroimaging technology with treatment.

As Strangman et al. point out, FreeSurfer is freely available and open for use by any radiology department or neuroimaging center, with the image analysis largely automated, allowing for the potential use of such software in aspects of clinical decision-making. DTI and fMRI analysis programs are also becoming more available, accessible, automated and inexpensive, and combinations of these technologies may be quite fruitful in the future. As always, there are complexities in imaging–behavior relationships that require careful consideration and additional research. For example, the Strangman et al. study utilized participants with chronic TBI, at an interval where both morphometric changes and cognitive recovery have been considered to be relatively stable. However, the relationship between brain morphometry and the cognitive capacity to benefit from rehabilitation strategies may be much more dynamic and complex in the earlier phases of recovery. In the first few months post-injury, an interval where clinical decision-making is being applied for many individuals with TBI, the trajectories of rapid cognitive recovery and incomplete degenerative tissue change may be progressing in opposing directions and to differing degrees. Clearly, much still needs to be understood about this relationship in order for these techniques to be widely utilized as predictive tools in a clinical setting. Additionally, and as mentioned in the study, automated programs for volumetric measurement such as FreeSurfer may have some accuracy limitations in regions of the brain that are difficult to model (e.g., medial temporal areas) or that contain very large lesions or types of pathology that the software cannot accurately distinguish. Furthermore, the general artifacts and limitations of MRI (e.g., metal artifact, motion, etc.) may limit its use in some individuals with TBI, particularly in early subacute stages. Finally, determination of TBI-related cortical change requires an appropriate normative comparison given the dynamic developmental changes that occur throughout the lifespan, particularly in children, adolescents, and the elderly. Despite these caveats, advanced quantitative imaging such as the techniques applied in the Strangman et al. study will likely advance our understanding of the effects of brain injury and how best to conduct rehabilitation therapies in the TBI patient.

## REFERENCES


Brown, G. G., Mathalon, D. H., Stern, H., Ford, J., Mueller, B., Greve, D. N., McCarthy, G., Voyvodic, J., Glover, G., Diaz, M., Yetter, E., Ozyurt, I. B., Jorgensen, K. W., Wible, C. G., Turner, J, A., Thompson, W. K., and Potkin, S. G. (2010). Function Biomedical Informatics Research Network. Multisite reliability of cognitive BOLD data. *Neuroimage*. PMID: 20932915. [Epub ahead of print].

Cicerone, K. D., Dahlberg, C., Malec, J. F., Langenbahn, D. M., Felicetti, T., Kneipp, S., Ellmo, W., Kalmar, K., Giacino, J. T., Harley, J. P., Laatsch, L., Morse, P. A., and Catanese, J. (2005). Evidence-based cognitive rehabilitation: updated review of the literature from 1998 through 2002. *Arch. Phys. Med. Rehabil.* 86, 1681–1692.

Laatsch, L. (2007). The use of functional MRI in traumatic brain injury diagnosis and treatment. *Phys. Med. Rehabil. Clin. N. Am.* 18, 69–85.

Niogi, S. N., and Mukherjee, P. (2010). Diffusion tensor imaging of mild traumatic brain injury. *J. Head Trauma Rehabil.* 25, 241–255.

O'Neil-Pirozzi, T. M., Strangman, G. E., Goldstein, R., Katz, D. I., Savage, C. R., Kelkar, K., Supelana, C., Burke, D., Rauch, S. L., and Glenn, M. B. (2010). A controlled treatment study of internal memory strategies (I-MEMS) following traumatic brain injury. *J. Head Trauma Rehabil.* 25, 43–51.

Skandsen, T., Kvistad, K. A., Solheim, O., Strand, I. H., Folvik, M., and Vik, A. (2010). Prevalence and impact of diffuse axonal injury in patients with moderate and severe head injury: a cohort study of early magnetic resonance imaging findings and 1-year outcome. *J. Neurosurg.* 113, 556–563.

Strangman, G. E., O'Neil-Pirozzi, T. M., Supelana, C., Goldstein, R., Katz, D., and Glenn, M. B. (2010). Regional brain morphometry predicts memory rehabilitation outcome after traumatic brain injury. *Front. Hum. Neurosci.* 4:182. doi: 10.3389/fnhum.2010.00182.

Stuss, D. T., Winocur, G., and Robertson, I. H. (2008). *Cognitive Neurorehabilitation*, 2nd Edn. New York: Cambridge University Press.

Svestkova, O., Angerova, Y., Sladkova, P., Bickenbach, J. E., and Raggi, A. (2010). Functioning and disability in traumatic brain injury. *Disabil. Rehabil.* 32(Suppl. 1), S68–S77.

Warner, M. A., Marquez de la Plata, C., Spence, J., Wang, J. Y., Harper, C., Moore, C., Devous, M., and Diaz-Arrastia, R. (2010a). Assessing spatial relationships between axonal integrity, regional brain volumes, and neuropsychological outcomes after traumatic axonal injury. *J. Neurotrauma*. doi: 10.1089/neu. 2010. 1429. [Epub ahead of print].

Warner, M. A., Youn, T. S., Davis, T., Chandra, A., Marquez de la Plata, C., Moore, C., Harper, C., Madden, C. J., Spence, J., McColl, R., Devous, M., King, R. D., and Diaz-Arrastia, R. (2010b). Regionally selective atrophy after traumatic axonal injury. *Arch Neurol.* 67, 1336–1344.


*Received: 27 October 2010; accepted: 30 November 2010; published online: 23 December 2010.*

*Citation: Bigler ED and Wilde EA (2010) Quantitative neuroimaging and the prediction of rehabilitation outcome following traumatic brain injury. Front. Hum. Neurosci. 4:228. doi: 10.3389/fnhum.2010.00228*

frontiers in
HUMAN NEUROSCIENCE

REVIEW ARTICLE
published: 14 October 2010
doi: 10.3389/fnhum.2010.00177

# Rhythmic pulsing: linking ongoing brain activity with evoked responses

***Ali Mazaheri*[1,2]*** ***and Ole Jensen*[1]**

[1] *Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, Nijmegen, Netherlands*
[2] *Center for Mind and Brain, University of California Davis, Davis, USA*

***Edited by:***
*Michael X. Cohen, University of Amsterdam, Netherlands*

***Reviewed by:***
*Paul Sauseng, University Medical Center Hamburg-Eppendorf, Germany*
*Catherine Tallon-Baudry, Universite Pierre et Marie Curie, France*

***Correspondence:**
*Ali Mazaheri, Donders Institute for Brain, Cognition and Behaviour. Radboud University Nijmegen, P. O. Box 9101, 6500 HB Nijmegen, Netherlands.*
*e-mail: ali.mazah@gmail.com*

The conventional assumption in human cognitive electrophysiology using EEG and MEG is that the presentation of a particular event such as visual or auditory stimuli evokes a "turning on" of additional brain activity that adds to the ongoing background activity. Averaging multiple event-locked trials is thought to result in the cancellation of the seemingly random phased ongoing activity while leaving the evoked response. However, recent work strongly challenges this conventional view and demonstrates that the ongoing activity is not averaged out due to specific non-sinusoidal properties. As a consquence, systematic modulations in ongoing activity can produce slow cortical evoked responses reflecting cognitive processing. In this review we introduce the concept of "rhythmic pulsing" to account for this specific non-sinusoidal property. We will explain how rhythmic pulsing can create slow evoked responses from a physiological perspective. We will also discuss how the notion of rhythmic pulsing provides a unifying framework linking ongoing oscillations, evoked responses and the brain's capacity to process incoming information.

**Keywords: alpha oscillations, evoked responses, inhibition, amplitude asymmetry**

EEG and MEG are used in neuroimaging studies to provide a real-time measure of respectively electric and magnetic fields produced by neuronal activity in the brain. The majority of EEG/MEG research attempting to link human cognition to neuronal activity examines the neuronal events (i.e., evoked response) *after* the occurrence of an event. The ongoing activity (already present prior to the stimulus) is often considered irrelevant and is assumed to be "averaged out" across trials (**Figure 1C**). Recent work has challenged the dogma that ongoing activity can simply be averaged out across trials. The key aspect of these studies is the revelation that the ongoing activity in the frequency of 10 Hz (i.e., alpha band) contains a non-sinusoidal property referred to as "amplitude asymmetry" or "baseline shift" (**Figure 1B**) (Nikulin et al., 2007; Mazaheri and Jensen, 2008; van Dijk et al., 2010b).

The amplitude asymmetry of ongoing oscillations entails that the peaks are modulated stronger than troughs (or *vice versa*) (**Figure 1B**) irrespective of the DC offset/zero-line of the signal. Amplitude asymmetric oscillations have profound consequences for event-related averaging. For example if the amplitude of ongoing oscillations is suppressed such that only the peaks are reduced in magnitude but not the troughs, averaging across trials would result in the emergence of an evoked response with a shape that is similar to the time course of amplitude suppression of the oscillation (**Figure 1D**). Had the magnitudes of the peaks and troughs been symmetrically decreased, no evoked responses would have been generated. The critical prerequisite for this mechanism is the differential modulation of the peaks and troughs which we propose is caused by unidirectional primary currents in pyramidal cells producing the oscillations (to be discussed further).

Because of the presence of amplitude asymmetry, slow evoked responses can be generated from simple changes in the amplitude of ongoing alpha activity without any turning "on" of additional brain activity.

In this review we will discuss the concept of amplitude asymmetry and its implications to the brain's post-stimulus responses. Furthermore, we propose that ongoing alpha activity along with amplitude asymmetric properties can be conceived as rhythmic pulsing which produces bouts of inhibition every 100 ms. Processing of a stimulus can only occur between two alpha pulses; this is consistent with the view that visual processing relies on discrete processing (VanRullen and Koch, 2003). Importantly we conjecture that this rhythmic inhibition only occurs when the "pulses" of alpha activity are at a sufficiently high amplitude. We will attempt to demonstrate that the rhythmic pulsing framework fits nicely with other theoretical models of alpha activity and could potentially offer a unified account of the brain's ongoing activity, discrete stimulus processing and evoked responses.

## ONGOING OSCILLATIONS AND EVOKED RESPONSES

The evoked response reflects the brain's transient time-locked response to a stimulus or event. Evoked responses are calculated by averaging several trials (typically 30 to a 100) time-locked to a given stimulus or event. Subsequently the pre-stimulus activity (typically in a 100- to 200-ms interval) is subtracted from the trial average. The relationship between ongoing activity and evoked responses has been a matter of debate for several decades. There are currently three different theories which attempt to account for how evoked responses are generated: *additivity, phase-resetting* and the recently proposed *amplitude asymmetry* mechanism (also termed *baseline-shift*).

The additive and phase-resetting models focus on the early evoked components. These early ERPs/ERFs (sometimes referred to as "exogenous components") are transient components that occur within the first few hundred milliseconds of stimulus presentation, and are widely believed to index the arrival of information to the cortex (Coles and Rugg, 1995). The amplitude asymmetry theory focuses on the late occurring components (often referred to as "endogenous") which emerge at least 100 ms after stimulus onset and are often sustained for hundred milliseconds or longer. There a numerous examples of the slow late components being modulated by various cognitive tasks and as such they can be viewed as a link between electrophysiology and cognition (Walter et al., 1964; Kutas and Hillyard, 1980; Sanquist et al., 1980; Hagoort and Brown, 2000; Soltani and Knight, 2000; Kilner et al., 2004; Vogel et al., 2005; Takashima et al., 2006; Rugg and Curran, 2007).

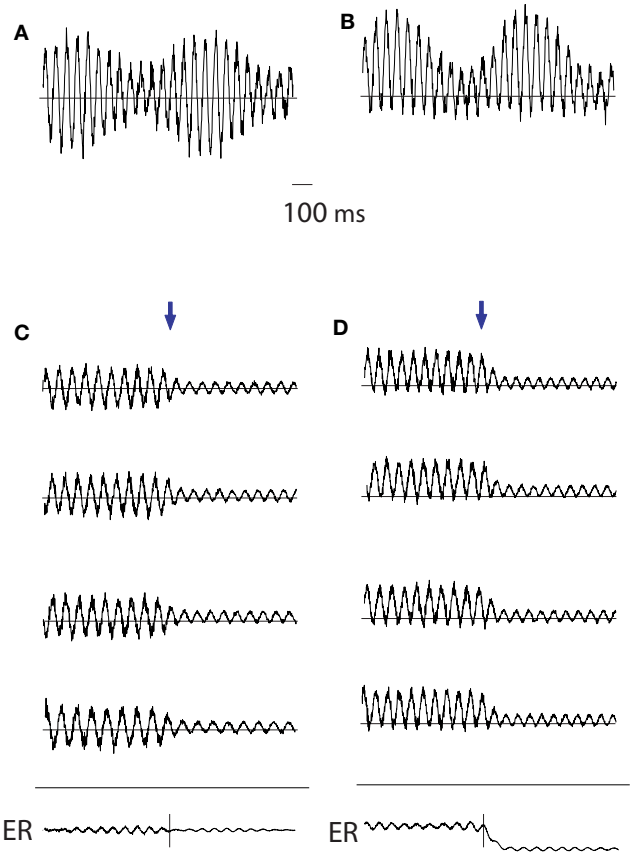



**FIGURE 1 | The concept of "amplitude asymmetry." (A)** The amplitude modulation of neuronal oscillatory activity is conventionally viewed as being symmetric around zero. **(B)** We propose that the amplitude modulations of the oscillatory activity are *asymmetric* such that the peaks are more strongly modulated than the troughs. For the 10-Hz alpha activity, this could be explained by bouts of activity every ~100 ms. **(C)** The conventional view ignoring asymmetric modulations of oscillatory activity would mean that averaging across trials (the arrow representing the start of the evoked response) would not result in the generation of slow fields. **(D)** As a direct consequence of amplitude asymmetry, a depression (or increase) in alpha activity in response to a stimulus will result in the generation of slow fields when multiple trials are averaged. Adapted from Mazaheri and Jensen (2008).

## ADDITIVE THEORY VERSUS PHASE-RESETTING THEORY OF EVOKED RESPONSES

The additive theory implies that evoked and ongoing activities are separate and distinct neuronal phenomena. According to this view the stimulus "evokes" a phase-locked response adding to the activity in each trial (**Figure 2A**). In contrast, according to the phase-resetting theory, the phases of the ongoing background oscillations become aligned (phase-reset or partial phase-reset) to the stimulus (**Figure 2B**). By averaging the stimulus-locked trials, the phase-locked oscillatory activity emerges as the evoked component in the average. Since alpha oscillations (8–12 Hz) are the predominant ongoing activity in the EEG/MEG, it is believed that the phase-resetting of these oscillations is particularly relevant for producing the evoked activity (Makeig et al., 2002; Klimesch et al., 2004; Gruber et al., 2005).

Part of the issue in disambiguating the additive and phase-resetting theory is that the addition of a response can look much the same as a phase-reset of the oscillatory background activity. There are a number of informative and critical discussions about the merits of the phase-resetting versus additive modeling of evoked response generations (Fell et al., 2004; Shah et al., 2004; Makinen et al., 2005; Mazaheri and Jensen, 2006; Klimesch et al., 2007b, 2009; Min et al., 2007; Becker et al., 2008; Risner et al., 2009; Ritter and Becker, 2009). It is still a matter of debate how general the phase-resetting mechanism is for the generation of evoked responses.

## AMPLITUDE ASYMMETRY AS A MECHANISM FOR THE GENERATION OF EVOKED RESPONSES

The amplitude fluctuations of oscillatory activity are conventionally viewed as being symmetric around zero (**Figure 1A**). Amplitude asymmetry refers to the observation that modulations of ongoing activity affect the peaks and the troughs of the alpha activity to different extents. This would mean that only the peak values

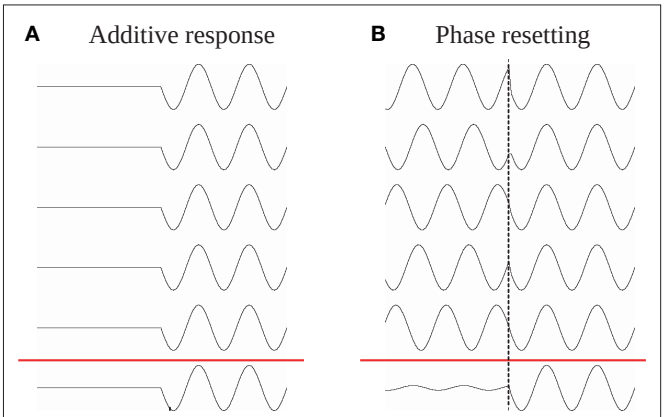



**FIGURE 2 | The additive and phase-resetting theory of evoked response generation. (A)** The additive theory implies that evoked and ongoing activities are separate and distinct neuronal phenomena. The stimulus "evokes" an additive, phase-locked response in each trial. **(B)** According to the phase-resetting view, upon the onset of a stimulus, the phases of the ongoing background oscillations become aligned (phase-reset or partial phase-reset) to the stimulus. By averaging the stimulus-locked trials, the phase-locked oscillatory activity emerges as an evoked component.

increase or decrease, while the trough values stay the same (*or vice versa*) (**Figure 1B**). Amplitude asymmetry can be found in a lot of systems around us. One way to conceptualize amplitude fluctuation asymmetry is to think of a bouncing ball: how high it jumps varies, but it cannot go lower than the floor. Another example is the light intensity in your office over days: the light at noon will vary a lot more over days compared to the light at mid-night.

One important consequence of this amplitude asymmetry of ongoing oscillations is that event-related modulations of oscillatory activity would not be "averaged out" over trials, but instead lead to the formation of slow evoked responses (**Figure 1D**). This is explained by a systematic depression of the peaks in response to the stimuli while the troughs remain the same. Since the phases of the individual trials are different, this will produce a slow deflection in the evoked response after trial averaging.

## EVIDENCE FOR THE AMPLITUDE ASYMMETRY OF OSCILLATIONS

The amplitude asymmetry of the ongoing EEG was appreciated by Stam et al. using a measure that demonstrated that the predictability of the EEG signal in time was reduced when the signal was inverted from "peak to trough" (Stam et al., 1999b). However, the link between evoked responses and the "non-zero mean" property in oscillations was first discussed by Nikulin et al. (2007). The authors found a correlation between low frequency drifts (referred to as baseline shifts) and the ~10 Hz somatosensory rhythm. The authors speculated that the resulting baseline shifts could play a role in the formation of somatosensory evoked responses.

A direct link between the amplitude asymmetry property of oscillations and evoked responses was empirically demonstrated by Mazaheri and Jensen (2008). In this study, a measure called the Amplitude Fluctuation Asymmetry Index ($AFA_{index}$) was first developed to quantify the asymmetry of amplitude fluctuations. The $AFA_{index}$ compares the variance of the peaks with the variance of the troughs by considering the normalized difference between the two measures. Later we will elaborate on the details of this measure. Using this $AFA_{index}$ Mazaheri and Jensen (2008) were able to show that in ongoing posterior alpha activity the peaks and the troughs were indeed modulated differently. More importantly, it was shown that the direction (i.e., stronger modulation of peaks than troughs or *vice versa*) and magnitude of the $AFA_{index}$ during a rest condition correlated with respectively the amplitude and polarity of slow ERFs in response to simple visual stimuli. Thus this study provided strong support for the notion that systematic modulations of oscillatory activity with amplitude asymmetry can produce slow evoked responses.

Recently, van Dijk et al. (2010b) were able to extend this link by demonstrating that an evoked response modulated by a cognitive task could be explained by systematic modulations in oscillatory activity. In particular the study focused on the contralateral delayed activity (CDA) component, which is a slow evoked response typically observed in working memory tasks in which hemifield specific attention is manipulated (Vogel and Machizawa, 2004; Vogel et al., 2005; Fukuda et al., 2010). The key finding was that the emergence of the CDA could be explained by modulations in alpha activity (**Figure 3**). Previous studies have suggested alpha activity to be

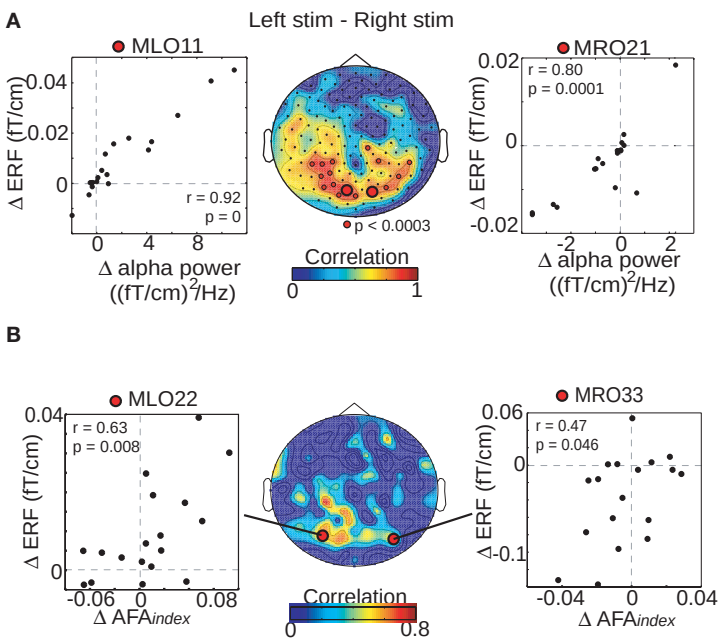



**FIGURE 3 | Relation between the evoked component produced in a working memory task and modulations in hemispheric alpha power lateralization. (A)** Correlation over subjects of differences in ERFs and alpha power modulations [left minus right target stimuli (stim)] for a left sensor and a right sensor (Left and Right, respectively). The correlations were highly significant. The topography (Middle) of the correlation coefficients. **(B)** Correlations over subjects between the differences (left minus right stimuli) in sustained ERFs and the differences in absolute alpha amplitude asymmetry ($AFA_{index}$) during retention. Adapted from van Dijk et al. (2010b).

involved in the functional inhibition of task-irrelevant regions (Klimesch, 1999; Klimesch et al., 1999, 2007a; Jensen et al., 2002; Jokisch and Jensen, 2007; Tuladhar et al., 2007; Jensen and Mazaheri, in press). Thus van Dijk et al. (2010b) proposed that the CDA is not attributable to an additive process reflecting memory maintenance *per se* but rather to changes in ongoing oscillatory alpha activity reflecting inhibition of task-irrelevant regions, while routing information to task-relevant regions. This view is further supported by recent findings of Sauseng et al. (2009) in which a similar design was used to demonstrate that modulation of hemispheric alpha lateralization predicted memory capacity based on efficient suppression of irrelevant information in short-term memory.

Although the amplitude asymmetry model suggests that certain endogenous responses are produced by modulations in oscillatory power, this does not question the merits of previous ERP/ERF studies. If a specific slow ERP/ERF is revealed to be produced by modulations in oscillatory activity, this does not mean that the conventional ERP/ERF analysis is inappropriate as a tool in cognitive neuroscience. Rather, if the mechanism of a particular evoked response can be linked to a modulation of ongoing activity, it could provide a stronger account for the neuronal substrate generating the slow evoked responses. Also, it can help interpreting the functional role of the evoked responses given that one can build on the insight gained from the role of oscillatory activity (see, e.g., van Dijk et al., 2010b).

## PREREQUISITE FOR LINKING THE MODULATION OF ONGOING OSCILLATIONS TO EVOKED COMPONENTS

How is it possible to determine if the mechanism behind a specific evoked response is due to modulation of ongoing activity with amplitude asymmetry? We introduce four prerequisites for linking modulations of oscillatory activity to evoked component generation.

1. The ongoing MEG/EEG oscillations must be modulated in amplitude by the stimuli or event
2. This amplitude modulation of the ongoing activity must correlate with the time course of the evoked response (over trials or subjects)
3. The ongoing oscillations must have an amplitude asymmetry
4. The magnitude and/or polarity of the amplitude asymmetry must relate to the amplitude and/or polarity of the evoked responses (over trials or subjects)

Event-related changes in oscillatory alpha activity have been demonstrated in many cognitive paradigms (Klimesch et al., 1997; Salenius et al., 1997; Klimesch, 1999; Jensen et al., 2002; Makeig et al., 2004; Mazaheri and Picton, 2005; Bastiaansen and Hagoort, 2006; Sauseng et al., 2006; Jokisch and Jensen, 2007). Perhaps the most challenging step in linking oscillatory changes to evoked responses is the quantification of the amplitude asymmetry of the ongoing signal. For this step, two different measures have so far been proposed: the $AFA_{index}$ (Mazaheri and Jensen, 2008) and baseline shift index (BSI) (Nikulin et al., 2007).

## ANALYTICAL METHODS FOR MEASURING AMPLITUDE ASYMMETRY

The $AFA_{index}$ quantifies the ratio of the variance of the peaks and troughs where $S_{peaks}$ and $S_{troughs}$ refer to the peak and trough values identified in the ongoing oscillatory signal:

$$AFA_{index} = \frac{Var(S_{peaks}) - Var(S_{troughs})}{Var(S_{peaks}) + Var(S_{troughs})} \quad (1)$$

When using the $AFA_{index}$, the signal is first bandpass-filtered in the frequency band of interest (**Figure 4A**). Next the time points for the peaks and troughs of the bandpassed data can be identified. These time points are then used to obtain the signal values of peaks and troughs in the raw data. The variance of these values is then used to calculate the $AFA_{index}$.

Accordingly, positive $AFA_{index}$ values indicate a stronger modulation of the peaks and negative values indicate a stronger modulation of the troughs, while values close to zero indicate symmetrical modulation.

By considering the difference between the variance of the peaks and troughs to the ratio of amplitudes, this measure is relatively immune to DC-like offsets that sometimes appear in MEG and EEG measurements. In order to ensure that this measure was not a consequence of a slow DC offset interacting with the alpha rhythm we (Mazaheri and Jensen, 2008) investigated various principles of actions using constructed surrogate signals. These simulations demonstrated that the $AFA_{index}$ successfully can detect true asymmetric amplitude fluctuation while being immune to slow multiplicative or additive modulations. The simulations can be seen in **Figure 5**.

When applying the $AFA_{index}$ one issue to consider is the time window in which the variance across peaks and troughs are measured. This time window would be dependent on the frequency of interest. A very short time window is not optimal, since then there simply would not be enough peaks/troughs to reliably assess the variance. To quantity the amplitude asymmetry of alpha and beta oscillations Mazaheri and Jensen (2008) used a time window of 5 s to calculate the asymmetry index, whereas van Dijk et al. (2010b) was able to successfully do this using a period of 1 s. However, future empirical work needs to be done to precisely characterize the optimal temporal parameters relevant for detection of amplitude asymmetry.

It should be mentioned that the $AFA_{index}$ can potentially be sensitive to harmonics present in the signal (Nikulin et al., 2010). For instance, 20 Hz harmonics can in some cases produce a non-zero $AFA_{index}$ in the 10 Hz band. However, a non-zero $AFA_{index}$ due to harmonics cannot produce slow evoked responses (Jensen et al., 2010; Nikulin et al., 2010) and as such would not have a relationship to the amplitude and polarity of evoked responses. One way to reduce the effect of harmonics on the $AFA_{index}$ is to lowpass filter the raw data at the frequency below the harmonic but above the frequency of interest (in the case of the alpha band it would be ~15 Hz). Importantly the $AFA_{index}$ can be correlated with both magnitude and/or the polarity of evoked responses (**Figures 3 and 6**). This has been successfully done for both visual evoked responses and working memory related evoked responses (Mazaheri and Jensen, 2008; van Dijk et al., 2010b). Thus, given that the $AFA_{index}$ was strongly correlated with the evoked responses, these findings cannot be explained by harmonics of the alpha oscillation.

The BSI is an alternative measure to the $AFA_{index}$ and quantifies the correlation between oscillatory activity in a given frequency band and low frequency fluctuations (**Figure 7**). First, the ongoing activity is filtered in two ways: using bandpass filters (e.g., 8–12 Hz) and a lowpass filter at 3 Hz. The BSI reflects the regression between the bandpassed and lowpass filtered signal's amplitude values

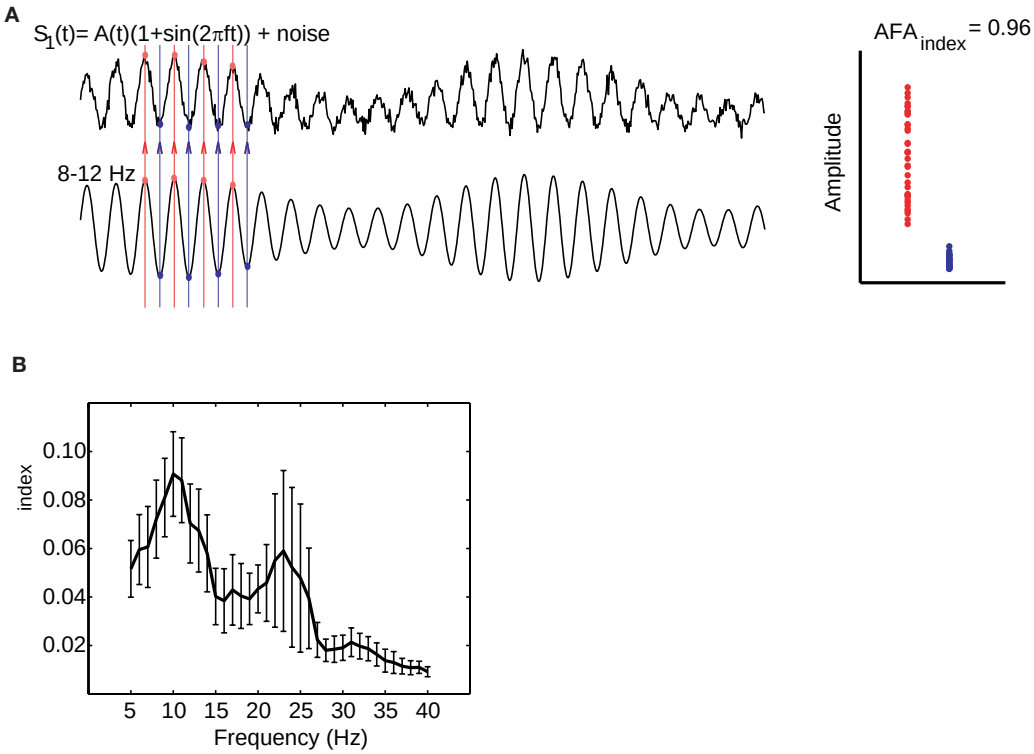



**FIGURE 4 | (A)** The $AFA_{index}$ quantifies the ratio of the variance of the peaks (red dots) versus the variance of the troughs (blue dots) identified in the ongoing oscillatory signal. The signal is first bandpass-filtered in the frequency band of interest. Next the time points for the peaks and troughs of the bandpassed data can be identified. These time points can then be used to obtain the signal values of peaks and troughs in the raw data. **(B)** The $AFA_{index}$ can be used to calculated amplitude asymmetry spectra of spontaneous MEG data during rest. Note that the amplitude asymmetry is constrained to the alpha and beta band. Adapted from Mazaheri and Jensen (2008).

(Nikulin et al., 2007). The measure does depend on extensive preprocessing and it would be highly interesting to investigate if it has a quantitative relationship with evoked responses.

In conclusion amplitude asymmetry has been demonstrated to be an intrinsic property of ongoing oscillatory activity in the alpha band using different analytic techniques. Moreover a strong link-between the amplitude asymmetry of alpha activity and evoked responses has also been established. We propose that amplitude asymmetric oscillations can be viewed as rhythmic pulsing through which information processing is facilitated or inhibited. In the next section we will discuss the hypothesized underlying physiology of amplitude asymmetric alpha activity and its function in rhythmic pulsing.

## THE UNDERLYING PHYSIOLOGY OF RHYTHMIC PULSING

Which physiological mechanisms can account for amplitude asymmetry? EEG and MEG signals are primarily thought to be produced by dendritic currents in pyramidal cells (Hamalainen et al., 1993). These intracellular currents are typically generated by electrical events at the apical dendrite or the soma. Such events include excitatory synaptic input at distal dendrites and after-hyperpolarization currents (Wu and Okada, 1999; Murakami and Okada, 2006). In order to generate oscillatory activity with symmetric amplitude fluctuations (**Figure 1A**) there must be intracellular currents propagating forward and backward down the dendrites with the same magnitude (respectively producing the troughs and peaks). However, even though back-propagating dendrite currents are known to exist, it would be unlikely that they exactly match the synaptic forward propagating currents. We conjecture that the currents producing the alpha activity are asymmetric, i.e., primarily explained by forward propagating dendritic currents most likely due to excitatory synaptic inputs and after-hyperpolarization currents. It is this asymmetry between the magnitude of forward and backwards current flow that gives the ongoing alpha activity its amplitude asymmetry property (**Figures 1B and 8B**). It should be mentioned that these primary intracellular dendritic currents produce instantaneous return currents (also known as volume currents). While MEG primarily detects the magnetic fields produced by the intracellular dendritic currents, EEG detects the differences in scalp potentials arising from the return currents. An illustration of this proposed neurobiological mechanism of asymmetry and its implications for MEG and EEG measurements can be seen in **Figure 8**.

We postulate that the amplitude asymmetric alpha activity can be viewed as rhythmic pulses producing bouts of inhibition repeated every 100 ms. How might this inhibition come about? GABAergic feedbacks from interneurons have been strongly implicated in the physiological mechanism generating the alpha rhythm (Lopes da Silva et al., 1976; Crunelli and Leresche, 1991; Jones et al., 2000). Thus it is possible that rhythmic neuronal activity generating the alpha oscillations is a consequence of a GABAergic inhibitory feedback paced by neocortical or thalamic rhythm generators (Hughes and Crunelli, 2005; Lorincz et al., 2008, 2009). This

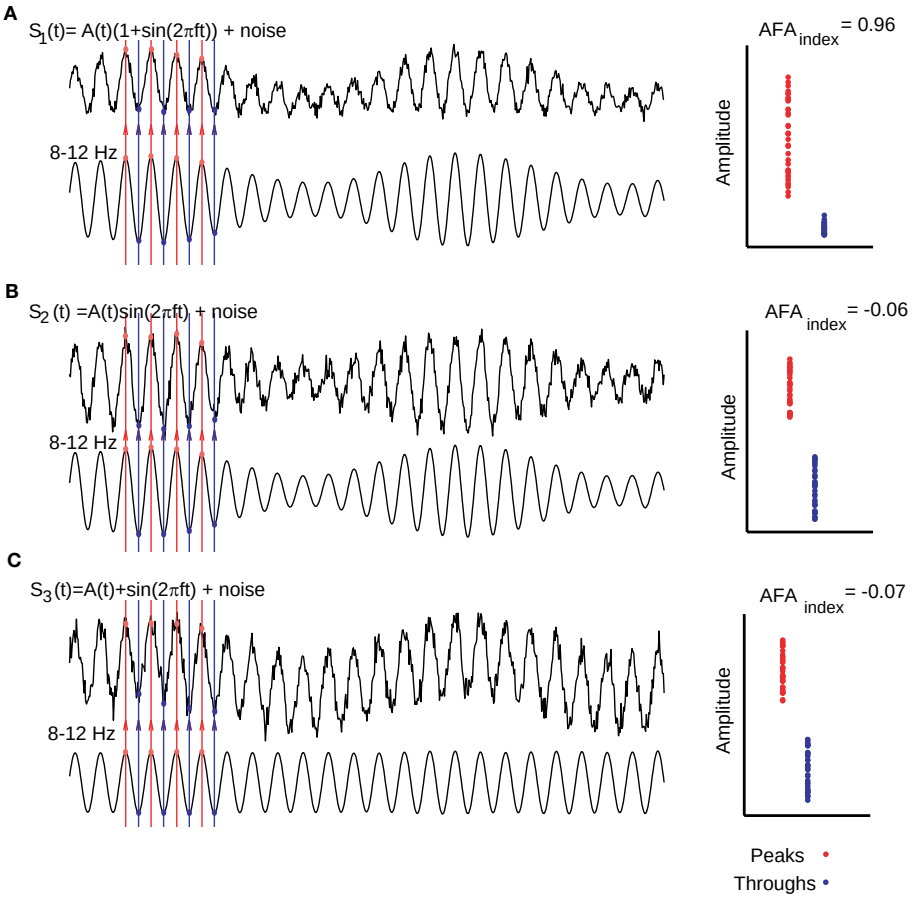



**FIGURE 5 | Various simulations in which surrogate signals were used to test the $AFA_{index}$. (A)** The signal, $s_1(t)$, was designed to have an amplitude asymmetry. The amplitude modulation was determined by a slower signal A(t). Clearly the peaks (red dots) are more modulated than the troughs (blue dots) yielding a strong $AFA_{index}$. **(B)** The signal, $s_2(t)$, was designed such that the slow modulations, A(t), affected the alpha rhythm in a multiplicative manner. Thus peaks and troughs are modulated symmetrically over time yielding an $AFA_{index}$ close to 0. **(C)** In signal $s_3(t)$ slow modulations added to the alpha oscillations (DC-like offset of the signal). This affected peaks and troughs in the same direction producing an $AFA_{index}$ close to 0. Adapted from Mazaheri and Jensen (2008).

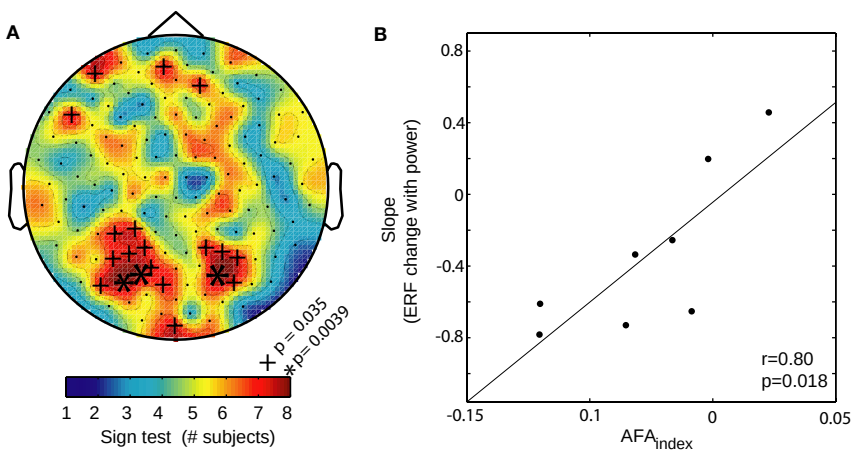



**FIGURE 6 | The $AFA_{index}$ can used to predict the polarity and magnitude of evoked responses. (A)** The consistency between the sign of the $AFA_{index}$ (from eyes closed data) and the sign of the modulation in slow visually evoked ERF with alpha power. The color code represents the number of consistent signs over the eight subjects. More than six consistent signs are considered significant (binomial test). **(B)** The correlation between the $AFA_{index}$ and the slope of slow ERF modulation with alpha power. The high correlation strongly suggests that the slow modulations in the ERFs are produced by changes in asymmetric amplitude changes of alpha power. Adapted from Mazaheri and Jensen (2008).

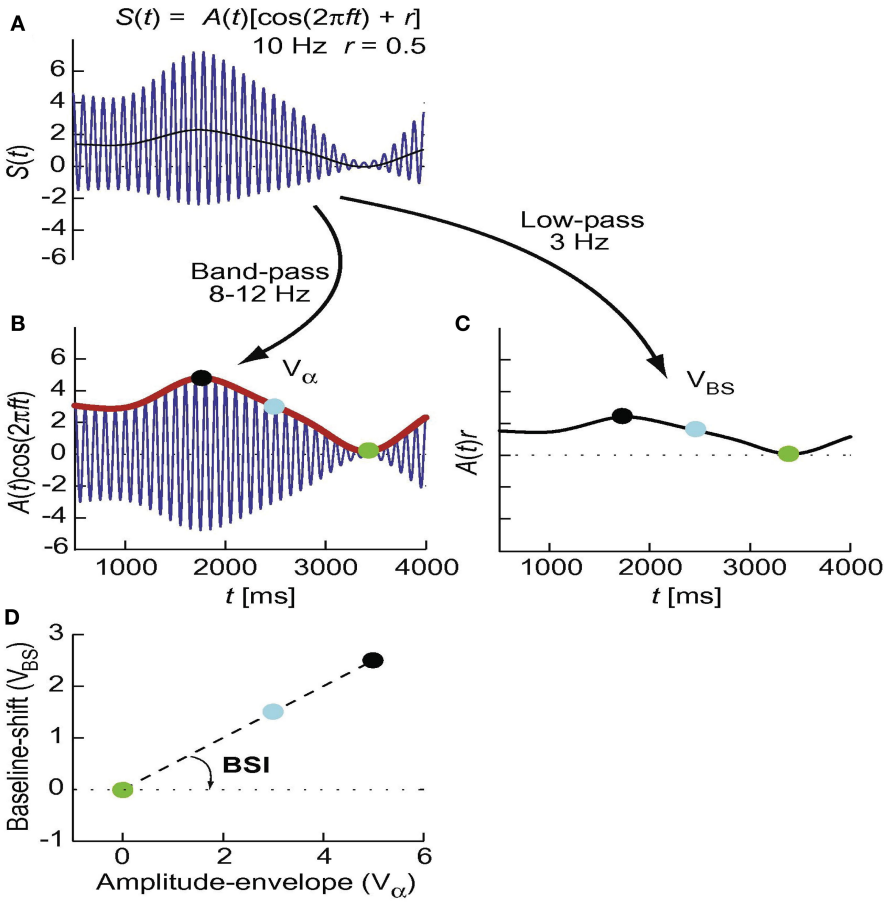



**FIGURE 7 | Schematic presentation of the derivation of the baseline shift index (BSI).** First, the signal in **(A)** is filtered in two ways to separate the amplitude envelope of the oscillations and the baseline shift A(t)r [thin black line in **(A)**]. The amplitude envelope ($V_\alpha$) is obtained using bandpass filters (e.g., 8–12 Hz) and the Hilbert transform [**(B)**, red thick line]. The baseline shifts ($V_{BS}$) are part of the signal at considerably lower frequencies than the oscillatory "carrier" signal and, therefore, are obtained by lowpass filtering, e.g., at 3 Hz **(C)**. The amplitude-envelope values are sorted into bins (e.g., 20 bins) and values from the baseline shift signal are sorted into the corresponding bins (note the temporal correspondence of dots with the same color). **(D)** Within-bin averaged values for the amplitude envelope and baseline shifts are plotted against each other and the slope of the regression line is calculated. This slope is the baseline shift index or BSI. Adapted from Nikulin et al. (2007).

GABAergic feedback could serve to directly silence processing in pyramidal neurons or reduce the efficacy of excitatory input by shunting inhibition.

In order to further explain how pulsed inhibition could be reflected in the neuronal firing of pyramidal cells and subsequently the MEG signal, we constructed a simple model (**Figure 9**). In this model, 500 pyramidal cells are modulated by a 10 Hz inhibitory signal. The signal serves to silence the firing of the pyramidal cells in a phasic manner. As a consequence a pattern of rhythmic pulsing emerges. The MEG signal is thought to be produced by the dendritic current in the pyramidal cells. Thus the inhibition results in an MEG signal having amplitude asymmetry. This model might also explain the inverse relationship between the alpha activity and the BOLD signal (Goldman et al., 2002; Goncalves et al., 2006; Laufs et al., 2006; Ritter et al., 2009; Scheeringa et al., 2009; Yuan et al., 2010). In the example in **Figure 9**, the firing rate is highest (e.g., 0.4–0.6 s) when the "10 Hz signal" is low compared to when the "10 Hz signal" is high (e.g., 0.8–1.0 s). Since firing rate is linked to the BOLD signal, it would have a negative relationship to the alpha activity.

Interestingly, amplitude asymmetry can emerge from even very simple models of oscillatory activity. In 1976, Lopes da Silva et al. proposed a computational model that could account for certain features of the alpha activity. The model was composed of thalamocortical relay neurons coupled with inhibitory interneurons (Lopes da Silva et al., 1976; Stam et al., 1999a). As seen in **Figure 10**, this model produced alpha oscillations with amplitude asymmetry (even though this was not essential for the authors). This underscores that amplitude asymmetry is a natural phenomenon that easily can arise from physiological models of oscillatory activity. In fact, we posit that amplitude asymmetry is the norm and amplitude symmetry is the exception.

## EVIDENCE SUPPORTING RHYTHMIC PULSING

### CROSS-FREQUENCY PHASE-AMPLITUDE COUPLING

The rhythmic pulsing view fits conceptually well with recent evidence that the power of gamma oscillations is phase-locked to posterior alpha activity. Gamma oscillations have long been implicated in neuronal processing in various tasks (Fell et al., 2003;

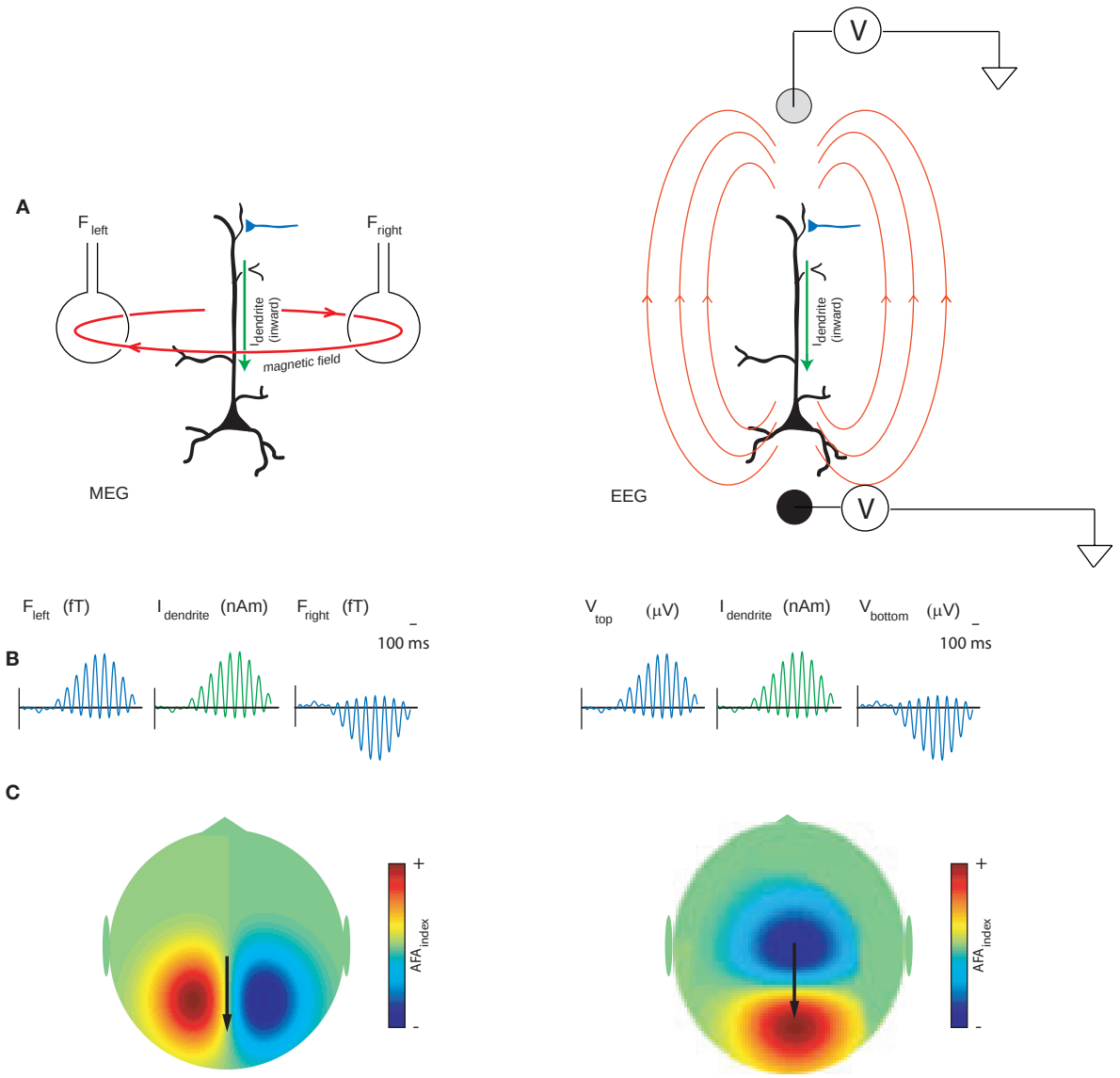
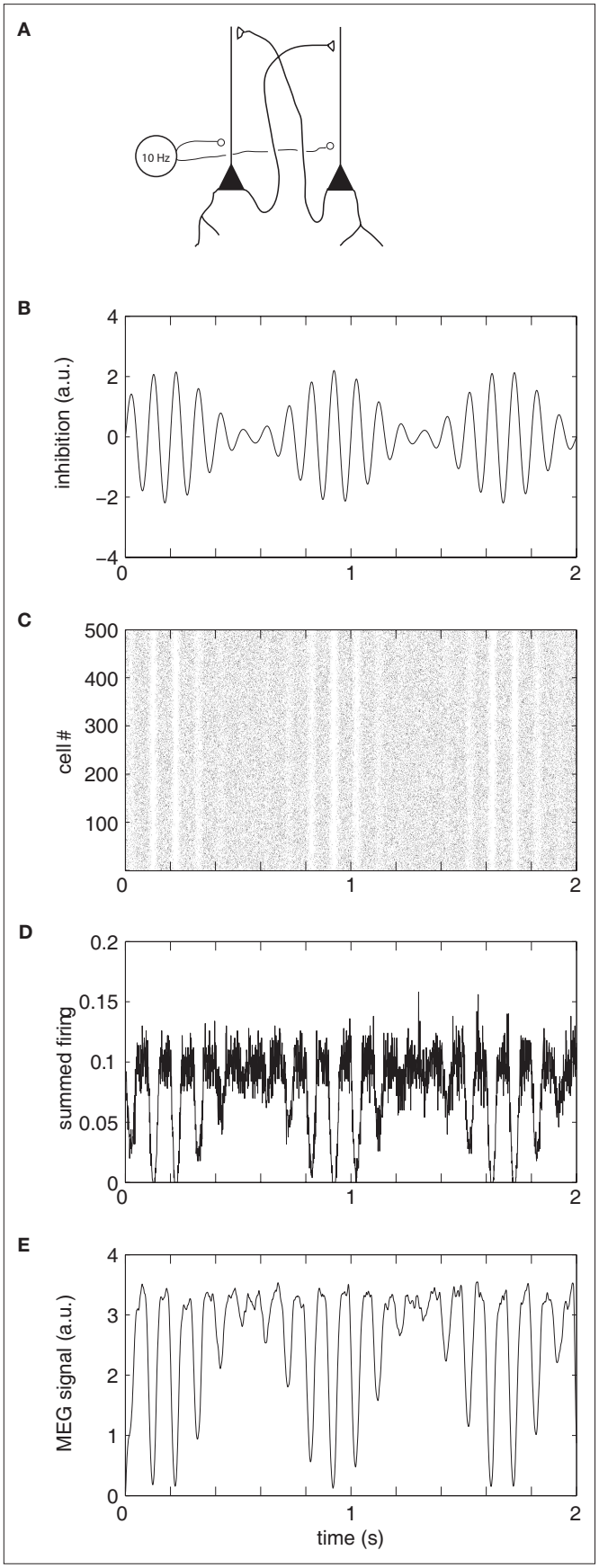



**FIGURE 8 | (A)** MEG signals measured outside the human scalp are primarily thought to be produced by magnetic fields induced by dendritic intracellular currents ($I_{dendrite}$) in pyramidal cells. EEG detects the differences in scalp potentials arising from the return currents. The polarity of the field will be determined by the direction (inward or outward) of the currents in the dendrites. **(B)** We propose that the amplitude modulations of the oscillatory activity are asymmetric such that the peaks are modulated stronger than the troughs. For the 10 Hz alpha activity this is explained by bouts of inward dendritic currents every ~100 ms. **(C)** The topography of the amplitude asymmetry of the MEG fields and EEG potentials. Adapted and modified from Mazaheri and Jensen (2008).

Lachaux et al., 2005; Womelsdorf et al., 2006; Fries et al., 2007; Jensen et al., 2007). One of the models proposed for explaining this interaction entails the amplitude fluctuations of alpha activity to be asymmetric (**Figure 11**). According to this model the gamma burst responsible for visual processing can only occur at alpha phases when the inhibition is sufficiently low (i.e., the trough). When the alpha inhibition is very weak, gamma can occur over the entire cycle of alpha. Such a model of alpha phase to gamma power interaction is consistent with many studies that have found a decrease in alpha to accompany and increase in gamma (Basar et al., 2001; Jokisch and Jensen, 2007; de Lange et al., 2008; Shahin et al., 2009). However, it must be noted that although we conjecture that amplitude asymmetry fits readily with phase-power coupling across oscillations, it is not necessary for oscillations to be amplitude asymmetric to exhibit phase-power relationships. Moreover a direct a functional link between alpha phase and gamma power still remains to be empirically determined.

## ONGOING ALPHA AMPLITUDE AND INHIBITION

Klimesch et al. (2007a) recently proposed the inhibition-timing hypothesis where alpha oscillations play an important role in the brain's capacity to process information. They postulated that a reduction in the amplitude of the ongoing alpha activity reflects a state of comparatively high excitability, whereas high amplitudes reflect a state of inhibition (comparatively low excitability).




In fact it was first suggested almost 80 years ago that that the cortex exhibits cyclic changes between maximal and minimal responsiveness (Bishop, 1933). Since then a number of studies have showed an influence of the phase of the ongoing alpha activity on processing of visual stimuli (Callaway and Yeager, 1960; Varela et al., 1981). Two very recent EEG studies by Mathewson et al. (2009) and Busch et al. (2009) have reported a functional link between the phase of the pre-stimulus alpha oscillations and conscious perception by using threshold stimuli. The authors conjecture that the alpha inhibition is limited to parts of the cycle, generating a form of "pulsed inhibition." We suggest that this pulsed inhibition occurs as a function of the amplitude asymmetric property of the ongoing oscillations.

Evidence supporting the inhibitory nature of alpha activity has been found in a number of studies that demonstrate an increase or decrease in the amplitude of alpha activity depending on the visual stream engaged in a given task. (Worden et al., 2000; Thut et al., 2003, 2006; Kelly, et al., 2006; Jokisch and Jensen, 2007; Medendorp et al., 2007; Rihs et al., 2007; Romei et al., 2007, 2008; van Dijk et al., 2008; Zhang et al., 2008). The functional role of alpha in defining the brain state does not seem to be restricted to visual tasks. A recent study has found alpha power lateralization at parieto-occipital sites to be modulated by the direction of auditory attention to continuous speech in space (Kerlin et al., 2010). In a working memory study on maintaining pitches, alpha activity from left superior temporal areas increased during the retention interval (van Dijk et al., 2010a). This temporal alpha activity, possibly being the tau-rhythm (Lehtela et al., 1997) is likely to reflect inhibition of the left auditory cortex in order to allocate resources to the right auditory cortex involved in pitch processing. In a somatosensory working memory task, the alpha activity decreased in the primary sensorimotor cortex contralateral to the engaged hand while it increased in the ipsilateral hemisphere (Haegens et al., 2010). In summary these studies are consistent with the idea that the functional inhibition of a region in the cortex is mediated by increasing oscillatory activity in the alpha band (8-13 Hz) (Jensen and Mazaheri, in press).

**FIGURE 9 | A simple model explaining how pulsed inhibition is reflected in neuronal firing of pyramidal cells and subsequently the MEG signal. (A)** Pyramidal cells are mutually synaptically coupled. The synaptic currents produce the MEG signal. We here assume a Poisson distribution of firing which is phasically silenced by an inhibitory GABAergic signal. **(B)** The signal at 10 Hz exercises an inhibitory drive silencing the firing of the pyramidal cells in a phasic manner. The magnitude of the inhibition is modulated by an arbitrary slower signal. **(C)** A raster diagram showing the firing of 500 pyramidal cells. The bouts of inhibition serve to silence the firing. In periods when the inhibition is low, the firing will persist (e.g., 0.4–0.6 s). **(D)** The summed firing of all the pyramidal cells (summed using a $\Delta t$ = 1 ms sliding time window). **(E)** Each spike was convolved with an alpha-function ($\alpha(t) = (t/\tau)\, e^{1-(t/\tau)}$, where $\tau$ = 5 ms) in order to approximate the post-synaptic current. These post-synaptic currents were then averaged over the pyramidal neurons. The "MEG signal" (arbitrary units, arbitrary offset) is proportional to the total dendritic current. Note that the inhibition results in a rhythmically pulsed MEG signal with amplitude asymmetry.

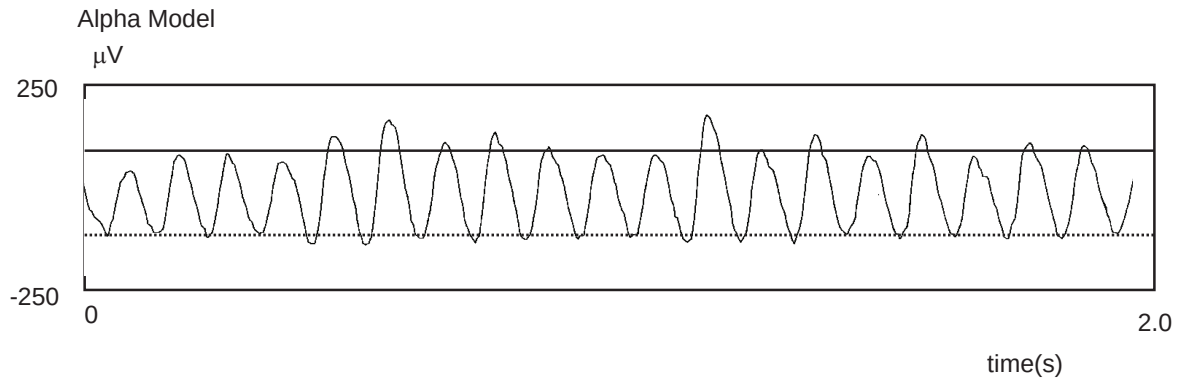



**FIGURE 10 | An alpha rhythm time-series displaying non-sinusoidal properties.** Prior to the discovery of its link to the formation of evoked responses, amplitude asymmetry had been observed in models designed to produce alpha activity. Here is an example adapted from Stam et al. (1999a).

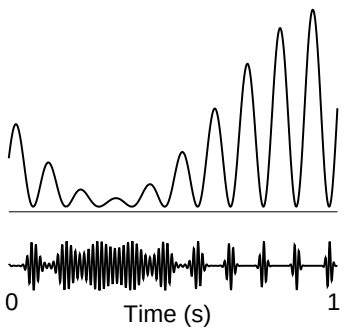



**FIGURE 11 | Rhythmic pulsing and it implications for processing – a hypothesis.** The gamma burst (e.g., involved in visual processing) can only occur when the alpha signal is low enough, e.g., at the troughs. Thus, the periods of gamma activity become briefer with stronger alpha. When the alpha activity is sufficiently weak, gamma can occur during the whole cycle. This scheme is consistent with the notion of rhythmic pulsing. Adapted from Osipova et al. (2008).

## OPEN QUESTIONS AND CAVEATS FOR FUTURE RESEARCH

### RELATIONSHIP BETWEEN OSCILLATIONS AND BOLD

Understanding the relationship between the brain's ongoing oscillations and evoked activity is quite important to the field of human electrophysiology since it is bound to lead to a fundamentally better understanding of how the signals measured by MEG/EEG link to cognition. Moreover understanding this relationship can extend beyond the realm of EEG and MEG research because evoked responses provide a critical link between the hemodynamic response measured by the fMRI and the underlying temporal dynamics of neuronal activity. There are now a number of studies that have correlated the amplitude of alpha activity with the BOLD signal (Goldman et al., 2002; Goncalves et al., 2006; Laufs et al., 2006; Ritter et al., 2009; Scheeringa et al., 2009; Yuan et al., 2010). However, the relationships between the pre-stimulus phase and amplitude of the ongoing alpha activity and the stimulus evoked BOLD have thus far not been investigated. If a relationship between pre-stimulus alpha phase and BOLD is demonstrated it would support the phasic aspect of the alpha oscillations and demonstrate that alpha activity has direct consequences for neuronal excitability. Moreover, the combination of brain stimulation by transcranial magnetic stimulation (TMS) with EEG can provide real-time information on the phasic aspects of cortical reactivity (Thut and Miniussi, 2009; Miniussi and Thut, 2010). Indeed, recent studies have found that the phase of the spinal beta rhythm in which the input (a TMS pulse) arrives modulated the gain of this input (Maki and Ilmoniemi, 2010; van Elswijk et al., 2010).

### WHAT ABOUT OTHER RHYTHMS?

The mechanism behind amplitude asymmetry of oscillatory activity need not be specific to alpha oscillations. The unidirectional primary currents in pyramidal cells could also be responsible for asymmetry in other frequencies bands a well. There is a body of research pointing to delta oscillations serving a fundamental role in the active input selection at primary sensory cortex level (Lakatos et al., 2007, 2008; Schroeder and Lakatos, 2009). According to this view the phase of the delta rhythm serves as a master controller of neuronal excitability. The relationship between delta phase and alpha activity needs to be explored. For example delta activity might reflect a top-down regulations of the alpha activity. Thus the time course of delta active could be comparable to the ERPs/ERFs produced by the modulation of the amplitude asymmetric alpha activity. The late occurring slow evoked responses can be viewed as the link between electrophysiology and cognition. These responses have been found to index working memory capacity (Vogel et al., 2005), long-term memory encoding and recognition (Sanquist et al., 1980; Takashima et al., 2006; Rugg and Curran, 2007), action monitoring (Kilner et al., 2004), language comprehension (Kutas and Hillyard, 1980; Hagoort and Brown, 2000), response preparation (Walter et al., 1964), and novelty detection (Soltani and Knight, 2000). Yet the exact physiological mechanism for how these responses are generated is still unknown.

Future work is required in order to investigate if the principle of amplitude asymmetry and the generation of evoked responses can be generalized to frequency bands beyond the alpha range. Averaging epochs of amplitude asymmetric oscillations will result in an of evoked response with a shape that is similar to the time course of amplitude suppression of the oscillation. A number of

studies have now shown that modulations in oscillatory activity at various frequency bands to occur at the same time window as cognitive evoked responses such as the Dm, P300, N400, and P600 effect (Klimesch et al., 1996; Mazaheri and Picton, 2005; Davidson and Indefrey, 2007). It would of interest to see if the mechanism of amplitude asymmetry plays a role in the formation of these responses.

If the mechanism underlying these responses is revealed to be explained by amplitude asymmetry of ongoing oscillations, it would in no way discount the value of conventional ERP/ERF analysis. Rather this would serve to take the human electrophysiology a very significant step further into linking the oscillatory firing of neuronal populations to human cognition.

## ACKNOWLEDGMENTS

This research was supported by The Netherlands Organization for Scientific Research (NWO): Rubicon Grant for AM, Innovational Research Incentive Schemes, VICI grant; ALW Open Competition, grant number 817.02.010. We would like to thank Dr. Risa Sawaki for comments on an earlier version of the manuscript.

## REFERENCES

Basar, E., Basar-Eroglu, C., Karakas, S., and Schurmann, M. (2001). Gamma, alpha, delta, and theta oscillations govern cognitive processes. *Int. J. Psychophysiol.* 39, 241–248.

Bastiaansen, M., and Hagoort, P. (2006). Oscillatory neuronal dynamics during language comprehension. *Prog. Brain Res.* 159, 179–196.

Becker, R., Ritter, P., and Villringer, A. (2008). Influence of ongoing alpha rhythm on the visual evoked potential. *Neuroimage* 39, 707–716.

Bishop, G. H. (1933). Cyclic change in the excitability of the optic pathway of the rabbit. *Am. J. Physiol.* 103, 213–224.

Busch, N. A., Dubois, J., and VanRullen, R. (2009). The phase of ongoing EEG oscillations predicts visual perception. *J. Neurosci.* 29, 7869–7876.

Callaway, E. 3rd, and Yeager, C. L. (1960). Relationship between reaction time and electroencephalographic alpha phase. *Science* 132, 1765–1766.

Coles, G. H., and Rugg, M. D. (1995). "ERPs: an introduction," in *Electrophysiology of Mind: Event Related Brain Potentials and Cognition*, eds M. D. Rugg and G. H. Coles (Oxford: Oxford University Press), 16–17.

Crunelli, V., and Leresche, N. (1991). A role for GABAB receptors in excitation and inhibition of thalamocortical cells. *Trends Neurosci.* 14, 16–21.

Davidson, D. J., and Indefrey, P. (2007). An inverse relation between event-related and time-frequency violation responses in sentence processing. *Brain Res.* 1158, 81–92.

de Lange, F. P., Jensen, O., Bauer, M., and Toni, I. (2008). Interactions between posterior gamma and frontal alpha/beta oscillations during imagined actions. *Front. Hum. Neurosci.* 2:7. doi: 10.3389/neuro.09.007.2008.

Fell, J., Dietl, T., Grunwald, T., Kurthen, M., Klaver, P., Trautner, P., Schaller, C., Elger, C. E., and Fernandez, G. (2004). Neural bases of cognitive ERPs: more than phase reset. *J. Cogn. Neurosci.* 16, 1595–1604.

Fell, J., Fernandez, G., Klaver, P., Elger, C. E., and Fries, P. (2003). Is synchronized neuronal gamma activity relevant for selective attention? *Brain Res. Brain Res. Rev.* 42, 265–272.

Fries, P., Nikolic, D., and Singer, W. (2007). The gamma cycle. *Trends Neurosci.* 30, 309–316.

Fukuda, K., Awh, E., and Vogel, E. K. (2010). Discrete capacity limits in visual working memory. *Curr. Opin. Neurobiol.* 20, 177–182.

Goldman, R. I., Stern, J. M., Engel, J. Jr., and and Cohen, M. S. (2002). Simultaneous EEG and fMRI of the alpha rhythm. *Neuroreport* 13, 2487–2492.

Goncalves, S. I., de Munck, J. C., Pouwels, P. J., Schoonhoven, R., Kuijer, J. P., Maurits, N. M., Hoogduin, J. M., Van Someren, E. J., Heethaar, R. M., and Lopes da Silva, F. H. (2006). Correlating the alpha rhythm to BOLD using simultaneous EEG/fMRI: inter-subject variability. *Neuroimage* 30, 203–213.

Gruber, W. R., Klimesch, W., Sauseng, P., and Doppelmayr, M. (2005). Alpha phase synchronization predicts P1 and N1 latency and amplitude size. *Cereb. Cortex* 15, 371–377.

Haegens, S., Osipova, D., Oostenveld, R., and Jensen, O. (2010). Somatosensory working memory performance in humans depends on both engagement and disengagement of regions in a distributed network. *Hum. Brain Mapp.* 31, 26–35.

Hagoort, P., and Brown, C. M. (2000). ERP effects of listening to speech compared to reading: the P600/SPS to syntactic violations in spoken sentences and rapid serial visual presentation. *Neuropsychologia* 38, 1531–1549.

Hamalainen, M. S., Hari, R., Ilmoniemi, R. J., Knuutila, J., and Lounasmaa. (1993). Magnetoencephalography. Theory, instrumentation and applications to tne noninvasive study of brain function. *Rev. Mod. Phys.* 65, 413–497.

Hughes, S. W., and Crunelli, V. (2005). Thalamic mechanisms of EEG alpha rhythms and their pathological implications. *Neuroscientist* 11, 357–372.

Jensen, O., Gelfand, J., Kounios, J., and Lisman, J. E. (2002). Oscillations in the alpha band (9–12 Hz) increase with memory load during retention in a short-term memory task. *Cereb. Cortex* 12, 877–882.

Jensen, O., Kaiser, J., and Lachaux, J. P. (2007). Human gamma-frequency oscillations associated with attention and memory. *Trends Neurosci.* 30, 317–324.

Jensen, O., van Dijk, H., and Mazaheri, A. (2010). Amplitude asymmetry as a mechanism for the generation of slow evoked responses. *Clin. Neurophysiol.* 121, 1148–1149.

Jensen, O., and Mazaheri, A. (2010). Shaping functional architecture by oscillatory alpha activity: gating by inhibition. *Front. Hum. Neurosci.* 4:186. doi:10.3389/fnhum.2010.00186

Jokisch, D., and Jensen, O. (2007). Modulation of gamma and alpha activity during a working memory task engaging the dorsal or ventral stream. *J. Neurosci.* 27, 3244–3251.

Jones, S. R., Pinto, D. J., Kaper, T. J., and Kopell, N. (2000). Alpha-frequency rhythms desynchronize over long cortical distances: a modeling study. *J. Comput. Neurosci.* 9, 271–291.

Kelly, S. P., Lalor, E. C., Reilly, R. B., and Foxe, J. J. (2006). Increases in alpha oscillatory power reflect an active retinotopic mechanism for distracter suppression during sustained visuospatial attention. *J. Neurophysiol.* 95, 3844–3851.

Kerlin, J. R., Shahin, A. J., and Miller, L. M. (2010). Attentional gain control of ongoing cortical speech representations in a "cocktail party". *J. Neurosci.* 30, 620–628.

Kilner, J. M., Vargas, C., Duval, S., Blakemore, S. J., and Sirigu, A. (2004). Motor activation prior to observation of a predicted movement. *Nat. Neurosci.* 7, 1299–1301.

Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain Res. Brain Res. Rev.* 29, 169–195.

Klimesch, W., Doppelmayr, M., Schimke, H., and Ripper, B. (1997). Theta synchronization and alpha desynchronization in a memory task. *Psychophysiology* 34, 169–176.

Klimesch, W., Doppelmayr, M., Schwaiger, J., Auinger, P., and Winkler, T. (1999). 'Paradoxical' alpha synchronization in a memory task. *Brain Res. Cogn. Brain Res.* 7, 493–501.

Klimesch, W., Sauseng, P., and Gruber, W. (2009). The functional relevance of phase reset: a comment to Risner et al. (2009): The visual evoked potential of surface alpha rhythm phase. *Neuroimage* 47, 5–7.

Klimesch, W., Sauseng, P., and Hanslmayr, S. (2007a). EEG alpha oscillations: the inhibition-timing hypothesis. *Brain Res. Rev.* 53, 63–88.

Klimesch, W., Sauseng, P., Hanslmayr, S., Gruber, W., and Freunberger, R. (2007b). Event-related phase reorganization may explain evoked neural dynamics. *Neurosci. Biobehav. Rev.* 31, 1003–1016.

Klimesch, W., Schack, B., Schabus, M., Doppelmayr, M., Gruber, W., and Sauseng, P. (2004). Phase-locked alpha and theta oscillations generate the P1-N1 complex and are related to memory performance. *Brain Res. Cogn. Brain Res.* 19, 302–316.

Klimesch, W., Schimke, H., Doppelmayr, M., Ripper, B., Schwaiger, J., and Pfurtscheller, G. (1996). Event-related desynchronization (ERD) and the Dm effect: does alpha desynchronization during encoding predict later recall performance? *Int. J. Psychophysiol.* 24, 47–60.

Kutas, M., and Hillyard, S. A. (1980). Reading senseless sentences: brain potentials reflect semantic incongruity. *Science* 207, 203–205.

Lachaux, J. P., George, N., Tallon-Baudry, C., Martinerie, J., Hugueville, L., Minotti, L., Kahane, P., and Renault, B. (2005). The many faces of the gamma band response to complex visual stimuli. *Neuroimage* 25, 491–501.

Lakatos, P., Chen, C. M., O'Connell, M. N., Mills, A., and Schroeder, C. E. (2007). Neuronal oscillations and multisensory interaction in primary auditory cortex. *Neuron* 53, 279–292.

Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I., and Schroeder, C. E. (2008). Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science* 320, 110–113.

Laufs, H., Holt, J. L., Elfont, R., Krams, M., Paul, J. S., Krakow, K., and Kleinschmidt, A. (2006). Where the BOLD signal goes when alpha EEG leaves. *Neuroimage* 31, 1408–1418.

Lehtela, L., Salmelin, R., and Hari, R. (1997). Evidence for reactive magnetic 10-Hz rhythm in the human auditory cortex. *Neurosci. Lett.* 222, 111–114.

Lopes da Silva, F. H., van Rotterdam, A., Barts, P., van Heusden, E., and Burr, W. (1976). Models of neuronal populations: the basic mechanisms of rhythmicity. *Prog. Brain Res.* 45, 281–308.

Lorincz, M. L., Crunelli, V., and Hughes, S. W. (2008). Cellular dynamics of cholinergically induced alpha (8–13 Hz) rhythms in sensory thalamic nuclei in vitro. *J. Neurosci.* 28, 660–671.

Lorincz, M. L., Kekesi, K. A., Juhasz, G., Crunelli, V., and Hughes, S. W. (2009). Temporal framing of thalamic relay-mode firing by phasic inhibition during the alpha rhythm. *Neuron* 63, 683–696.

Makeig, S., Delorme, A., Westerfield, M., Jung, T. P., Townsend, J., Courchesne, E., and Sejnowski, T. J. (2004). Electroencephalographic brain dynamics following manually responded visual targets. *PLoS Biol.* 2, e176. doi: 10.1371/journal.pbio.0020176.

Makeig, S., Westerfield, M., Jung, T. P., Enghoff, S., Townsend, J., Courchesne, E., and Sejnowski, T. J. (2002). Dynamic brain sources of visual evoked responses. *Science* 295, 690–694.

Maki, H., and Ilmoniemi, R. J. (2010). EEG oscillations and magnetically evoked motor potentials reflect motor system excitability in overlapping neuronal populations. *Clin. Neurophysiol.* 121, 492–501.

Makinen, V., Tiitinen, H., and May, P. (2005). Auditory event-related responses are generated independently of ongoing brain activity. *Neuroimage* 24, 961–968.

Mathewson, K. E., Gratton, G., Fabiani, M., Beck, D. M., and Ro, T. (2009). To see or not to see: prestimulus alpha phase predicts visual awareness. *J. Neurosci.* 29, 2725–2732.

Mazaheri, A., and Jensen, O. (2006). Posterior alpha activity is not phase-reset by visual stimuli. *Proc. Natl. Acad. Sci. U.S.A.* 103, 2948–2952.

Mazaheri, A., and Jensen, O. (2008). Asymmetric amplitude modulations of brain oscillations generate slow evoked responses. *J. Neurosci.* 28, 7781–7787.

Mazaheri, A., and Picton, T. W. (2005). EEG spectral dynamics during discrimination of auditory and visual targets. *Brain Res. Cogn. Brain Res.* 24, 81–96.

Medendorp, W. P., Kramer, G. F., Jensen, O., Oostenveld, R., Schoffelen, J. M., and Fries, P. (2007). Oscillatory activity in human parietal and occipital cortex shows hemispheric lateralization and memory effects in a delayed double-step saccade task. *Cereb Cortex* 17, 2364–2374.

Min, B. K., Busch, N. A., Debener, S., Kranczioch, C., Hanslmayr, S., Engel, A. K., and Herrmann, C. S. (2007). The best of both worlds: phase-reset of human EEG alpha activity and additive power contribute to ERP generation. *Int. J. Psychophysiol.* 65, 58–68.

Miniussi, C., and Thut, G. (2010). Combining TMS and EEG offers new prospects in cognitive neuroscience. *Brain Topogr.* 22, 249–256.

Murakami, S., and Okada, Y. (2006). Contributions of principal neocortical neurons to magnetoencephalography and electroencephalography signals. *J. Physiol. (Lond.)* 575(Pt 3), 925–936.

Nikulin, V. V., Linkenkaer-Hansen, K., Nolte, G., and Curio, G. (2010). Non-zero mean and asymmetry of neuronal oscillations have different implications for evoked responses. *Clin. Neurophysiol.* 121, 186–193.

Nikulin, V. V., Linkenkaer-Hansen, K., Nolte, G., Lemm, S., Müller K. R., Ilmoniemi R. J., and Curio, G. (2007). A novel mechanism for evoked responses in the human brain. *Eur. J. Neurosci.* 25, 3146–3154.

Osipova, D., Hermes, D., and Jensen, O. (2008). Gamma power is phase-locked to posterior alpha activity. *PLoS ONE* 3, e3990. doi: 10.1371/journal.pone.0003990.

Rihs, T. A., Michel, C. M., and Thut, G. (2007). Mechanisms of selective inhibition in visual spatial attention are indexed by alpha-band EEG synchronization. *Eur. J. Neurosci.* 25, 603–610.

Risner, M. L., Aura, C. J., Black, J. E., and Gawne, T. J. (2009). The visual evoked potential is independent of surface alpha rhythm phase. *Neuroimage* 45, 463–469.

Ritter, P., and Becker, R. (2009). Detecting alpha rhythm phase reset by phase sorting: caveats to consider. *Neuroimage* 47, 1–4.

Ritter, P., Moosmann, M., and Villringer, A. (2009). Rolandic alpha and beta EEG rhythms' strengths are inversely related to fMRI-BOLD signal in primary somatosensory and motor cortex. *Hum. Brain Mapp.* 30, 1168–1187.

Romei, V., Brodbeck, V., Michel, C., Amedi, A., Pascual-Leone, A., and Thut, G. (2007). Spontaneous fluctuations in posterior {alpha}-band EEG activity reflect variability in excitability of human visual areas. *Cereb. Cortex.* 18, 2010–2018.

Romei, V., Rihs, T., Brodbeck, V., and Thut, G. (2008). Resting electroencephalogram alpha-power over posterior sites indexes baseline visual cortex excitability. *Neuroreport* 19, 203–208.

Rugg, M. D., and Curran, T. (2007). Event-related potentials and recognition memory. *Trends Cogn. Sci. (Regul. Ed.)* 11, 251–257.

Salenius, S., Schnitzler, A., Salmelin, R., Jousmaki, V., and Hari, R. (1997). Modulation of human cortical rolandic rhythms during natural sensorimotor tasks. *Neuroimage* 5, 221–228.

Sanquist, T. F., Rohrbaugh, J. W., Syndulko, K., and Lindsley, D. B. (1980). Electrocortical signs of levels of processing: perceptual analysis and recognition memory. *Psychophysiology* 17, 568–576.

Sauseng, P., Klimesch, W., Freunberger, R., Pecherstorfer, T., Hanslmayr, S., and Doppelmayr, M. (2006). Relevance of EEG alpha and theta oscillations during task switching. *Exp. Brain Res.* 170, 295–301.

Sauseng, P., Klimesch, W., Heise, K. F., Gruber, W. R., Holz, E., Karim, A. A., Glennon, M., Gerloff, C., Birbaumer, N., and Hummel, F. C. (2009). Brain oscillatory substrates of visual short-term memory capacity. *Curr. Biol.* 19, 1846–1852.

Scheeringa, R., Petersson, K. M., Oostenveld, R., Norris, D. G., Hagoort, P., and Bastiaansen, M. C. (2009). Trial-by-trial coupling between EEG and BOLD identifies networks related to alpha and theta EEG power increases during working memory maintenance. *Neuroimage* 44, 1224–1238.

Schroeder, C. E., and Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. *Trends Neurosci.* 32, 9–18.

Shah, A. S., Bressler, S. L., Knuth, K. H., Ding, M., Mehta, A. D., Ulbert, I., and Schroeder, C. E. (2004). Neural dynamics and the fundamental mechanisms of event-related brain potentials. *Cereb. Cortex* 14, 476–483.

Shahin, A. J., Picton, T. W., and Miller, L. M. (2009). Brain oscillations during semantic evaluation of speech. *Brain Cogn.* 70, 259–266.

Soltani, M., and Knight, R. T. (2000). Neural origins of the P300. *Crit. Rev. Neurobiol.* 14, 199–224.

Stam, C. J., Pijn, J. P., Suffczynski, P., and Lopes da Silva, F. H. (1999a). Dynamics of the human alpha rhythm: evidence for non-linearity? *Clin. Neurophysiol.* 110, 1801–1813.

Stam, C. J., Vliegen, J. H., and Nicolai, J. (1999b). Investigation of the dynamics underlying periodic complexes in the EEG. *Biol. Cybern.* 80, 57–69.

Takashima, A., Jensen, O., Oostenveld, R., Maris, E., van de Coevering, M., and Fernandez, G. (2006). Successful declarative memory formation is associated with ongoing activity during encoding in a distributed neocortical network related to working memory: a magnetoencephalography study. *Neuroscience* 139, 291–297.

Thut, G., and Miniussi, C. (2009). New insights into rhythmic brain activity from TMS-EEG studies. *Trends Cogn. Sci. (Regul. Ed.)* 13, 182–189.

Thut, G., Nietzel, A., Brandt, S. A., and Pascual-Leone, A. (2006). Alpha-band electroencephalographic activity over occipital cortex indexes visuospatial attention bias and predicts visual target detection. *J. Neurosci.* 26, 9494–9502.

Thut, G., Theoret, H., Pfennig, A., Ives, J., Kampmann, F., Northoff, G., and Pascual-Leone, A. (2003). Differential effects of low-frequency rTMS at the occipital pole on visual-induced alpha desynchronization and visual-evoked potentials. *Neuroimage* 18, 334–347.

Tuladhar, A. M., ter Huurne, N., Schoffelen, J. M., Maris, E., Oostenveld, R., and Jensen, O. (2007). Parieto-occipital sources account for the increase in alpha activity with working memory load. *Hum. Brain Mapp.* 28, 785–792.

van Dijk, H., Schoffelen, J. M., Oostenveld, R., and Jensen, O. (2008). Prestimulus oscillatory activity in the alpha band predicts visual discrimination ability. *J. Neurosci.* 28, 1816–1823.

van Dijk, H., Nieuwenhuis, I. L., and Jensen, O. (2010a). Left temporal alpha band activity increases during working memory retention of pitches. *Eur. J. Neurosci.* 31, 1701–1707.

van Dijk, H., van der Werf, J., Mazaheri, A., Medendorp, W. P., and Jensen, O. (2010b). Modulations in oscillatory activity with amplitude asymmetry can produce cognitively relevant event-related responses. *Proc. Natl. Acad. Sci. U.S.A.* 107, 900–905.

van Elswijk, G., Maij, F., Schoffelen, J. M., Overeem, S., Stegeman, D. F., and Fries, P. (2010). Corticospinal beta-band synchronization entails rhythmic gain modulation. *J. Neurosci.* 30, 4481–4488.

VanRullen, R., and Koch, C. (2003). Is perception discrete or continuous? *Trends Cogn. Sci.* 7, 207–213.

Varela, F. J., Toro, A., John, E. R., and Schwartz, E. L. (1981). Perceptual framing and cortical alpha rhythm. *Neuropsychologia* 19, 675–686.

Vogel, E. K., and Machizawa, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature* 428, 748–751.

Vogel, E. K., McCollough, A. W., and Machizawa, M. G. (2005). Neural measures reveal individual differences in controlling access to working memory. *Nature* 438, 500–503.

Walter, W. G., Cooper, R., Aldridge, V. J., McCallum, W. C., and Winter, A. L. (1964). Contingent negative variation: an electric sign of sensorimotor association and expectancy in the human brain. *Nature* 203, 380–384.

Womelsdorf, T., Fries, P., Mitra, P. P., and Desimone, R. (2006). Gamma-band synchronization in visual cortex predicts speed of change detection. *Nature* 439, 733–736.

Worden, M. S., Foxe, J. J., Wang, N., and Simpson, G. V. (2000). Anticipatory biasing of visuospatial attention indexed by retinotopically specific alpha-band electroencephalography increases over occipital cortex. *J. Neurosci.* 20, RC63.

Wu, J., and Okada, Y. C. (1999). Roles of a potassium afterhyperpolarization current in generating neuromagnetic fields and field potentials in longitudinal CA3 slices of the guinea-pig. *Clin. Neurophysiol.* 110, 1858–1867.

Yuan, H., Liu, T., Szarkowski, R., Rios, C., Ashe, J., and He, B. (2010). Negative covariation between task-related responses in alpha/beta-band activity and BOLD in human sensorimotor cortex: an EEG and fMRI study of motor imagery and movements. *Neuroimage* 49, 2596–2606.

Zhang, Y., Wang, X., Bressler, S. L., Chen, Y., and Ding, M. (2008). Prestimulus cortical activity is correlated with speed of visuomotor processing. *J. Cogn. Neurosci.* 20, 1915–1925.

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

*Received: 21 May 2010; paper pending published: 12 June 2010; accepted: 25 August 2010; published online: 14 October 2010.*
*Citation: Mazaheri A and Jensen O (2010) Rhythmic pulsing: linking ongoing brain activity with evoked responses. Front. Hum. Neurosci. **4**:177. doi: 10.3389/fnhum.2010.00177*

frontiers in
HUMAN NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 19 October 2010
doi: 10.3389/fnhum.2010.00173

# What have we learned from "perturbing" the human cortical motor system with transcranial magnetic stimulation?

***Philippe A. Chouinard[1]* and Tomáš Paus[2,3]***

[1] *Department of Psychology, Centre for Brain and Mind, University of Western Ontario, London, ON, Canada*
[2] *Department of Psychology, The Rotman Research Institute, University of Toronto, Toronto, ON, Canada*
[3] *Department of Psychiatry, The Rotman Research Institute, University of Toronto, Toronto, ON, Canada*

***Edited by:***
*Michael X. Cohen, University of Amsterdam, Netherlands*

***Reviewed by:***
*Hartwige Siebner, Danish Research Centre for Magnetic Resonance, Denmark*
*Katsuyuki Sakai, The University of Tokyo, Japan*

****Correspondence:***
*Philippe A. Chouinard, Centre for Brain and Mind, Department of Psychology, University of Western Ontario, 1151 Richmond Street, London, ON N6A 5C2, Canada.*
*e-mail: pchouin@uwo.ca*

The purpose of this paper is twofold. First, we will review different approaches that one can use with transcranial magnetic stimulation (TMS) to study both its effects on motor behavior and on neural connections in the human brain. Second, we will present evidence obtained in TMS-based studies showing that the dorsal premotor area (PMd), the ventral premotor area (PMv), the supplementary motor area (SMA), and the pre-supplementary motor area (pre-SMA) each have different roles to play in motor behavior. We highlight the importance of the PMd in response selection based on arbitrary cues and in the control of arm movements, the PMv in grasping and in the discrimination of bodily actions, the SMA in movement sequencing and in bimanual coordination, and the pre-SMA in cognitive control. We will also discuss ways in which TMS can be used to chart "true" cerebral reorganization in clinical populations and how TMS might be used as a therapeutic tool to facilitate motor recovery after stroke. We will end our review by discussing some of the methodological challenges and future directions for using this tool in basic and clinical neuroscience.

**Keywords: transcranial magnetic stimulation, motor system, functional connectivity, effective connectivity, premotor area, supplementary motor area, stroke recovery, functional neuroimaging**

## INTRODUCTION

Hughlings Jackson proposed that the central nervous system was composed of a number of hierarchical levels: each level containing a complete set of representations of the next lower level that enables it to exert influence on motor behavior (see Hughlings Jackson, 1958). This hierarchical organization of the motor system was challenged in the 1990s with the emergence of anatomical studies in the monkey that demonstrated a number of cortical areas other than the primary motor cortex (M1) with direct projections to the spinal cord (Dum and Strick, 1991; He et al., 1993, 1995). We now know that several areas in the frontal lobe have the anatomical substrate to influence motor output both through connections with M1 and through direct projections to the spinal cord. These non-primary motor areas include the premotor, the supplementary motor, and the cingulate motor areas. These areas can be further divided into caudal and rostral subdivisions based on the degree to which they can influence motor output (Picard and Strick, 1996, 2001). Caudal subdivisions for each of these areas exert a much stronger influence on motor output than their rostral subdivisions (Barbas and Pandya, 1987). The latter exert little or no direct influence on motor output. The presence of analogous areas in the human brain has been proposed based on a series of meta-analyses carried out to characterize functional activation during motor tasks (Picard and Strick, 1996, 2001). These areas are shown in **Figure 1**. They include the dorsal premotor area (PMd), the ventral premotor area (PMv), the supplementary motor area (SMA), the pre-SMA, and the cingulate motor areas (CMAs). The purpose of this review is twofold. First, we will describe the different approaches that one can use with transcranial magnetic stimulation (TMS) to study both its effects on motor behavior and neural connections in the human brain. Second, we will present the evidence obtained in TMS-based studies showing that PMd, PMv, SMA, and pre-SMA each have different roles to play in motor behavior. We will also discuss ways in which TMS can be used to chart "true" cerebral reorganization in clinical populations and how TMS might be used as a therapeutic tool to help motor recovery after stroke. We will end our review by discussing some of the future avenues for using this tool in basic and clinical neuroscience.

## WHAT IS TMS?

Transcranial magnetic stimulation is a "perturbation" technique. It perturbs neural activity in space and time by inducing brief electrical currents in a restricted region of the cerebral cortex. A brief current passes through a stimulating coil, which is placed over the person's scalp, that then induces a rapid rise of magnetic field, and this transient field in turn induces electrical current in the underlying brain tissue. Barker et al. (1985) performed the first TMS experiment and the technique has since acquired importance as a non-invasive method for examining motor, perceptual, and cognitive processes in the human brain. Transcranial direct current stimulation (tDCS) is another perturbation technique that is used to study brain processes. Unlike TMS, tDCS is used to modulate the activity of neurons by applying weak electrical currents through an electrode placed on the scalp (for review, see Been et al., 2007). The underlying principle of both techniques is simple. Using a "perturb-and-measure" approach, one can perturb one brain region and measure the consequences that this manipulation has on either behavior or brain activity (Paus, 2005). As perturbation techniques, TMS and tDCS are designed to study consequences of modulating brain activity and, as such, to inject a certain level

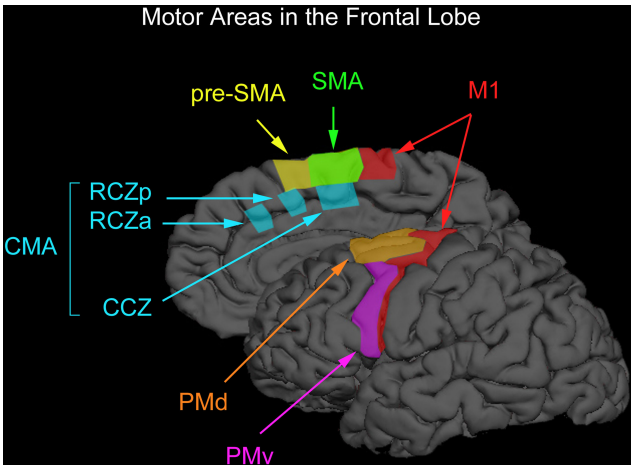



**FIGURE 1 | Motor areas in the frontal lobe.** The premotor cortex on the lateral surface of the brain can be divided into the dorsal and ventral premotor areas (PMd and PMv) and the supplementary motor cortex on the medial wall of the brain can be divided into the supplementary motor and pre-supplementary motor areas (SMA and pre-SMA). Premotor cortex below the superior frontal sulcus is typically considered PMv whereas premotor cortex above this anatomical landmark is typically considered PMd. The vertical anterior-commissural line is often used to denote the boundary between SMA and pre-SMA. One can further divide PMd according to a rostral subdivision located along the superior frontal gyrus and a caudal subdivision located along the precentral gyrus. However, one cannot dissociate these two subdivisions with TMS easily and we therefore do not discuss them separately. There also exists two cingulate motor areas (RCZa and RZp) anterior to the vertical anterior-commissural line and one cingulate motor area (CCZ) posterior to the vertical anterior-commissural line. This parcellation of non-primary motor areas in the human was proposed by Picard and Strick (1996, 2001). We have arbitrarily drawn boundaries on a surface-rendered cortical surface loosely based on definitions proposed by Picard and Strick (1996, 2001).

of causality in investigations of brain-behavior relationships. In this manner, they are fundamentally different than functional neuroimaging. Both functional magnetic resonance imaging (fMRI) and positron emission tomography (PET) are measurement techniques that are designed to answer the following question: what regions in the brain are engaged during a behavior? We will focus our review on TMS studies.

Although TMS enables one to make conclusions about the necessity of a particular brain region for a given behavior (i.e., stimulated region 'X' is necessary), it does not enable one to conclude that the region is "sufficient" for this behavior (i.e., whether or not other brain regions may also be necessary). To illustrate this point, consider the following analogy that is sometimes cited in the literature (Huettel et al., 2009). Damage to one part of a radio, such as the speakers, the tuner, or the power switch, will result in an inability to play music. Because damage to any of these parts will cause the radio to stop playing music, one should not go about damaging one of these parts and then claim that this manipulation knocked out the "music-playing" region in the radio. The same logic also applies to TMS. Nonetheless, because performance on a task often relies on multiple brain regions, TMS can provide opportunities to examine how different brain regions might interact with each other to accomplish a behavior – this is known as functional connectivity. In this paper, we will review the various ways in which TMS can be used to measure its consequences on behavior and how the technique can also be used to examine both functional and effective (i.e., the influence that one brain region exerts over another) connectivity.

## HOW IS TMS APPLIED?

There are a number of ways that one can perturb neural activity with TMS. One can use a small number of pulses applied at a high frequency (5 Hz or more) to disrupt transiently neural activity in a brain region "on-line" during task performance. Using this same approach, one can further reduce the number of pulses delivered and/or increase the frequency with which the pulses are delivered to examine when in time a stimulated brain region is engaged during a given behavior. Similar "event-related" TMS can also be performed using single-pulse stimulation. One important disadvantage of the on-line approach, however, is that the TMS creates acoustic artifacts and tactile sensations on the scalp that can interfere with task performance by distracting the participant. Although these effects are usually controlled for with additional experiments, some researchers will opt to use "off-line" TMS. For off-line TMS, one can use certain stimulation protocols to test the effects of stimulation on task performance after the stimulation has been applied. This can be accomplished with either one continuous train of low-frequency (~1 Hz) stimulation or with multiple bursts of high-frequency (~50 Hz) stimulation that are spaced in time. The latter is known as theta-burst stimulation (TBS; Huang et al., 2005, 2008). The precise physiological mechanisms that underlie TMS-induced perturbations for each of these different applications of TMS are not completely understood (for review on putative mechanisms, see Ridding and Rothwell, 2007; Bestmann, 2008; Miniussi et al., 2009; Siebner et al., 2009). We should point out that TBS is relatively new. At the present time, the safety of TBS is not completely understood and there are no recommended guidelines on how to administer this type of stimulation safely (for the latest guidelines on TMS safety, see Rossi et al., 2009). Furthermore, we are not aware of any studies that have compared directly, in the same group of participants, the efficacy of TBS in disrupting task performance against other forms of brain stimulation.

Targeting specific cortical regions with TMS can be achieved with a variety of approaches – some of which work better than others. Using power analysis, Sack et al. (2009) revealed that out of the four most commonly used methods for targeting cortical regions with TMS, functional localization was the method that resulted in the most powerful effects. This approach consists of localizing targets with fMRI and then using an MRI-guided navigation system to guide the TMS coil over the functionally defined regions. Alternative approaches consist of guiding the TMS coil to a location defined by anatomical landmarks on an individual's MRI (the second most effective approach tested by Sack et al., 2009), guiding the TMS coil to a location defined in Talairach coordinates (the third most effective approach tested by Sack et al., 2009, which has less spatial precision compared with the previous two methods given variability in brain anatomy), and guiding the TMS coil relative to positions in the 10-20 EEG system (not surprisingly, the least effective approach tested by Sack et al., 2009, which has little spatial precision). On the topic of spatial precision, one should also

consider the spread of current induced by TMS. Spread of current can be minimized using a figure-of-eight coil, which has become standard practice, and reducing the intensity of TMS. All work that we will cover here have used a figure-of-eight coil and stimulated non-primary motor areas at reduced levels of stimulation so as to avoid encroaching on M1 and other adjacent cortical structures. Figure-of-eight coils are commercially available in different sizes so it might be worthwhile to invest in a smaller figure-of-eight coil to allow for more focal stimulation.

Despite these recommendations, TMS does not rival fMRI (or PET) in its spatial resolution. In fact, the gap between the two is widening. High-resolution fMRI (i.e., voxels smaller than 2 mm in isotropic size) is becoming increasingly more common. In comparison, the spread of current induced by TMS with a standard figure-of-eight coil is ~1.5 cm at motor-threshold intensities (Thielscher and Kammer, 2004). Nonetheless, the two techniques can be used in a complementary manner to make up for the shortcomings of the other. For example, the temporal response of the blood supply underlying fMRI (~5 s) is much slower than the electrical signals that define neuronal communication. In contrast, TMS offers far better temporal resolution (~1 ms). Also, fMRI is a measurement technique and, therefore, it can only measure correlates of behavior and does not allow one to make inferences about causality. In contrast, TMS is a perturbation technique and it can therefore be used to study causal relationships (Paus, 2005). The two approaches are therefore useful for providing converging evidence to argue for or against any functional attributions inferred by the other.

To illustrate the complementary roles of the perturbation approach using TMS and the activation approach using functional neuroimaging, let us consider some of the first author's fMRI (Chouinard et al., 2008) and TMS (Chouinard et al., 2009a) work on object identification. FMRI usually reveals that the presentation of objects in central vision engages extrastriate visual areas in the two hemispheres – "activation" is typically seen bilaterally in a ventral-stream area known as the lateral-occipital complex (LOC), which is thought to be important for analyzing the form of objects (Malach et al., 1995; Kanwisher et al., 1996). But it has been this author's experience that TMS applied to LOC in either hemisphere (as defined by fMRI) has little effect in disrupting object identification when these objects are presented in central vision (unpublished data). Yet, when the same stimulation is applied when objects are presented in the contralateral but not in the ipsilateral hemifield, TMS applied over LOC in either hemisphere can produce deficits in identifying objects (Chouinard et al., 2009a). This finding is hardly surprising if one considers that extrastriate visual areas, such as LOC, are retinotopically organized (Op De Beeck and Vogels, 2000). It also suggests, however, that the unstimulated LOC could perhaps stand in for the stimulated LOC when objects are presented in central vision. This idea fits well with the notion that ventral-stream damage in the two hemispheres must occur to produce visual-form agnosia (Farah, 2004).

## THE NON-PRIMARY MOTOR AREAS IN BEHAVIOR

**Tables 1–4** provide a summary of all TMS experiments that examined the behavioral consequences of stimulating a non-primary motor area ($N = 50$). Here, we will focus our discussion on the most consistent findings. Our review highlights the importance of PMd in response selection based on arbitrary cues and in the control of arm movements, the importance of PMv in grasping and in the discrimination of bodily actions, the importance of SMA in movement sequencing and in bimanual coordination, and the importance of pre-SMA in cognitive control processes.

### PMd IN MOTOR BEHAVIOR

People often select motor responses according to arbitrary rules. For example, our movements while driving a car can be instructed by color cues that we see on traffic lights. These associations are learned and are thought to involve neural processes that differ from the ones that are used in standard visuomotor transformations (Chouinard and Goodale, 2009). As it turns out, a number of studies show that response selection to arbitrary visual cues can be disrupted when TMS is applied to the left PMd irrespectively of whether participants have to make selections using button pressing (Schluter et al., 1998, 1999; Rushworth et al., 2002; O'Shea et al., 2007a,b), hand postures (Taubert et al., 2010), or fingertip forces during object lifting (Chouinard et al., 2005). Response selection to auditory-presented arbitrary cues can also be disrupted with TMS to the left PMd (Mochizuki et al., 2005). Taken together, the left PMd seems to be necessary for selecting responses to arbitrary cues irrespectively of whether these cues are visual or auditory and the type of actions that are made in response to these cues.

The TMS literature also highlights the importance of PMd in arm movements during pointing tasks (van Donkelaar et al., 2002; Busan et al., 2009) including a role in the visual online control of these movements (Lee and van Donkelaar, 2006). This is not surprising given that PMd in the macaque monkey is reciprocally connected to areas MIP and V6A in the parietal lobe, which together form a frontal-parietal circuit important for the visual guidance of arm movement trajectories (Colby and Duhamel, 1991; Galletti et al., 1996; Matelli and Luppino, 2000). Note that arm movements are important not only for reaching but also for lifting objects; this is because we tend to lift objects using proximal muscles of the arm. Davare et al. (2006) demonstrated that TMS to PMd disrupts the generation of proximal arm movements when people lift objects. Namely, electromyography of the arm revealed that participants took a longer time to move the arm when TMS was applied to PMd during the lifting phase of the movement as compared with the same stimulation applied earlier in time. Taken together, these studies are consistent with the notion that PMd is important for both the online control and the execution of arm movements (Kalaska et al., 1997).

### PMv IN MOTOR BEHAVIOR

Research in the monkey shows that PMv makes an important contribution in object grasping (Jeannerod et al., 1995). The first TMS study to confirm this in humans was performed by Davare et al. (2006); they demonstrated that participants took a longer time to position their fingers on a task object when TMS was applied to PMv. Measurements acquired with force transducers mounted inside the task object revealed that this same stimulation also caused participants to place their index finger and thumb less accurately on the object. As it turns out, many PMv neurons that discharge while monkeys grasp objects also discharge while monkeys see either another monkey or human perform the same action. This class of

**Table 1 | TMS studies of PMd.**

| Reference | TMS used | Behavioral deficits induced |
|---|---|---|
| Brighina et al. (2002) | HF-RTMS: 25 Hz, 10 stimuli, 115% rMT | Visuospatial perception |
| Busan et al. (2009) | SP-TMS: 110% rMT | Response preparation, visually-guided reaching |
| Chouinard et al. (2005) | LF-RTMS: 1 Hz, 900 stimuli, 90% rMT | Response selection to arbitrary visual cues, object lifting |
| Davare et al. (2006) | HF-RTMS: 10 Hz, 6 stimuli, 120% rMT | Lift production on objects |
| Davare et al. (2006) | DP-TMS: 120% rMT, ISI 5 ms | Lift production on objects |
| Giovannelli et al. (2006) | LF-RTMS: 1 Hz, 900 stimuli, 115% rMT | Inhibition of mirror movements, finger movements |
| Herwig et al. (2003) | HF-RTMS: 15 Hz, 45 stimuli, 110% rMT | Verbal working memory |
| Koski et al. (2005) | SP-TMS: 110% rMT | Inhibition of automatic stimulus-response associations |
| Lee and Van Donkelaar (2006) | SP-TMS: 110% rMT | On-line control of movements, visually-guided reaching |
| Liuzzi et al. (2010) | DS PMd-M1 TMS: c-TMS varied, t-TMS MEPs of 1 mV at rest, ISI 10 ms | Response preparation, button pressing |
| Mochizuki et al. (2005) | TBS: 50 Hz, 3 stimuli, every 200 ms for 20 s, 90% aMT | Response selection to arbitrary sound cues, button pressing |
| Mochizuki et al. (2005) | DP-TMS: 120% rMT, ISI 25 ms | Response selection to arbitrary sound cues, button pressing |
| O'Shea et al. (2007a) | LF-RTMS: 1 Hz, 900 stimuli, 90% aMT | Response selection to arbitrary visual cues, button pressing |
| O'Shea et al. (2007b) | DS PMd-M1 TMS: c-TMS 110% rMT, t-TMS MEPs of 1 mV at rest, ISI 8 ms | Response selection to arbitrary visual cues, button pressing |
| Pollok et al. (2008) | LF-RTMS: 1 Hz, 1200 stimuli, 90% aMT | Movement timing, finger tapping to sound cues |
| Praamstra et al. (1999) | HF-RTMS: 20 Hz, 4 stimuli, 90% rMT | Inhibition of automatic stimulus-response associations |
| Rushworth et al. (2002) | HF-RTMS: 5 Hz, 4 stimuli, 105% rMT (foot) | Response selection to arbitrary visual cues, button pressing |
| Schlaghecken et al. (2003) | LF-RTMS: 1 Hz, 1200 stimuli, 80% aMT | Response selection to directional visual cues, button pressing |
| Schluter et al. (1998) | SP-TMS: 100% rMT | Response selection to arbitrary visual cues, button pressing |
| Schluter et al. (1999) | SP-TMS: 100% rMT | Response preparation, button pressing |
| Schluter et al. (1999) | SP-TMS: 100% rMT | Response selection to arbitrary visual cues, button pressing |
| Tanaka et al. (2005) | LF-RTMS: 0.9 Hz, 420 stimuli, 70% stimulator output | Udating spatial information |
| Taubert et al. (2010) | DP-TMS: 100% rMT, ISI 100 ms | Response selection to arbitrary visual cues, hand postures |
| Van den Berg et al. (in press) | DP-TMS: 110% rMT | Bimanual coordination, finger tapping |
| Van Donkelaar et al. (2002) | SP-TMS: 110% rMT | Eye-hand coordination, visually-guided reaching |

*aMT, active motor threshold; c-TMS, conditioning pulse TMS; DS, dual-site TMS; DP-TMS, double-pulse TMS; HF-TMS, High-frequency TMS; ISI, inter-stimulus interval; LF-TMS, Low-frequency TMS; MEPs, motor evoked potentials; rMT, resting motor threshold; SP-TMS, single-pulse TMS; t-TMS, test pulse TMS.*

neurons is called mirror neurons (Rizzolatti and Luppino, 2001) and are thought by some to be important for understanding the meaning of actions (Umilta et al., 2001). In support of this idea, TMS studies show that PMv plays an important role in the discrimination of bodily actions (Urgesi et al., 2007a,b; Candidi et al., 2008). For example, Candidi et al. (2008) stimulated PMv while participants had to decide which of two images presented in either possible or impossible hand configurations corresponded to the same action that was presented to them in an earlier probe image. Stimulation of PMv impaired performance on this discrimination task but only when the possible hand postures were presented. Thus, PMv's involvement in the discrimination of bodily actions seems to be specific to meaningful actions.

## SMA IN MOTOR BEHAVIOR

In the early "activation" studies carried out with PET, it was reported that SMA is strongly activated when people imagine themselves performing complex sequences of finger movements (Roland et al., 1980). This observation, along with other lines of evidence, lead to theories that motor areas in the medial wall of the cortex are important for the internal generation of complex movements (Goldberg, 1985). But more recent work, including that carried out with TMS, highlights the importance of SMA in movement sequencing and in bimanual coordination. Gerloff et al. (1997) showed that TMS applied over SMA interferes with the generation of complex sequences of finger movements but not with the generation of simple repetitive finger movements. In a different TMS study, Verwey et al. (2002) examined whether or not SMA played a role in initiating "chunks" of movements in a complex sequence of finger movements. Behavioral research shows that people bin complex movement sequences into motor chunks (Verwey, 1996, 1999) in the same way that people bin phone numbers into shorter series. Contrary to the prediction, however, stimulating SMA did not disrupt the initiation of chunks but instead disrupted the execution of each finger movement. As we will see later, this finding differs from what was obtained in a different study that stimulated the pre-SMA (Kennerley et al., 2004). Other studies show that TMS applied over SMA can also disrupt performance on bimanual tasks as a function of task complexity (Obhi et al., 2002; Serrien et al., 2002; Steyvers et al., 2003). For example, both Serrien et al. (2002) and Steyvers et al. (2003) showed that TMS applied over SMA disrupts bimanual movements of the index finger

**Table 2 | TMS studies of PMv.**

| Reference | TMS used | Behavioral deficits induced |
|---|---|---|
| Buch et al. (2010) | DS PMv-M1 TMS: c-TMS 110% rMT, t-TMS MEPs of 1 mV at rest, ISI 8 ms | Task switching, grasp production |
| Candidi et al. (2008) | DP-TMS: 120% rMT, ISI 100 ms | Discrimination of bodily postures |
| Cattaneo et al. (2010) | SP-TMS: 65% of stimulator output | Semantic processing specific to tools |
| Dafotakis et al. (2008) | SP-TMS: 100% rMT | Sensorimotor memory, object lifting |
| Davare et al. (2006) | HF-RTMS: 10 Hz, 6 stimuli, 120% rMT | Grasp production on objects |
| Davare et al. (2006) | DP-TMS: 120% rMT, ISI 5 ms | Grasp production on objects |
| Davare et al. (2008) | DS PMv-M1 TMS: c-TMS 80% rMT, t-TMS 120% rMT, ISI 1–15 ms | Grasp production on objects |
| Davare et al. (2009) | DS PMv-M1 TMS: c-TMS 80% rMT, t-TMS 120% rMT, ISI 1–15 ms | Grasp production on objects |
| Davare et al. (2010) | DS PMv-M1 TMS: c-TMS 80% rMT, t-TMS 120% rMT, ISI 1–15 ms | Grasp production on objects |
| Kansaku et al. (2007) | DP-TMS: 70% of stimulator output, ISI 10 ms | Counting long sequences of numbers |
| Lago et al. (2010) | DS PMv-M1 TMS: c-TMS 90% aMT, t-TMS MEPs of 1 mV at rest, ISI 6 ms | Action observation |
| Meister et al. (2007) | LF-RTMS: 1 Hz, 900 stimuli, 90% rMT | Speech perception |
| Sato et al. (2009) | LF-RTMS: 1 Hz, 600 stimuli, 110% rMT | Speech perception |
| Tunik et al. (2008) | SP-TMS: 110% rMT | Response selection to manipulable objects, Manual tasks |
| Urgesi et al. (2007a) | DP-TMS: 120% rMT, ISI 100 ms | Discrimination of bodily postures |
| Urgesi et al. (2007b) | DP-TMS: 120% rMT, ISI 100 ms | Discrimination of bodily postures |

*aMT, active motor threshold; c-TMS, conditioning pulse TMS; DS, dual-site TMS; DP-TMS, double-pulse TMS; HF-TMS, High-frequency TMS; ISI, inter-stimulus interval; LF-TMS, Low-frequency TMS; MEPs, motor evoked potentials; rMT, resting motor threshold; SP-TMS, single-pulse TMS; t-TMS, test pulse TMS.*

**Table 3 | TMS studies of SMA.**

| Reference | TMS used | Behavioral deficits induced |
|---|---|---|
| Gerloff et al (1997) | HF-RTMS: 15–20 Hz, 1.4–2.4 s, 96–110% rMT | Complex movement sequences, button pressing |
| Jones et al. (2004) | HF-RTMS: 20 Hz, 4 stimuli, 90% aMT (leg) | Time reproduction of movements, button pressing |
| Obhi et al. (2002) | LF-RTMS: 1 Hz, 300 stimuli, 110% rMT | Bimanual coordination, manual tasks |
| Perez et al. (2008) | LF-RTMS: 1 Hz, 1200 stimuli, 80% rMT | Intermanual transfer, button pressing |
| Serrien et al. (2002) | HF-RTMS: 20 Hz, 50 stimuli, 90% aMT | Bimanual coordination, finger tapping |
| Steyvers et al. (2003) | HF-RTMS: 20 Hz, 10 stimuli, 120% rMT | Bimanual coordination, finger tapping |
| Verwey et al. (2002) | LF-RTMS: 1 Hz, 1200 stimuli, 90% rMT | Complex movement sequences, button pressing |

*aMT, active motor threshold; HF-TMS, High-frequency TMS; LF-TMS, Low-frequency TMS; rMT, resting motor threshold.*

when these movements are performed in the opposite anti-phase direction but not when these movements are performed in the same in-phase direction. TMS over SMA can also disrupt inter-manual transfer (Perez et al., 2008), which relates to a phenomenon whereby training one hand on a manual task can lead to improvements in the untrained hand.

### PRE-SMA IN MOTOR BEHAVIOR

The vertical anterior-commissural line is sometimes used to denote the boundary between SMA and pre-SMA. This division was proposed by Picard and Strick (1996, 2001) based on a meta-analysis carried out to characterize functional activation on the medial wall of the cerebral cortex. They showed that "more complex" tasks tended to engage cortex anterior to this boundary while "less complex" tasks tended to engage more posterior cortex. In support of this view, TMS applied over pre-SMA tends to disrupt processes that are "higher-order" than those that are disrupted from stimulating SMA. These include processes of "cognitive control" such as task switching, response inhibition, and conflict resolution. With respect to task switching, Rushworth et al. (2002) examined the role of pre-SMA in a response selection task in which the stimulus-response rules were sometimes switched during the course of the experiment. For example, participants could begin the experiment with a triangle instructing them to respond with their right hand and a square instructing them to respond with their left hand. After several consecutive trials, they were then instructed to do the reverse. By doing this, Rushworth et al. (2002) revealed that TMS applied over pre-SMA affected response selection but only after a switch was introduced. Similarly, Kennerley et al. (2004) revealed that TMS over pre-SMA can disrupt the initiation of a complex sequence of finger movements but only after participants were instructed to switch between two learned sequences. Interestingly, Kennerley

**Table 4 | TMS studies of pre-SMA.**

| Reference | TMS used | Behavioral deficits induced |
|---|---|---|
| Chen et al. (2009) | DP-TMS: 60% of stimulator output, ISI 100 ms | Response inhibition, button pressing |
| Kennerley et al. (2004) | HF-RTMS: 10 Hz, 5 stimuli, 110% aMT | Complex movement sequences, button pressing |
| Kennerley et al. (2004) | HF-RTMS: 10 Hz, 5 stimuli, 110% aMT | Task switching, button pressing |
| Lau et al. (2007) | SP-TMS: 105% aMT (foot) | Perception of motor intention and actions |
| Mars et al. (2009) | DS preSMA-M1 TMS: c-TMS 120% rMT, t-TMS MEPs of 1–1.5 mV at rest, ISI 6 ms | Conflict resolution in response selection, button pressing |
| Oliveri et al. (2003) | DS preSMA-M1 TMS: c-TMS 70–110% rMT, t-TMS 110% rMT, ISI 4 ms | Response execution to emotional cues, button pressing |
| Rushworth et al. (2002) | HF-RTMS: 5 Hz, 4 stimuli, 105% rMT (foot) | Task switching, button pressing |
| Tanaka et al. (2005) | LF-RTMS: 0.9 Hz, 420 stimuli, 70% of stimulator output | Updating verbal information |
| Taylor et al. (2007) | HF-RTMS: 10 Hz, 3 stimuli, 110% rMT | Conflict resolution in response selection, button pressing |
| Tremblay and Gracco (2009) | HF-RTMS: 10 Hz, 5 stimuli, 110% rMT | Verbal response production |

*aMT, active motor threshold; c-TMS, conditioning pulse TMS; DS, dual-site TMS; DP-TMS, double-pulse TMS; HF-TMS, High-frequency TMS; ISI, inter-stimulus interval; LF-TMS, Low-frequency TMS; rMT, resting motor threshold; SP-TMS, single-pulse TMS; t-TMS, test pulse TMS.*

et al. (2004) also revealed that TMS over pre-SMA applied at "chunk" points in these sequences could disrupt task performance, which is different than what Verwey et al. (2002) had shown from stimulating SMA. Chen et al. (2009) further revealed that the pre-SMA has a role to play in response inhibition. They found that stimulating pre-SMA impaired people's performance to react to a stop signal (presented on 25% of the trials) that instructed them to refrain from responding to a cue that they had just seen. Other studies have shown that TMS applied over pre-SMA also plays a role in conflict resolution (Taylor et al., 2007; Mars et al., 2009).

## PAIRED-PULSE TMS AND DUAL-SITE TMS

Paired-pulse TMS consists of applying two TMS pulses separated closely in time. This type of TMS has been widely used to examine intra-cortical circuits in M1 (Kujirai et al., 1993). Typically, a sub-threshold "conditioning" stimulus is applied over M1 at different time periods before applying a supra-threshold "test" stimulus through the same coil. Given that the intensity of the conditioning pulse is too small to produce motor output from M1, it is generally accepted that any influence that the conditioning pulse has on motor excitability (as assessed by the amplitude of muscle responses invoked by the test pulse) is at the level of neural circuits in M1 but not at the level of neural circuits in the spinal cord. The influence of the conditioning pulse on motor excitability can be either inhibitory, if the inter-stimulus interval (ISI) between the two pulses ranges between 1 and 4 ms, or facilitatory, if the ISI ranges between 8 and 20 ms.

One can also use paired-pulse TMS to examine the time course of interactions in a "two-node" neural circuit using two different coils (**Figure 2**); this approach is known as "dual-site" TMS. The first demonstration of this approach was carried out by Ferbert et al. (1992). Ferbert et al. (1992) showed that applying a conditioning pulse over M1 in one hemisphere suppressed the excitability of M1 in the opposite hemisphere. Importantly, Ferbert et al. (1992) also demonstrated that this suppression was mediated by cortical but not by spinal mechanisms. Namely, the conditioning pulse applied over M1 in one hemisphere had no effect on motor output induced by electrical stimulation of the opposite M1, which is thought to activate corticospinal projections directly, nor did it have any effect on the H-reflex of ipsilateral hand muscles. Subsequent studies have revealed that a conditioning pulse applied over either PMv or pre-SMA can suppress or facilitate the excitability of M1 in the same hemisphere (Oliveri et al., 2003; Davare et al., 2008; Mars et al., 2009). Likewise, a conditioning pulse applied to PMd can

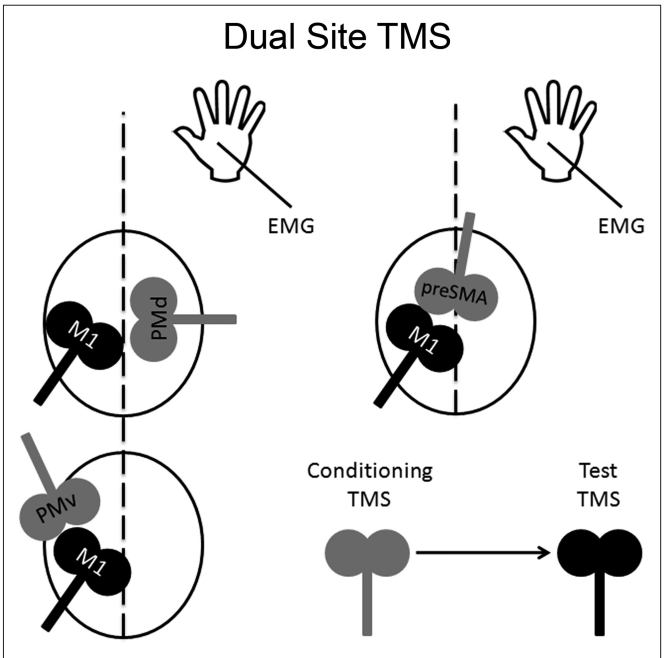



**FIGURE 2 | Dual-site TMS approach.** Dual-site TMS approach can be used to examine the time course of interactions in a particular neural circuit containing a non-primary motor area and M1. The idea is to stimulate a non-primary motor area with a conditioning pulse to examine its effect on a subsequent supra-threshold test pulse to M1. Changes in M1 can be inferred by measuring any possible changes in its motor excitability on a hand muscle using electromyography.

suppress or facilitate the excitability of M1 in the opposite hemisphere (Mochizuki et al., 2004; Baumer et al., 2006). Dual-site TMS can also be used to test whether or not these connections can be modulated by task demands. In the next three subsections, we provide examples of this type of modulation. We should point out that all these studies used figure-of-eight coils to achieve more focal stimulation.

#### PMd – M1 CONNECTIVITY

O'Shea et al. (2007b) examined whether or not functional connectivity between PMd and M1 can be modulated while participants performed a task that required them to choose between different responses based on arbitrary visual cues. The authors revealed that dual-site TMS (conditioning pulse over PMd and test pulse over M1 in the opposite hemisphere) disrupted behavior when this stimulation was applied 100 ms after cue presentation. The authors also revealed that motor excitability of M1 was facilitated during task performance when dual-site TMS was applied 75 ms after cue presentation. Taken together, the authors concluded that selecting responses on the basis of arbitrary visual cues engages PMd 75–100 ms after cue presentation and that during this time period PMd exerts an influence on M1. It is worthwhile to point out that, at present, examination of PMd-M1 connectivity within the same hemisphere is not feasible with dual-site TMS. This is because the physical size of figure-of-eight coils precludes one to target PMd with one coil and M1 with the other coil. Although Civardi et al. (2001) did claim to have examined PMd-M1 connectivity within the same hemisphere using dual-site TMS, it would appear from their figures that a more rostral area in the prefrontal cortex was primarily targeted with the conditioning pulse instead. The same issue also applies for targeting SMA (proper) and M1 with dual-site TMS.

#### PMv – M1 CONNECTIVITY

In three different studies, Davare et al. (2008, 2009, 2010) examined whether or not functional connectivity between PMv and M1 can be modulated differentially when people grasp objects using either a precision grip or a whole-hand grasp. In all three studies, TMS was applied over the left M1 either in isolation or after a conditioning stimulus was applied over the left PMv. In the first study, Davare et al. (2008) showed that a conditioning stimulus to PMv during no movement suppressed motor excitability of M1 whereas a conditioning stimulus to PMv during a whole-hand grasp on a tennis ball caused this inhibition to disappear. They also showed that a conditioning stimulus to PMv during precision grip on a small cube facilitated motor excitability of M1. Given these differential effects exerted by the conditioning stimulus (to PMv) on motor excitability (of M1), the authors concluded that PMv-M1 circuits are differentially modulated under different types of grasps. In the second study, Davare et al. (2009) repeated a similar experiment in which they had participants grasp either a pen using a precision grip or a circular disk using a whole-hand grasp. They also recorded EMG activity in different hand muscles. In doing so, what they had found was that a conditioning stimulus to PMv facilitated motor excitability of M1 but only in the specific hand muscles that were employed in the types of grasp that were carried out. In the third study, Davare et al. (2010) repeated the same experiment before and after applying TBS over the left anterior intra-parietal sulcus (aIPS); TBS to aIPS reduced all effects of the conditioning stimulus to PMv. Based on these results, the authors suggested that the modulations of the PMv-M1 circuit observed before TBS (and in their earlier studies) might be driven by sensory information about the objects' geometrical properties it receives from aIPS.

#### PRE-SMA-M1 CONNECTIVITY

Oliveri et al. (2003) examined whether or not functional connectivity between pre-SMA and the left M1 is differentially modulated by simple button responses to photographs that were either neutral or unpleasant in emotional content. They reasoned that pre-SMA might play a greater role in the execution of movements in response to visual stimuli with stronger emotional content. Consistent with their hypothesis, their results revealed that a conditioning pulse of TMS applied over pre-SMA facilitated motor excitability of M1 only when participants saw the emotionally unpleasant photographs as compared with either the presentation of these same photographs during TMS applied over M1 alone or when dual-site TMS was applied during the presentation of the emotionally neutral photographs. In a very different study, Mars et al. (2009) examined the timing of the pre-SMA effects on M1 in a response selection task during conflict resolution. Two different colored-flanker stimuli were presented in each hemifield (e.g., green on the left and red on the right) and participants were required to respond to a central cue that became either green or red, which instructed them to respond with the index finger of the hand that was on the same side as the flanker of the same color. The critical manipulation was that the central cue took the same color for several consecutive trials before switching to the next color. Their results revealed that a conditioning stimulus to pre-SMA facilitated motor excitability of the left M1 during task performance but only during the switch trials. Based on these findings, the authors concluded that the pre-SMA influences the left M1 during a response selection task in cases when re-programming is necessary.

## COMBINING TMS AND FUNCTIONAL NEUROIMAGING IN NORMAL INDIVIDUALS

The second half of the 1990s saw the emergence of studies that combined TMS and functional neuroimaging concurrently (Fox et al., 1997; Paus et al., 1997; Bohning et al., 1998; Siebner et al., 1998). The underlying logic is simple. Using a perturb-and-measure approach, one can alter neural activity in one brain region to evaluate the effects that this manipulation has on neural activity elsewhere in the brain (Paus, 2005). Changes in neural activity in regions in the brain other than the one that was stimulated can be inferred as being connected to the region that was stimulated. In other words, one can examine connections in the brain by stimulating a target region of the cortex with TMS and measuring changes in neural activity elsewhere in the brain with either PET or fMRI. In this section, we will review some studies that used this combined approach for examining PMd connectivity.

In a combined TMS/PET study, we examined the effects of applying off-line 1-Hz repetitive TMS to either the left PMd or the left M1 on both the motor excitability of the left M1 and regional cerebral blood flow (CBF) throughout the brain (Chouinard et al., 2003). Specifically, we mapped networks of brain regions in which

changes in CBF correlated with changes in the motor excitability of the left M1 after applying repetitive TMS over either the left PMd or the left M1. We interpreted these correlations as an index of neural modulation induced by the repetitive TMS. Although repetitive stimulation at the two adjacent cortical sites produced the same effects on motor excitability, statistical maps of correlations between the magnitude of MEP (motor evoked potential) suppression and changes in CBF revealed two distinct patterns of distal neural modulation (**Figure 3**). Neural modulation occurred in a small number of brain regions after 1-Hz repetitive TMS to the left M1, many of these confined to the non-primary motor areas and sub-cortical motor structures. In contrast, neural modulation occurred in multiple regions after 1-Hz repetitive TMS to the left PMd; these included motor areas in the frontal cortex as well as more associational regions in the parietal and prefrontal cortices. We concluded that these findings were consistent with known differences between PMd and M1 in the extent of their anatomical connectivity in the macaque monkey. Lee et al. (2003) and Siebner et al. (2003) have also performed similar TMS/PET studies that examined the effects of applying off-line 1-Hz repetitive TMS over either the left PMd or the left M1. In addition, Bestmann et al. (2005, 2008) have done some interesting work stimulating the left PMd during fMRI. Although this procedure is technically challenging, activation patterns induced by stimulating the left PMd concurrently during fMRI acquisition appears to be in agreement with those obtained in the TMS/PET studies.

In a different study, O'Shea et al. (2007a) demonstrated that a network of cortical regions can compensate for function when the left PMd is disrupted by TMS. In one experiment, without fMRI, the authors showed that after applying 1-Hz repetitive TMS over the left PMd, performance on a response-selection task was disrupted temporarily – but recovered after only 4 min. Equally important, this disruption in behavior did not coincide with reductions in motor excitability of the left M1, which remained suppressed for a considerable amount of time after performance had recovered. Taken together, this suggested to the authors that some sort of adaptive compensation had taken place. Namely, a different region might be taking over function of the left PMd, which they believed was still disrupted as indexed by the suppression in the motor excitability of the left M1. In a different experiment, the authors then used fMRI to measure changes in BOLD after performance had recovered from this TMS-induced disruption. They found changes in BOLD in the right PMd (as well as in other distal brain regions) during the performance of the response-selection task. In a final experiment,

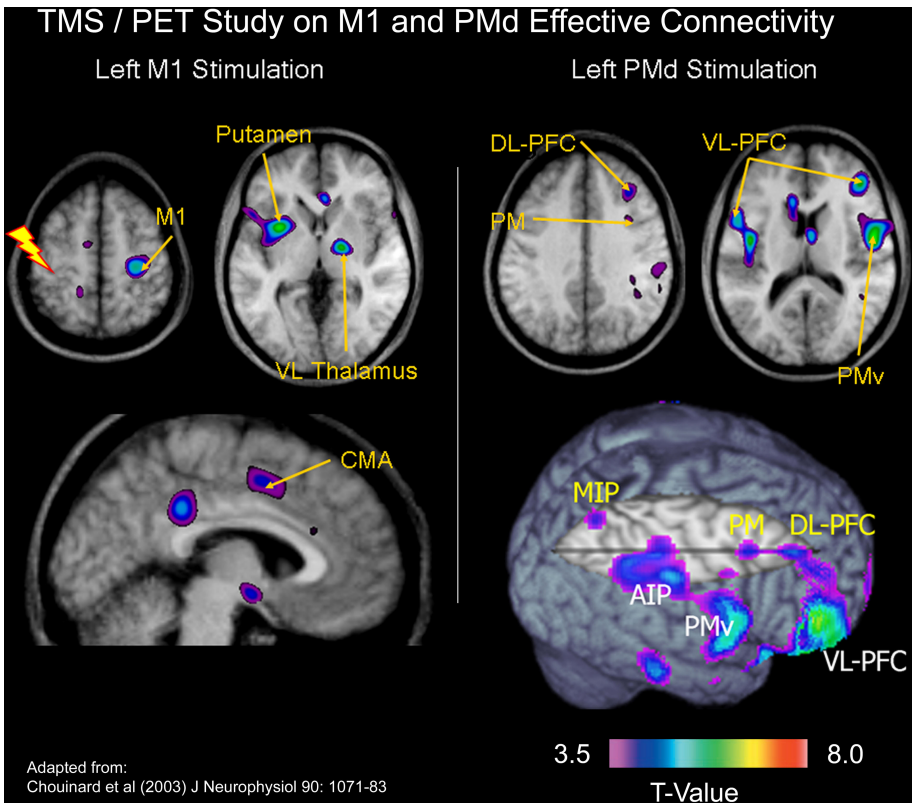



**FIGURE 3 | TMS/PET study on M1 and PMd effective connectivity.** In a combined TMS/PET study, we mapped networks of brain regions in which changes in cerebral blood flow correlated with changes in the motor excitability of the left M1 after applying repetitive TMS over either the left PMd or the left M1 (Chouinard et al., 2003). We interpreted these correlations as an index of neural modulation induced by the repetitive TMS. Although repetitive stimulation at the two adjacent cortical sites produced the same effects on motor excitability, statistical maps of correlations between the magnitude of MEP suppression and changes in cerebral blood flow revealed two distinct patterns of distal neural modulation. Abbreviations: MIP = medial intraparietal area; DL-PFC = dorsolateral prefrontal cortex; AIP = anterior intra-parietal area; VL-PFC = ventrolateral prefrontal cortex; VL Thalamus = ventrolateral thalamus.

without fMRI, O'Shea et al. (2007a) found that delivering TMS over the right PMd by itself did not disrupt response selection but doing exactly the same thing after first disrupting the left PMd did result in deficits. Taken together, the authors concluded from these series of experiments that the observed compensation in performance following TMS-induced disruption of neural processing in the left PMd depended critically on intact neural processing in the right PMd.

## COMBINING TMS AND FUNCTIONAL NEUROIMAGING IN STROKE PATIENTS

One of the biggest problems that researchers face in charting recovery in the brain after cerebral damage is how to disentangle compensatory mechanisms from "true" cerebral reorganization (Krakauer, 2007). Patients with brain damage are impaired in behavior and they will perform a task in a way that is, by necessity, different from that performed by neurologically intact controls. This difference in how a task is carried out can lead to differences in brain activation between these two groups of participants that might reflect different neural operations to complete the task – also known as compensatory mechanisms. Not only is this a problem for cross-sectional studies but it is also a problem for longitudinal studies. As patients make a meaningful recovery, by definition, their performance improves and they will perform a task differently after their recovery. It is often unclear whether any resulting differences in brain activation might reflect compensatory mechanisms or cerebral reorganization or both. The question then arises as to how one can examine cerebral reorganization without examining changes in compensation? One solution is to use TMS to examine the integrity of neural circuits independently of behavior. In this section, we present a study that we conducted a few years ago that illustrates how one can combine TMS and functional neuroimaging concurrently to provide a behavior-independent assay of effective connectivity to chart motor recovery after stroke. The point that we wish to highlight is that TMS can offer the opportunity to chart changes in effective connectivity without any task confounds by examining its effects on hemodynamic measurements while patients are not engaged in a task.

In a TMS/PET study, we examined changes in the effective connectivity of M1 in seven patients with capsular strokes who underwent constraint-induced movement therapy (CI therapy; Taub et al., 2002) one year after their stroke (Chouinard et al., 2006). We adopted a TMS/PET paradigm that had previously been used by our research group. In normal volunteers, Paus et al. (1998) applied sub-threshold 10-Hz repetitive TMS over M1 and varied the number of TMS trains delivered during each block of PET scanning. In doing so, the CBF response correlated negatively with the number of stimulus trains delivered both at the site of stimulation and in several distal brain regions known to be connected transsynaptically in the monkey (**Figure 4A**). Based on these findings, Paus et al. (1998) speculated that the trains of stimulation resulted in an activation of local inhibitory mechanisms and a subsequent reduction of excitatory synaptic activity in the stimulated region and in an interconnected network. Given the success of this protocol in normal volunteers, we carried out the same procedures before and after the stroke patients had their CI therapy.

We analyzed the CBF data in the patients in two different ways. In the first analysis, we asked whether or not correlations between CBF and TMS trains differed for any brain regions after CI therapy from those seen before CI therapy. This analysis (**Figure 4B**) revealed that both the stimulated ipsilesional M1 (i.e., M1 in the damaged hemisphere) and a more distal ipsilesional CMA reverted back to the more normal inverse relationship between CBF and TMS trains previously seen in healthy participants (Paus et al., 1998). We speculated that that these findings reflected a strengthening of local inhibitory neurons in the ipsilesional M1, which have been shown to be important for the fractionation or the isolation of proximal and distal muscles (Keller, 1993), as well as a strengthening of connections between the ipsilesional M1 and CMA in the same hemisphere. In the second analysis, we examined the relationship between motor improvement and changes in the CBF response to TMS between the two PET sessions. This analysis revealed an inverse relationship locally in the stimulated ipsilesional M1 (**Figure 4C**) which suggested to us that the observed changes in M1 were adaptive. This relationship, however, did not reach significance in the ipsilesional CMA – which may relate to a lack of power from having only seven patients.

Also, we should mention that there is an emerging interest in the use of TMS as a therapeutic tool to drive motor recovery after stroke. Not all motor deficits after stroke relate directly to the damaged brain tissue. Abnormal interactions with cortical regions remote from the site of damage can also contribute to motor deficits (for review, see Nowak et al., 2009). There have been reports that capsular strokes can lead to greater trans-callosal inhibition in the ipsilesional M1 originating from the contralesional M1 as revealed by dual-site TMS (Murase et al., 2004). This observation has led some researchers to inquire whether or not TMS-induced down regulation of the contralesional M1 might aid motor recovery after capsular stroke. Nowak et al. (2008) revealed that 10 min of off-line 1-Hz repetitive TMS to the contralesional M1 can lead to motor improvements. The study also found that when the same patients performed a grip task with the affected hand, task performance after the TMS was applied invoked smaller blood oxygenation-level dependent (BOLD) responses in the contralesional motor areas and greater BOLD responses in ipsilesional motor areas as compared with task performance after a sham stimulation was delivered. In a different paper, Grefkes et al. (2010) reanalyzed these data using dynamic causal modeling. The reanalysis revealed a reduction in neuronal "coupling" between the two primary motor areas and an increase in neuronal coupling between the ipsilesional M1 and the ipsilesional non-primary motor areas. In another study, Ameli et al. (2009) demonstrated that an off-line TMS protocol consisting of short bursts of 10-Hz TMS applied to the ipsilesional M1 over a period of several minutes can also lead to motor improvements. The study also revealed that when the same patients performed a grip task with the affected hand, task performance after the TMS was applied invoked smaller BOLD responses in the contralesional motor areas as compared with task performance after a sham stimulation. These studies concluded that TMS can drive motor improvements in patients with capsular strokes via a TMS-induced down regulation of the contralesional M1 and a TMS-induced up regulation of the ipsilesional M1.

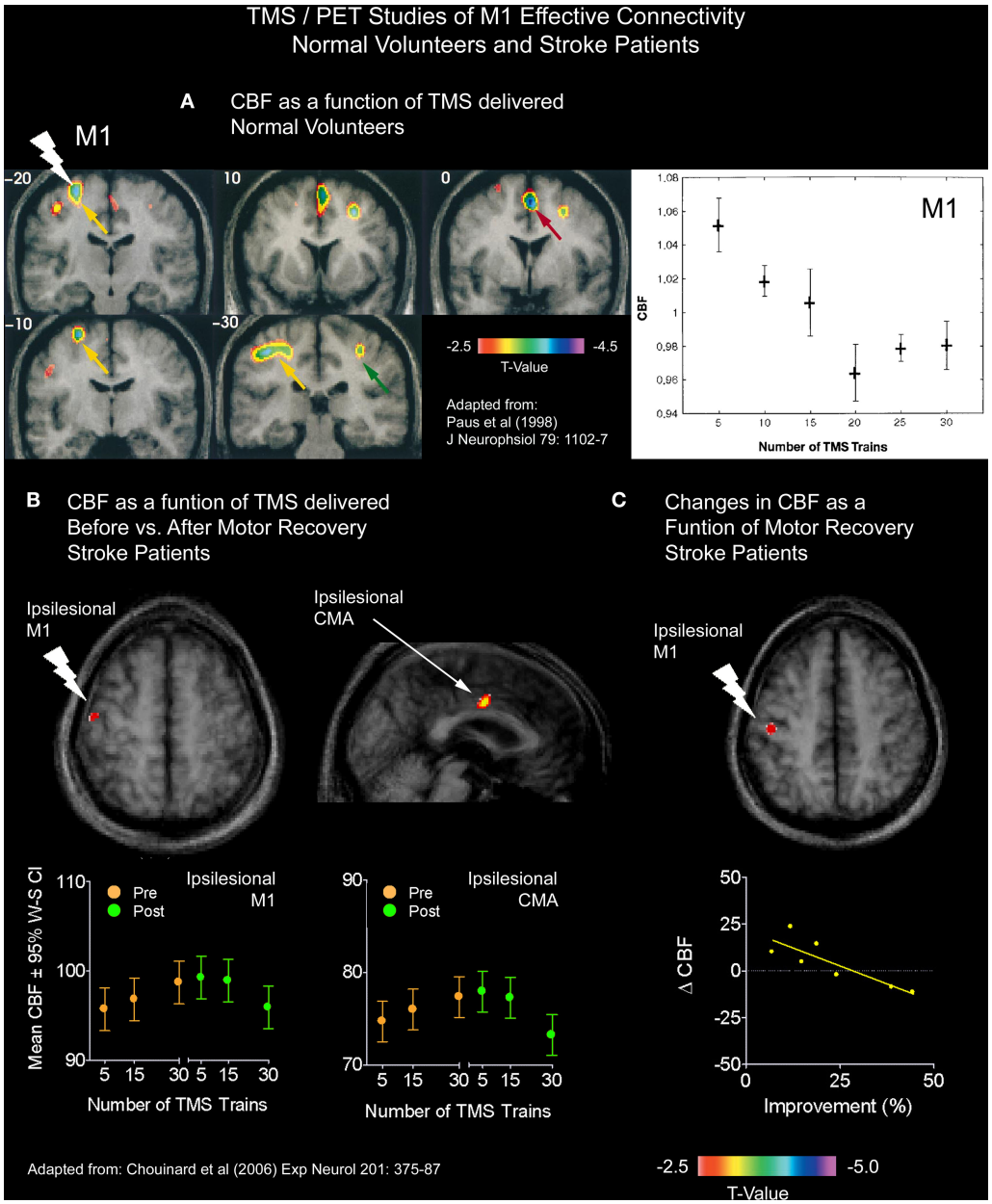



**FIGURE 4 | TMS/PET studies of M1 effective connectivity on normal volunteers and stroke patients. (A)** Paus et al. (1998) applied sub-threshold 10-Hz repetitive TMS over M1 and varied the number of TMS trains delivered during each block of PET scanning. In doing so, the CBF response correlated negatively with the number of stimulus trains delivered both at the site of stimulation and in several distal brain regions. Given the success of this protocol in normal volunteers, we carried out the same procedures before and after the stroke patients had their CI therapy. **(B)** In one analysis, we asked whether or not correlations between CBF and TMS trains differed for any brain regions after CI therapy from those seen before CI therapy. This analysis revealed that both the stimulated ipsilesional M1 and a more distal ipsilesional CMA reverted back to the more normal inverse relationship between CBF and TMS trains. **(C)** In another analysis, we examined the relationship between motor improvement and changes in the CBF response to TMS between the two PET sessions. This analysis revealed an inverse relationship locally in the stimulated ipsilesional M1, which suggests that the observed changes in M1 were adaptive.

## METHODOLOGICAL CHALLENGES AND FUTURE DIRECTIONS FOR USING TMS

More work is needed to demonstrate functional specificity between non-primary motor areas through double dissociations. Although there have been some notable demonstrations (e.g., Davare et al., 2006 showed that PMd plays a role in arm reaching but not in grasping while PMv plays a role in grasping but not in reaching; Chouinard et al., 2005 showed with TMS that when people lift objects, M1 but not PMd plays a role in scaling lifting forces based on information acquired during a previous lift and that PMd but not M1 plays a role in scaling forces based on arbitrary visual cues), many of the TMS studies that we reviewed only stimulated one

cortical site. Designing a TMS study aimed at testing for double dissociations is a good idea for a number of other reasons. One of the reasons is that stimulating two different cortical areas can serve as a better control than sham stimulation. Sham stimulation does not control for all peripheral effects associated with the real stimulation. It releases magnetic fields around the head (not in the brain) and it feels noticeably different to participants than real TMS. Another issue to consider is that non-primary motor areas along the medial wall of the brain (SMA and pre-SMA) require higher levels of stimulation than those located on the lateral surface of the brain (PMd and PMv). This is due to the fact that these medial areas are further away from the TMS coil. Thus, it is conceivable that targeting a non-primary motor area on the medial wall could result in concurrent disruption of neural processing in the medial aspect of PMd close to the interhemispheric fissure. One solution to this problem would be to use PMd as a control site (e.g., Rushworth et al., 2002; Kennerley et al., 2004; Mars et al., 2009).

As we mentioned earlier, an important limitation of the dual-site approach is that the physical size of the figure-of-eight coils precludes one to target either the ipsilateral PMd or SMA (proper) with one coil and M1 with the other coil. Moreover, as we discussed in the Introduction, the more caudal portions of the non-primary motor areas exert a much stronger influence on M1 than the more rostral portions of the non-primary motor areas – at least in the monkey (Barbas and Pandya, 1987). It is therefore unclear as to whether or not conditioning TMS in dual-site TMS studies might encroach on more caudal regions of non-primary motor areas, which cannot be targeted directly, or whether or not the influences reported in studies reflect indirect connections between the cortical area that is being stimulated with the conditioning TMS pulse and M1. One way to resolve this issue would be to carry out dual-site TMS concurrently with functional neuroimaging. Also, the dual-site approach only enables one to examine the influence that one cortical area exerts over M1. In contrast, the combined TMS and functional neuroimaging approach enables one to examine the influence that one cortical area exerts on more than one area in the brain.

The sensitivity of TMS could improve. One important avenue for future research would be to compare variability and effect sizes induced by different stimulation protocols, which is somewhat lacking in the TMS literature, and to develop new approaches to TMS that might tap into more subtle forms of neuronal processing. For example, some researchers have started to incorporate principles similar to those underlying fMRI-adaptation as a way to tap into the processing of a small subset of neurons in a stimulated cortical region (Silvanto et al., 2007; Cattaneo et al., 2010). FMRI-adaptation is used in both vision (for review, see Grill-Spector et al., 2006) and motor (Chouinard and Goodale, 2009; Chouinard et al., 2009b) research as a way to tap into the processing of a small subset of neurons. In the case of fMRI, when a brain area contains neurons that code for a particular stimulus or action, the hemodynamic response is higher during conditions in which the stimulus or action changes across trials as compared to conditions in which the stimulus or action remains the same. In a similar way, TMS can be used to disrupt behavioral priming effects to test whether or not a particular brain area contains a subset of neurons that code for a particular stimulus or action. Cattaneo et al. (2010) recently used this approach to demonstrate category-specific neuronal representations in the left PMv in semantic processing for tools but not for animals.

Transcranial magnetic stimulation has two important advantages over functional neuroimaging: TMS offers better temporal resolution and allows one to examine causality. These two strengths can make TMS more suitable for examining effective connectivity than functional neuroimaging alone. The temporal response of the blood supply, which is the basis of fMRI, is much slower than the electrical signals that underlie neuronal communication (Kim et al., 1997). As a result, it is difficult to infer with fMRI how brain areas are interconnected. Although a number of mathematical methods have been put forth to make inferences about connections in fMRI data (e.g., structural equation modeling: McIntosh and Gonzalez-Lima, 1991; dynamic causal modeling: Friston et al., 2003; granger causality: Goebel et al., 2003; graph theory: Reijneveld et al., 2007; Bullmore and Sporns, 2009), results from these types of analyses, because they are based on computing correlations, may reflect relationships between different task components (and/or "third" parties) rather than true connections. Moreover, TBS offers potential to make the combined TMS and functional neuroimaging approach more accessible to laboratories that either do not have access to PET or do not wish to employ TMS inside an MRI scanner. This is because TBS protocols take less time to apply outside of the magnet than other off-line protocols such as in prolonged periods of stimulation at 1-Hz. For these reasons, we foresee an increase in the use of TMS for examining effective connectivity in the brain.

Last, we discussed the use of TMS for charting and driving motor recovery processes after stroke. One of the biggest problems that researchers face in charting recovery in the brain after cerebral damage is how to disentangle compensatory mechanisms from "true" cerebral reorganization (Krakauer, 2007). As we mentioned earlier, one solution would be to use TMS independently of behavior. The same principles that we discussed earlier can be equally applied for charting the progression of a number of brain disorders and the mechanisms of recovery that underlie their treatments. In recent years, we have seen a considerable amount of research devoted to developing TMS as a viable treatment option for stroke (Ridding and Rothwell, 2007; Nowak et al., 2009), movement disorders (Edwards et al., 2008), epilepsy (Kimiskidis, 2010), depression (Ridding and Rothwell, 2007; Kim et al., 2009), schizophrenia (Kim et al., 2009) and a number of other brain disorders (Ridding and Rothwell, 2007). This research has considered a number of ethical and safety issues: how to identify patients that might benefit from TMS treatment, how TMS might interact with concurrent treatments, what stimulation protocols are most effective in treating a particular disorder, what are the cost and benefits to the patients, and a better understanding of the physiological mechanisms that underlie changes in symptoms. Thus, we foresee a rise in the use of TMS in clinical practice.

## CLOSING REMARKS

We reviewed the different approaches that one can use with TMS to study both its effects on motor behavior and neural connections in the brain. We also presented the evidence obtained from TMS showing that PMd, PMv, SMA, and pre-SMA each have different roles to play in motor behavior. Namely, we highlighted the

importance of PMd in response selection based on arbitrary cues and in the control of arm movements, PMv in grasping and in the discrimination of bodily actions, SMA in movement sequencing and in bimanual coordination, and pre-SMA in cognitive control. We also presented the evidence from dual-site TMS that each of these areas can influence M1. We then went on to show that the combination of TMS and functional neuroimaging can provide opportunities to examine how a non-primary motor area (or any other region that can be stimulated with TMS for that matter) is connected to regions other than M1. We discussed some of the challenges that imagers face when charting "true" cerebral reorganization in the brain and proposed that the use of TMS can help eliminate these problems when used independently of task. We also discussed how TMS can be used as a therapeutic tool to aid motor recovery after stroke. We then ended our review by discussing some of the methodological challenges and future directions for using this tool in basic and clinical neuroscience.

## ACKNOWLEDGMENTS

Philippe Chouinard is supported by a post-doctoral award from the Ontario Mental Health Foundation. Tomás˘ Paus' work is supported by the Canadian Institutes of Health Research, the National Institutes of Health (United States), and the Sixth Framework Program of the European Union.

## REFERENCES

Ameli, M., Grefkes, C., Kemper, F., Riegg, F. P., Rehme, A. K., Karbe, H., Fink, G. R., and Nowak, D. A. (2009). Differential effects of high-frequency repetitive transcranial magnetic stimulation over ipsilesional primary motor cortex in cortical and subcortical middle cerebral artery stroke. *Ann. Neurol.* 66, 298–309.

Barbas, H., and Pandya, D. N. (1987). Architecture and frontal cortical connections of the premotor cortex (area 6) in the rhesus monkey. *J. Comp. Neurol.* 256, 211–228.

Barker, A. T., Jalinous, R., and Freeston, I. L. (1985). Non-invasive magnetic stimulation of human motor cortex. *Lancet* 1, 1106–1107.

Baumer, T., Bock, F., Koch, G., Lange, R., Rothwell, J. C., Siebner, H. R., and Munchau, A. (2006). Magnetic stimulation of human premotor or motor cortex produces interhemispheric facilitation through distinct pathways. *J. Physiol.* 572, 857–868.

Been, G., Ngo, T. T., Miller, S. M., and Fitzgerald, P. B. (2007). The use of tDCS and CVS as methods of non-invasive brain stimulation. *Brain Res. Rev.* 56, 346–361.

Bestmann, S. (2008). The physiological basis of transcranial magnetic stimulation. *Trends Cogn. Sci.* 12, 81–83.

Bestmann, S., Baudewig, J., Siebner, H. R., Rothwell, J. C., and Frahm, J. (2005). BOLD MRI responses to repetitive TMS over human dorsal premotor cortex. *Neuroimage* 28, 22–29.

Bestmann, S., Swayne, O., Blankenburg, F., Ruff, C. C., Haggard, P., Weiskopf, N., Josephs, O., Driver, J., Rothwell, J. C., and Ward, N. S. (2008). Dorsal premotor cortex exerts state-dependent causal influences on activity in contralateral primary motor and dorsal premotor cortex. *Cereb. Cortex* 18, 1281–1291.

Bohning, D. E., Shastri, A., Nahas, Z., Lorberbaum, J. P., Andersen, S. W., Dannels, W. R., Haxthausen, E. U., Vincent, D. J., and George, M. S. (1998). Echoplanar BOLD fMRI of brain activation induced by concurrent transcranial magnetic stimulation. *Invest. Radiol.* 33, 336–340.

Brighina, F., Bisiach, E., Piazza, A., Oliveri, M., La Bua, V., Daniele, O., and Fierro, B. (2002). Perceptual and response bias in visuospatial neglect due to frontal and parietal repetitive transcranial magnetic stimulation in normal subjects. *Neuroreport* 13, 2571–2575.

Buch, E. R., Mars, R. B., Boorman, E. D., and Rushworth, M. F. (2010). A network centered on ventral premotor cortex exerts both facilitatory and inhibitory control over primary motor cortex during action reprogramming. *J. Neurosci.* 30, 1395–1401.

Bullmore, E., and Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat. Rev. Neurosci.* 10, 186–198.

Busan, P., Barbera, C., Semenic, M., Monti, F., Pizzolato, G., Pelamatti, G., and Battaglini, P. P. (2009). Effect of transcranial magnetic stimulation (TMS) on parietal and premotor cortex during planning of reaching movements. *PLoS ONE* 4, e4621. doi:10.1371/journal.pone.0004621.

Candidi, M., Urgesi, C., Ionta, S., and Aglioti, S. M. (2008). Virtual lesion of ventral premotor cortex impairs visual perception of biomechanically possible but not impossible actions. *Soc. Neurosci.* 3, 388–400.

Cattaneo, Z., Devlin, J. T., Salvini, F., Vecchi, T., and Silvanto, J. (2010). The causal role of category-specific neuronal representations in the left ventral premotor cortex (PMv) in semantic processing. *Neuroimage* 49, 2728–2734.

Chen, C. Y., Muggleton, N. G., Tzeng, O. J., Hung, D. L., and Juan, C. H. (2009). Control of prepotent responses by the superior medial frontal cortex. *Neuroimage* 44, 537–545.

Chouinard, P. A., and Goodale, M. A. (2009). FMRI adaptation during performance of learned arbitrary visuomotor conditional associations. *Neuroimage* 48, 696–706.

Chouinard, P. A., Whitwell, R. L., and Goodale, M. A. (2009a). The lateral-occipital and the inferior-frontal cortex play different roles during the naming of visually presented objects. *Hum. Brain Mapp.* 30, 3851–3864.

Chouinard, P. A., Large, M. E., Chang, E. C., and Goodale, M. A. (2009b). Dissociable neural mechanisms for determining the perceived heaviness of objects and the predicted weight of objects during lifting: an fMRI investigation of the size-weight illusion. *Neuroimage* 44, 200–212.

Chouinard, P. A., Leonard, G., and Paus, T. (2005). Role of the primary motor and dorsal premotor cortices in the anticipation of forces during object lifting. *J. Neurosci.* 25, 2277–2284.

Chouinard, P. A., Leonard, G., and Paus, T. (2006). Changes in effective connectivity of the primary motor cortex in stroke patients after rehabilitative therapy. *Exp. Neurol.* 201, 375–387.

Chouinard, P. A., Morrissey, B. F., Kohler, S., and Goodale, M. A. (2008). Repetition suppression in occipital-temporal visual areas is modulated by physical rather than semantic features of objects. *Neuroimage* 41, 130–144.

Chouinard, P. A., Van Der Werf, Y. D., Leonard, G., and Paus, T. (2003). Modulating neural networks with transcranial magnetic stimulation applied over the dorsal premotor and primary motor cortices. *J. Neurophysiol.* 90, 1071–1083.

Civardi, C., Cantello, R., Asselman, P., and Rothwell, J. C. (2001). Transcranial magnetic stimulation can be used to test connections to primary motor areas from frontal and medial cortex in humans. *Neuroimage* 14, 1444–1453.

Colby, C. L., and Duhamel, J. R. (1991). Heterogeneity of extrastriate visual areas and multiple parietal areas in the macaque monkey. *Neuropsychologia* 29, 517–537.

Dafotakis, M., Sparing, R., Eickhoff, S. B., Fink, G. R., and Nowak, D. A. (2008). On the role of the ventral premotor cortex and anterior intraparietal area for predictive and reactive scaling of grip force. *Brain Res.* 1228, 73–80.

Davare, M., Andres, M., Cosnard, G., Thonnard, J. L., and Olivier, E. (2006). Dissociating the role of ventral and dorsal premotor cortex in precision grasping. *J. Neurosci.* 26, 2260–2268.

Davare, M., Lemon, R., and Olivier, E. (2008). Selective modulation of interactions between ventral premotor cortex and primary motor cortex during precision grasping in humans. *J. Physiol.* 586, 2735–2742.

Davare, M., Montague, K., Olivier, E., Rothwell, J. C., and Lemon, R. N. (2009). Ventral premotor to primary motor cortical interactions during object-driven grasp in humans. *Cortex* 45, 1050–1057.

Davare, M., Rothwell, J. C., and Lemon, R. N. (2010). Causal connectivity between the human anterior intraparietal area and premotor cortex during grasp. *Curr. Biol.* 20, 176–181.

Dum, R. P., and Strick, P. L. (1991). The origin of corticospinal projections from the premotor areas in the frontal lobe. *J. Neurosci.* 11, 667–689.

Edwards, M. J., Talelli, P., and Rothwell, J. C. (2008). Clinical applications of transcranial magnetic stimulation in patients with movement disorders. *Lancet Neurol.* 7, 827–840.

Farah, M. J. (2004). *Visual Agnosia*. Cambridge, MA: MIT Press.

Ferbert, A., Priori, A., Rothwell, J. C., Day, B. L., Colebatch, J. G., and Marsden, C. D. (1992). Interhemispheric inhibition of the human motor cortex. *J. Physiol.* 453, 525–546.

Fox, P., Ingham, R., George, M. S., Mayberg, H., Ingham, J., Roby, J., Martin, C., and Jerabek, P. (1997). Imaging human intra-cerebral connectivity by PET during TMS. *Neuroreport* 8, 2787–2791.

Friston, K. J., Harrison, L., and Penny, W. (2003). Dynamic causal modelling. *Neuroimage* 19, 1273–1302.

Galletti, C., Fattori, P., Battaglini, P. P., Shipp, S., and Zeki, S. (1996). Functional demarcation of a border between areas V6 and V6A in the superior parietal gyrus of the macaque monkey. *Eur. J. Neurosci.* 8, 30–52.

Gerloff, C., Corwell, B., Chen, R., Hallett, M., and Cohen, L. G. (1997). Stimulation over the human supplementary motor area interferes with the organization of future elements in complex motor sequences. *Brain* 120, 1587–1602.

Giovannelli, F., Borgheresi, A., Balestrieri, F., Ragazzoni, A., Zaccara, G., Cincotta, M., and Ziemann, U. (2006). Role of the right dorsal premotor cortex in "physiological" mirror EMG activity. *Exp. Brain Res.* 175, 633–640.

Goebel, R., Roebroeck, A., Kim, D. S., and Formisano, E. (2003). Investigating directed cortical interactions in time-resolved fMRI data using vector autoregressive modeling and Granger causality mapping. *Magn. Reson. Imaging* 21, 1251–1261.

Goldberg, G. (1985). Supplementary motor area structure and function: review and hypotheses. *Behav. Brain. Sci.* 8, 567–616.

Grefkes, C., Nowak, D. A., Wang, L. E., Dafotakis, M., Eickhoff, S. B., and Fink, G. R. (2010). Modulating cortical connectivity in stroke patients by rTMS assessed with fMRI and dynamic causal modeling. *Neuroimage* 50, 233–242.

Grill-Spector, K., Henson, R., and Martin, A. (2006). Repetition and the brain: neural models of stimulus-specific effects. *Trends Cogn. Sci.* 10, 14–23.

He, S. Q., Dum, R. P., and Strick, P. L. (1993). Topographic organization of corticospinal projections from the frontal lobe: motor areas on the lateral surface of the hemisphere. *J. Neurosci.* 13, 952–980.

He, S. Q., Dum, R. P., and Strick, P. L. (1995). Topographic organization of corticospinal projections from the frontal lobe: motor areas on the medial surface of the hemisphere. *J. Neurosci.* 15, 3284–3306.

Herwig, U., Abler, B., Schonfeldt-Lecuona, C., Wunderlich, A., Grothe, J., Spitzer, M., and Walter, H. (2003). Verbal storage in a premotor-parietal network: evidence from fMRI-guided magnetic stimulation. *Neuroimage* 20, 1032–1041.

Huang, Y. Z., Edwards, M. J., Rounis, E., Bhatia, K. P., and Rothwell, J. C. (2005). Theta burst stimulation of the human motor cortex. *Neuron* 45, 201–206.

Huang, Y. Z., Rothwell, J. C., Edwards, M. J., and Chen, R. S. (2008). Effect of physiological activity on an NMDA-dependent form of cortical plasticity in human. *Cereb. Cortex* 18, 563–570.

Huettel, S., Song, A. W., and McCarthy, G. (2009). Functional Magnetic Resonance Imaging (Sunderland, Mass: Sinauer Associates).

Hughlings Jackson, J. (1958). "On the anatomical and physiological localization of movements in the brain," in *Selected Writings of John Hughlings Jackson,* ed. J. Taylor, (New York: Basic Books), 37–76.

Jeannerod, M., Arbib, M. A., Rizzolatti, G., and Sakata, H. (1995). Grasping objects: the cortical mechanisms of visuomotor transformation. *Trends Neurosci.* 18, 314–320.

Jones, C. R., Rosenkranz, K., Rothwell, J. C., and Jahanshahi, M. (2004). The right dorsolateral prefrontal cortex is essential in time reproduction: an investigation with repetitive transcranial magnetic stimulation. *Exp. Brain Res.* 158, 366–372.

Kalaska, J. F., Scott, S. H., Cisek, P., and Sergio, L. E. (1997). Cortical control of reaching movements. *Curr. Opin. Neurobiol.* 7, 849–859.

Kansaku, K., Carver, B., Johnson, A., Matsuda, K., Sadato, N., and Hallett, M. (2007). The role of the human ventral premotor cortex in counting successive stimuli. *Exp. Brain Res.* 178, 339–350.

Kanwisher, N., Chun, M. M., McDermott, J., and Ledden, P. J. (1996). Functional imaging of human visual recognition. *Brain Res. Cogn. Brain Res.* 5, 55–67.

Keller, A. (1993). Intrinsic synaptic organization of the motor cortex. *Cereb. Cortex* 3, 430–441.

Kennerley, S. W., Sakai, K., and Rushworth, M. F. (2004). Organization of action sequences and the role of the pre-SMA. *J. Neurophysiol.* 91, 978–993.

Kim, D. R., Pesiridou, A., and O'Reardon, J. P. (2009). Transcranial magnetic stimulation in the treatment of psychiatric disorders. *Curr. Psychiatry Rep.* 11, 447–452.

Kim, S. G., Richter, W., and Ugurbil, K. (1997). Limitations of temporal resolution in functional MRI. *Magn. Reson. Med.* 37, 631–636.

Kimiskidis, V. K. (2010). Transcranial magnetic stimulation for drug-resistant epilepsies: rationale and clinical experience. *Eur. Neurol.* 63, 205–210.

Koski, L., Molnar-Szakacs, I., and Iacoboni, M. (2005). Exploring the contributions of premotor and parietal cortex to spatial compatibility using image-guided TMS. *Neuroimage* 24, 296–305.

Krakauer, J. W. (2007). Avoiding performance and task confounds: multimodal investigation of brain reorganization after stroke rehabilitation. *Exp. Neurol.* 204, 491–495.

Kujirai, T., Caramia, M. D., Rothwell, J. C., Day, B. L., Thompson, P. D., Ferbert, A., Wroe, S., Asselman, P., and Marsden, C. D. (1993). Corticocortical inhibition in human motor cortex. *J. Physiol.* 471, 501–519.

Lago, A., Koch, G., Cheeran, B., Marquez, G., Sanchez, J. A., Ezquerro, M., Giraldez, M., and Fernandez-Del-Olmo, M. (2010). Ventral premotor to primary motor cortical interactions during noxious and naturalistic action observation. *Neuropsychologia* 48, 1802–1806.

Lau, H. C., Rogers, R. D., and Passingham, R. E. (2007). Manipulating the experienced onset of intention after action execution. *J. Cogn. Neurosci.* 19, 81–90.

Lee, J. H., and van Donkelaar, P. (2006). The human dorsal premotor cortex generates on-line error corrections during sensorimotor adaptation. *J. Neurosci.* 26, 3330–3334.

Lee, L., Siebner, H. R., Rowe, J. B., Rizzo, V., Rothwell, J. C., Frackowiak, R. S., and Friston, K. J. (2003). Acute remapping within the motor system induced by low-frequency repetitive transcranial magnetic stimulation. *J. Neurosci.* 23, 5308–5318.

Liuzzi, G., Hörniß, V., Hoppe, J., Heise, K., Zimerman, M., Gerloff, C., and Hummel, F. C. (2010). Distinct temporospatial interhemispheric interactions in the human primary and premotor cortex during movement preparation. *Cereb. Cortex* 20, 1323–1331.

Malach, R., Reppas, J. B., Benson, R. R., Kwong, K. K., Jiang, H., Kennedy, W. A., Ledden, P. J., Brady, T. J., Rosen, B. R., and Tootell, R. B. (1995). Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proc. Natl. Acad. Sci. U.S.A.* 92, 8135–8139.

Mars, R. B., Klein, M. C., Neubert, F. X., Olivier, E., Buch, E. R., Boorman, E. D., and Rushworth, M. F. (2009). Short-latency influence of medial frontal cortex on primary motor cortex during action selection under conflict. *J. Neurosci.* 29, 6926–6931.

Matelli, M., and Luppino, G. (2000). Parietofrontal circuits: parallel channels for sensory-motor integrations. *Adv. Neurol.* 84, 51–61.

McIntosh, A. R., and Gonzalez-Lima, F. (1991). Structural modeling of functional neural pathways mapped with 2-deoxyglucose: effects of acoustic startle habituation on the auditory system. *Brain Res.* 547, 295–302.

Meister, I. G., Wilson, S. M., Deblieck, C., Wu, A. D., and Iacoboni, M. (2007). The essential role of premotor cortex in speech perception. *Curr. Biol.* 17, 1692–1696.

Miniussi, C., Ruzzoli, M., and Walsh, V. (2010). The mechanism of transcranial magnetic stimulation in cognition. *Cortex* 46, 128–130.

Mochizuki, H., Huang, Y. Z., and Rothwell, J. C. (2004). Interhemispheric interaction between human dorsal premotor and contralateral primary motor cortex. *J. Physiol.* 561, 331–338.

Mochizuki, H., Franca, M., Huang, Y. Z., and Rothwell, J. C. (2005). The role of dorsal premotor area in reaction task: comparing the "virtual lesion" effect of paired pulse or theta burst transcranial magnetic stimulation. *Exp. Brain Res.* 167, 414–421.

Murase, N., Duque, J., Mazzocchio, R., and Cohen, L. G. (2004). Influence of interhemispheric interactions on motor function in chronic stroke. *Ann. Neurol.* 55, 400–409.

Nowak, D. A., Grefkes, C., Ameli, M., and Fink, G. R. (2009). Interhemispheric competition after stroke: brain stimulation to enhance recovery of function of the affected hand. *Neurorehabil. Neural. Repair* 23, 641–656.

Nowak, D. A., Grefkes, C., Dafotakis, M., Eickhoff, S., Kust, J., Karbe, H., and Fink, G. R. (2008). Effects of low-frequency repetitive transcranial magnetic stimulation of the contralesional primary motor cortex on movement kinematics and neural activity in subcortical stroke. *Arch. Neurol.* 65, 741–747.

Obhi, S. S., Haggard, P., Taylor, J., and Pascual-Leone, A. (2002). rTMS to the supplementary motor area disrupts bimanual coordination. *Motor. Control.* 6, 319–332.

Oliveri, M., Babiloni, C., Filippi, M. M., Caltagirone, C., Babiloni, F., Cicinelli, P., Traversa, R., Palmieri, M. G., and Rossini, P. M. (2003). Influence of the supplementary motor area on primary motor cortex excitability during movements triggered by neutral or emotionally unpleasant visual cues. *Exp. Brain Res.* 149, 214–221.

Op De Beeck, H., and Vogels, R. (2000). Spatial sensitivity of macaque inferior temporal neurons. *J. Comp. Neurol.* 426, 505–518.

O'Shea, J., Johansen-Berg, H., Trief, D., Gobel, S., and Rushworth, M. F. (2007a). Functionally specific reorganization in human premotor cortex. *Neuron* 54, 479–490.

O'Shea, J., Sebastian, C., Boorman, E. D., Johansen-Berg, H., and Rushworth, M. F. (2007b). Functional specificity of human premotor-motor cortical interactions during action selection. *Eur. J. Neurosci.* 26, 2085–2095.

Paus, T. (2005). Inferring causality in brain images: a perturbation approach. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* 360, 1109–1114.

Paus, T., Jech, R., Thompson, C. J., Comeau, R., Peters, T., and Evans, A. C. (1997). Transcranial magnetic stimulation during positron emission

tomography: a new method for studying connectivity of the human cerebral cortex. *J. Neurosci.* 17, 3178–3184.

Paus, T., Jech, R., Thompson, C. J., Comeau, R., Peters, T., and Evans, A. C. (1998). Dose-dependent reduction of cerebral blood flow during rapid-rate transcranial magnetic stimulation of the human sensorimotor cortex. *J. Neurophysiol.* 79, 1102–1107.

Perez, M. A., Tanaka, S., Wise, S. P., Willingham, D. T., and Cohen, L. G. (2008). Time-specific contribution of the supplementary motor area to intermanual transfer of procedural knowledge. *J. Neurosci.* 28, 9664–9669.

Picard, N., and Strick, P. L. (1996). Motor areas of the medial wall: a review of their location and functional activation. *Cereb. Cortex* 6, 342–353.

Picard, N., and Strick, P. L. (2001). Imaging the premotor areas. *Curr. Opin. Neurobiol.* 11, 663–672.

Pollok, B., Rothkegel, H., Schnitzler, A., Paulus, W., and Lang, N. (2008). The effect of rTMS over left and right dorsolateral premotor cortex on movement timing of either hand. *Eur. J. Neurosci.* 27, 757–764.

Praamstra, P., Kleine, B. U., and Schnitzler, A. (1999). Magnetic stimulation of the dorsal premotor cortex modulates the Simon effect. *Neuroreport* 10, 3671–3674.

Reijneveld, J. C., Ponten, S. C., Berendse, H. W., and Stam, C. J. (2007). The application of graph theoretical analysis to complex networks in the brain. *Clin Neurophysiol* 118, 2317–2331.

Ridding, M. C., and Rothwell, J. C. (2007). Therapeutic use of rTMS. *Nat. Rev. Neurosci.* 8, 559–567.

Rizzolatti, G., and Luppino, G. (2001). The cortical motor system. *Neuron* 31, 889–901.

Roland, P. E., Larsen, B., Lassen, N. A., and Skinhoj, E. (1980). Supplementary motor area and other cortical areas in organization of voluntary movements in man. *J. Neurophysiol.* 43, 118–136.

Rossi, S., Hallett, M., Rossini, P. M., and Pascual-Leone, A. (2009). Safety, ethical considerations, and application guidelines for the use of transcranial magnetic stimulation in clinical practice and research. *Clin Neurophysiol* 120, 2008–2039.

Rushworth, M. F., Hadland, K. A., Paus, T., and Sipila, P. K. (2002). Role of the human medial frontal cortex in task switching: a combined fMRI and TMS study. *J. Neurophysiol.* 87, 2577–2592.

Sack, A. T., Cohen Kadosh, R., Schuhmann, T., Moerel, M., Walsh, V., and Goebel, R. (2009). Optimizing functional accuracy of TMS in cognitive studies: a comparison of methods. *J. Cogn. Neurosci.* 21, 207–221.

Sato, M., Tremblay, P., and Gracco, V. L. (2009). A mediating role of the premotor cortex in phoneme segmentation. *Brain Lang.* 111, 1–7.

Schlaghecken, F., Munchau, A., Bloem, B. R., Rothwell, J., and Eimer, M. (2003). Slow frequency repetitive transcranial magnetic stimulation affects reaction times, but not priming effects, in a masked prime task. *Clin. Neurophysiol.* 114, 1272–1277.

Schluter, N. D., Rushworth, M. F., Mills, K. R., and Passingham, R. E. (1999). Signal-, set-, and movement-related activity in the human premotor cortex. *Neuropsychologia* 37, 233–243.

Schluter, N. D., Rushworth, M. F., Passingham, R. E., and Mills, K. R. (1998). Temporary interference in human lateral premotor cortex suggests dominance for the selection of movements. A study using transcranial magnetic stimulation. *Brain* 121, 785–799.

Serrien, D. J., Strens, L. H., Oliviero, A., and Brown, P. (2002). Repetitive transcranial magnetic stimulation of the supplementary motor area (SMA) degrades bimanual movement control in humans. *Neurosci. Lett.* 328, 89–92.

Siebner, H. R., Filipovic, S. R., Rowe, J. B., Cordivari, C., Gerschlager, W., Rothwell, J. C., Frackowiak, R. S., and Bhatia, K. P. (2003). Patients with focal arm dystonia have increased sensitivity to slow-frequency repetitive TMS of the dorsal premotor cortex. *Brain* 126, 2710–2725.

Siebner, H. R., Hartwigsen, G., Kassuba, T., and Rothwell, J. C. (2009). How does transcranial magnetic stimulation modify neuronal activity in the brain? Implications for studies of cognition. *Cortex* 45, 1035–1042.

Siebner, H. R., Willoch, F., Peller, M., Auer, C., Boecker, H., Conrad, B., and Bartenstein, P. (1998). Imaging brain activation induced by long trains of repetitive transcranial magnetic stimulation. *Neuroreport* 9, 943–948.

Silvanto, J., Muggleton, N. G., Cowey, A., and Walsh, V. (2007). Neural adaptation reveals state-dependent effects of transcranial magnetic stimulation. *Eur. J. Neurosci.* 25, 1874–1881.

Steyvers, M., Etoh, S., Sauner, D., Levin, O., Siebner, H. R., Swinnen, S. P., and Rothwell, J. C. (2003). High-frequency transcranial magnetic stimulation of the supplementary motor area reduces bimanual coupling during anti-phase but not in-phase movements. *Exp. Brain Res.* 151, 309–317.

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.

Taub, E., Uswatte, G., and Elbert, T. (2002). New treatments in neurorehabilitation founded on basic research. *Nat. Rev. Neurosci.* 3, 228–236.

Taubert, M., Dafotakis, M., Sparing, R., Eickhoff, S., Leuchte, S., Fink, G. R., and Nowak, D. A. (2010). Inhibition of the anterior intraparietal area and the dorsal premotor cortex interfere with arbitrary visuo-motor mapping. *Clin. Neurophysiol.* 121, 408–413.

Taylor, P. C., Nobre, A. C., and Rushworth, M. F. (2007). Subsecond changes in top down control exerted by human medial frontal cortex during conflict and action selection: a combined transcranial magnetic stimulation electroencephalography study. *J. Neurosci.* 27, 11343–11353.

Thielscher, A., and Kammer, T. (2004). Electric field properties of two commercial figure-8 coils in TMS: calculation of focality and efficiency. *Clin. Neurophysiol.* 115, 1697–1708.

Tremblay, P., and Gracco, V. L. (2009). Contribution of the pre-SMA to the production of words and non-speech oral motor gestures, as revealed by repetitive transcranial magnetic stimulation (rTMS). *Brain Res.* 1268, 112–124.

Tunik, E., Lo, O. Y., and Adamovich, S. V. (2008). Transcranial magnetic stimulation to the frontal operculum and supramarginal gyrus disrupts planning of outcome-based hand-object interactions. *J. Neurosci.* 28, 14422–14427.

Umilta, M.A., Kohler, E., Gallese, V., Fogassi, L., Fadiga, L., Keysers, C., and Rizzolatti, G. (2001). I know what you are doing a neurophysiological study. *Neuron.* 31, 155–165.

Urgesi, C., Calvo-Merino, B., Haggard, P., and Aglioti, S. M. (2007a). Transcranial magnetic stimulation reveals two cortical pathways for visual body processing. *J. Neurosci.* 27, 8023–8030.

Urgesi, C., Candidi, M., Ionta, S., and Aglioti, S. M. (2007b). Representation of body identity and body actions in extrastriate body area and ventral premotor cortex. *Nat. Neurosci.* 10, 30–31.

van den Berg, F. E., Swinnen, S. P., and Wenderoth, N. (in press). Hemispheric asymmetries of the premotor cortex are task specific as revealed by disruptive TMS during bimanual versus unimanual movements. *Cereb. Cortex.* doi:10.1093/cercor/bhq034.

van Donkelaar, P., Lee, J. H., and Drew, A. S. (2002). Eye-hand interactions differ in the human premotor and parietal cortices. *Hum. Mov. Sci.* 21, 377–386.

Verwey, W. B. (1996). Buffer loading and chunking in sequential keypressing. *J. Exp. Psychol. Hum. Percept. Perform.* 22, 544–562.

Verwey, W. B. (1999). Evidence for a multistage model of practice in a sequential movement task. *J. Exp. Psychol. Hum. Percept Perform* 25, 1–16.

Verwey, W. B., Lammens, R., and van Honk, J. (2002). On the role of the SMA in the discrete sequence production task: a TMS study. *Transcranial Magnetic Stimulation. Neuropsychologia* 40, 1268–1276.

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

*Received: 25 May 2010; paper pending published: 14 June 2010; accepted: 13 August 2010; published online: 19 October 2010.*

*Citation: Chouinard PA and Paus T (2010) What have we learned from "perturbing" the human cortical motor system with transcranial magnetic stimulation? Front. Hum. Neurosci. **4**:173. doi:10.3389/fnhum.2010.00173.*

frontiers in
HUMAN NEUROSCIENCE

HYPOTHESIS AND THEORY ARTICLE
published: 22 November 2010
doi: 10.3389/fnhum.2010.00211

# Natural memory beyond the storage model: repression, trauma, and the construction of a personal past

***Nikolai Axmacher[1]*, Anne T. A. Do Lam[1], Henrik Kessler[2,3] and Juergen Fell[1]***

[1] *Department of Epileptology, University of Bonn, Bonn, Germany*
[2] *Department of Medical Psychology, University of Bonn, Bonn, Germany*
[3] *Department of Medical Psychology, University of Ulm, Ulm, Germany*

***Edited by:***
*Michael X. Cohen, University of Amsterdam, Netherlands*

***Reviewed by:***
*Robert Blumenfeld, University of California at Berkeley, USA*
*Patrick Khader, Philipps University, Germany*

***Correspondence:**
*Nikolai Axmacher, Department of Epileptology, University of Bonn, Sigmund-Freud-Str. 25, 53105 Bonn, Germany.*
*e-mail: nikolai.axmacher@ukb.uni-bonn.de*

Naturally occurring memory processes show features which are difficult to investigate by conventional cognitive neuroscience paradigms. Distortions of memory for problematic contents are described both by psychoanalysis (internal conflicts) and research on post-traumatic stress disorder (PTSD; external traumata). Typically, declarative memory for these contents is impaired – possibly due to repression in the case of internal conflicts or due to dissociation in the case of external traumata – but they continue to exert an unconscious pathological influence: neurotic symptoms or psychosomatic disorders after repression or flashbacks and intrusions in PTSD after dissociation. Several experimental paradigms aim at investigating repression in healthy control subjects. We argue that these paradigms do not adequately operationalize the clinical process of repression, because they rely on an intentional inhibition of random stimuli (suppression). Furthermore, these paradigms ignore that memory distortions due to repression or dissociation are most accurately characterized by a lack of self-referential processing, resulting in an impaired integration of these contents into the self. This aspect of repression and dissociation cannot be captured by the concept of memory as a storage device which is usually employed in the cognitive neurosciences. It can only be assessed within the framework of a constructivist memory concept, according to which successful memory involves a reconstruction of experiences such that they fit into a representation of the self. We suggest several experimental paradigms that allow for the investigation of the neural correlates of repressed memories and trauma-induced memory distortions based on a constructivist memory concept.

**Keywords: repression, memory distortions, PTSD, constructive memory, psychoanalysis, self-referential processing**

## INTRODUCTION

The current cognitive neurosciences have already revealed some mechanisms underlying the modulation of memory performance. However, several phenomena occurring in natural memory like memory distortions due to repression and due to trauma-related dissociations have not been adequately addressed yet by neurocognitive research.

Many studies have shown convincingly that emotionally arousing stimuli are better remembered than neutral images (Heuer and Reisberg, 1990; Bradley et al., 1992; Christianson, 1992; Ochsner, 2000; Buchanan and Lovallo, 2001; Kensinger and Corkin, 2003). This effect depends on a facilitation of hippocampus-dependent memory processes by the amygdala (Seidenbecher et al., 2003; Dolcos et al., 2004; Kensinger and Corkin, 2004; Phelps, 2004; Smith et al., 2006; **Figure 1A**) and is related to the action of glucocorticoids (Kim and Diamond, 2002; Sapolsky, 2003). Such an interaction between amygdala and hippocampus may be necessary for the enhanced declarative memory of emotional events because these two structures support complementary processes, as revealed by a hippocampal-amygdala double dissociation: While integrity of the hippocampus is necessary for the conscious memory that a particular stimulus was, during conditioning, associated with a shock, the amygdala is required for the unconsciously associated vegetative reaction (Bechara et al., 1995; LaBar et al., 1995).

While these studies provide compelling evidence for an enhanced memory of stimuli which induce moderately negative emotions, this is not necessarily true for two problematic cases involving extremely negative emotions: the emergence of an unconscious conflict, which is subject to *repression*, and traumatic events that overstress a person's executive capabilities and thus lead to *dissociation*. As a result, conscious recall of these contents is impaired, but they continue to exert an unconscious effect which dramatically influences subsequent life – for example, by uncontrollably occurring intrusions and dissociative flashbacks, panic attacks, or psychosomatic symptoms (see **Box 1**). In other instances, victims of a traumatic experience may be able to recall details from the trauma, but only in a contorted manner – for example, from a detached view outside of themselves, or without the associated emotions. These symptoms of people which have suffered from real traumatic experiences are subsumed under the diagnosis of post-traumatic stress disorder (PTSD; F43.1, ICD-10, World Health Organization, 1992; DSM-IV-TR, American Psychiatric Association, 2000; Elbert and Schauer, 2002; Maercker, 2009).

Therapeutic interventions on patients suffering from symptoms due to repressed conflicts or traumatic experiences require an understanding of the mechanisms of repression and dissociation, not only on the psychological but also on the neurophysiological level (for the general benefit of cognitive neuroscience

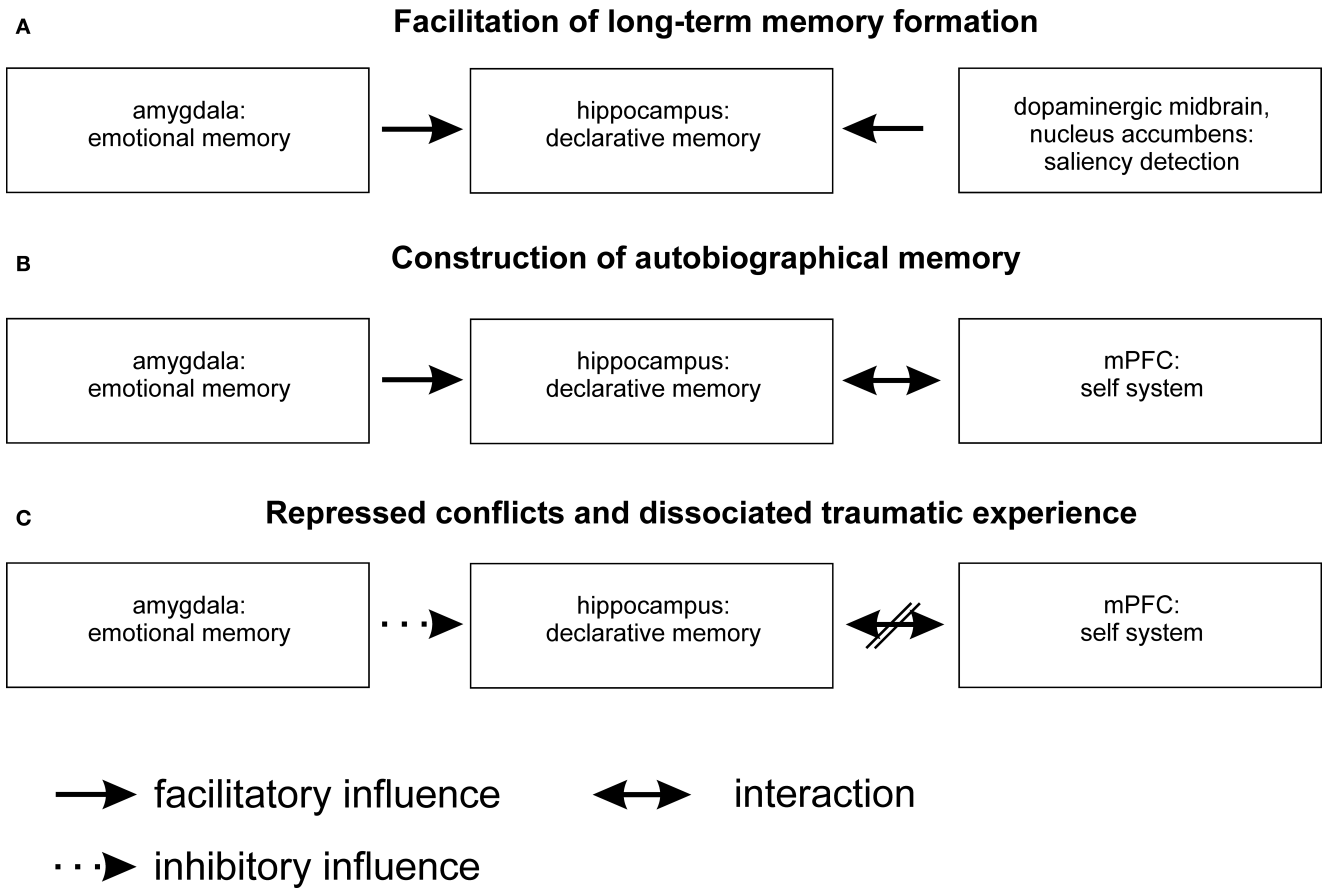



**FIGURE 1 | Emotional effects on memory. (A)** Storage model of memory: Moderately arousing emotional events are better encoded than neutral events. **(B)** The narrative construction of a personal past relies on an interaction between the hippocampus and self-referential processes in the medial prefrontal cortex (mPFC). This process is facilitated by (moderately) emotional events due to inputs from the amygdala. **(C)** In contrast, in the case of extremely negative (conflictual or traumatic) experiences, the amygdala inhibits declarative memory formation by the hippocampus and an integration of these memories with the self-image.

**BOX 1 | Memory distortions for problematic contents – repression and dissociation.**

There are at least two distinct processes leading to memory distortions in the case of problematic contents involving extremely negative emotions: repression and dissociation. Repression is the process by which *internal* conflicts are stored in the unconscious. Dissociation, on the other hand, is the process by which parts of *external* traumatic events are stored in a non-declarative memory system (see below). *Repression* is a typical defense mechanism thoroughly described by Freud to explain clinical symptoms such as neurotic depression or psychosomatic symptoms he observed in his patients. According to Freud, the starting point is an internal conflict arising when (mostly unconscious) wishes or drives are in a strict opposition to internalized norms or standards. If this conflict cannot be solved (e.g., because opposing elements cannot be integrated with self-referential processes), it automatically produces intense anxiety signaling danger for the subject ("Signalangst"). In an effort to avoid extremely negative emotions, the entire conflict and its associated emotions and memories are pushed into the unconscious (repression). This leads to a lack of declarative memory for the conflict and often the circumstances under which it emerged. The conflictual material itself, however, continues to exist in the unconscious and, more importantly, exerts a major influence on the subject by causing neurotic (e.g., depressed mood) or psychosomatic (e.g., paralyses) symptoms. According to Freud, the symptom is a symbolization of the internal conflict (for details, see Person et al., 2005). *Dissociation* is a process mainly investigated in the context of PTSD. When external traumata involving extremely negative emotions cannot be integrated with self-referential processes and no coherent narrative can be built, memories of this trauma become dissociated, i.e., they are stored in a system with no direct verbal access. Brewin (2001, 2003) uses the term "situationally accessible memory" to denote the memory system where such dissociated elements are stored. Verbal memories of the traumatic event are often vague and include gaps. The contents stored in the SAM, on the other hand, are the source for situationally triggered and hence not controllable intrusions and flashbacks typical for PTSD. Although distinct from the *process* of dissociation described here on a conceptual level, the clinical phenomena of "dissociations" in the traumatic situation (e.g., depersonalization, derealization) are empirically linked to the dissociation of memories in the SAM and hence the eventual development of a PTSD (Brewin and Holmes, 2003).

for understanding cognition see Henson, 2005; Axmacher et al., 2009). A number of studies investigated brain regions activated during presentation of trauma-related cues in PTSD patients by functional MRI (for a review, see Shin et al., 2005). Other studies used MEG recordings to explore the neurophysiological basis of the "trauma network" in PTSD. Elbert and colleagues found that in these patients, processing of arousing stimuli (e.g., IAPS pictures) relied more on a fast sensory processing pathway, which is uncoupled from prefrontal control, and less on elaborate processing along the ventral visual stream (Rockstroh and Elbert, 2010). This may be due to a reduced connectivity with neural assemblies representing context-related ("cold") and trauma-related ("hot") information; as a result, trauma victims are unable to locate their trauma memories in time and space and experience them as flashbacks (Elbert and Schauer, 2002). Psychotherapeutical treatment of PTSD patients (e.g., by narrative exposure therapy; Neuner et al., 2004) aims at re-integrating this network by placing the traumatic memories in a coherent context. However, despite the incontestable value of such studies directly investigating PTSD patients, they are necessarily retrospective and do not allow for an experimental control of the emotionally disturbing situation itself. Several experimental paradigms have been developed to capture the processes occurring during PTSD-*like* memory distortions in healthy control subjects.

In this article, we will first describe memory distortions due to repression of internal conflicts (Memory Distortions Due to Repression of an Internal Conflict) and due to dissociation of external traumata (Memory Distortions Due to Dissociation After an External Trauma). In the Section "Experimental Paradigms of Repression," we provide an overview of the experimental paradigms currently used in the cognitive neurosciences to study the neural correlates of repression in healthy volunteers. We will argue that these paradigms do not convincingly operationalize repression because they fall short of capturing key aspects of this clinical phenomenon. Next, we will broaden our scope and demonstrate that the concept of memory itself usually implied in cognitive neuroscience studies represents only one aspect of real-life memory, but disregards others and is therefore unable to capture several complex features of memory (Storage and Constructive Models of Memory): The criteria for successful memory retrieval – namely accurate recollections of particular events – and the corresponding measures in the cognitive neurosciences such as analyses of subsequent memory effects only capture one relevant dimension of memory but disregard its constructive aspects. This becomes most obvious in studies of autobiographical memory, which involves integration of experiences with self-referential processes (**Figure 1B**). In the Section "Trauma-Related Memory Distortions Due to Lack of Self-Referential Processing," we will argue that trauma-related memory distortions can only be adequately understood (and experimentally investigated) when they are conceptualized as a failure to construct autobiographical memories, i.e., to integrate these experiences with self-referential processes (**Figure 1C**). In other words, these memory distortions cannot be adequately understood as a failure to *store* memories, but as a failure to *integrate* them with self-referential processes. We will present a paradigm that takes these considerations into account. Finally (Promising Approaches for Studying Memory Distortions Due to Repression), we will suggest a paradigm for the investigation of memory distortions due to repression.

## MEMORY DISTORTIONS DUE TO REPRESSION OF AN INTERNAL CONFLICT

The concept of repression was originally suggested by Herbart (1824), but was introduced as a pathological process by Sigmund Freud (see **Box 1**). Most of his patients had symptoms that could not be explained by common logic or medical knowledge (e.g., paralyses of isolated limbs with no medical cause). In order to explain the etiology of such symptoms, he conceptualized the construct of repression as a mechanism being applied throughout child development. It means the storage of complete internal conflicts (mostly between drives or wishes and internalized norms or standards) and their surrounding emotions and memories in the unconscious (Freud, 1915). The developing child hence represses problematic content once an internal conflict is emergent. The classical view of repression as depicted in **Box 1** was significantly advanced by an increased consideration of the child's relationship to its parents and its developing self (Ferenczi, 1933; Balint, 1969). Internalization of a safe relationship to the parents is necessary for the development of agency and a stable self (Stern, 1985), and the infantile self is only developed through such parental feedback (Fonagy et al., 2005). In this new view, repression may occur when an internal conflict cannot be integrated with the image of the parents and/or self-referential processes. It is hence mainly this problematic integration in self-referential processes that give conflicts their true "pathological" value and trigger repression as a means to cope with them for the moment. The price to pay for the relative peace the developing self obtains by repressing internal conflicts is high, though: Clinically, it is important that the repressed material still exists in the unconscious and exerts a large influence on the subject by causing psychosomatic (e.g., paralyses) or neurotic (e.g., depressed mood) symptoms. The mechanism through which repressed conflicts gain access to the "outside" in the form of symptoms is called conversion (Breuer and Freud, 1895). By symbolizing the conflict or parts thereof in a symptomatic language, the patient can gain some relief at the price of clinical symptoms.

Although rooted in early twentieth century psychology and physiology, the concepts of repression and conversion via symbolization are still of high clinical relevance for recent models of psychotherapy (for a current clinical textbook see Person et al., 2005). Psychoanalytic therapies primarily working with the concept of repression are highly effective in treating patients with severe psychiatric problems (e.g., chronic depression, personality disorders), as reviewed in a recent meta-analysis (Leichsenring and Rabung, 2008). The recent trend that problematic integration in self-referential processes is regarded as the main virulent component of repressed conflicts is reflected in newer manualized psychodynamic treatments (e.g., Wöller and Kruse, 2010): The therapist should explicitly help the patient reorganize past experiences in a way that repressed conflicts can be integrated in self-referential processes. If successful, the patient not only gains clinical improvement on a symptomatic level (less or no need for conversion of an internal conflict in body symptoms), but also achieves a new and more coherent view of the self.

## MEMORY DISTORTIONS DUE TO DISSOCIATION AFTER AN EXTERNAL TRAUMA

In contrast to the psychoanalytical focus on internal conflicts early in development, the concept of PTSD was primarily developed with regard to external later-life traumata, initially those during the Vietnam war. Memory distortions are, however, also among the main symptoms of PTSD. On the one hand, memories of the traumatic situation occur involuntarily as intrusions or even as flashbacks, in which patients do not remember the traumatic event as something past, but re-experience it as if it occurred again, similar to a dissociative state. On the other hand, declarative memory for details of the traumatic situation is often impaired (Brewin, 2007; Jones et al., 2007): A large body of studies have reported general declarative memory deficits in PTSD patients (Clancy et al., 2000; Behrendt and Moritz, 2005; Jelinek et al., 2006; for a review, see Brewin, 2007), and in particular concerning verbal declarative memory (Yehuda et al., 1995; Bremner et al., 2004). These memory impairments emphasize initial encoding and retention of new contents (Bremner, 2002), for instance, long lasting deficits in short-term memory have been shown in adults suffering from abuse during childhood (Bremner et al., 1995) as well as in Vietnam veterans diagnosed with PTSD (Bremner et al., 1993). Additionally, source memory referring to specific details of the traumatic event is often distorted in this population (Johnson et al., 1993). However, as we will elaborate below, the specific memory deformations following a trauma can be most accurately conceptualized not as failures to recall specific information, but as an impairment to integrate these experiences with self-referential processes, i.e., in the framework of constructivist memory theories (Trauma-Related Memory Distortions Due to Lack of Self-Referential Processing).

During the traumatic experience, limited time for conscious processing of accompanying sensory or vegetative perceptions inhibits an adequate integration thereof into autobiographical memory. This effect may depend on two mechanisms. First, an impaired processing of peripheral details during a traumatic situation due to a narrowing of the spotlight of attention is known as the *weapon focusing* effect (Christianson, 1992). This effect reduces the integration of all trauma-related information into a coherent representation. Second, traumatic experiences may be so overwhelming that the executive processing capabilities of trauma victims fail, leading to peri-traumatic dissociative states. In these states, subjects describe that they view themselves from a detached standpoint (e.g., Bremner and Brett, 1997; Lanius et al., 2002). The dual representation theory (Brewin, 2003) accounts for these trauma-related memory distortions by suggesting that memory contents can be encoded either via a hippocampus-dependent mechanism, which leads to a narrative integration of these contents with other experiences and makes them accessible for declarative recall (*verbally accessible memory*, VAM), or via an amygdala-dependent process, which does not allow for a conscious control of memory retrieval; however, unconscious experiences encoded by the amygdala are automatically recalled whenever an associated cue appears. According to the dual representation theory, this latter memory system has been described as *situationally accessible memory* (SAM). This theory further suggests that amygdala and hippocampus inhibit each other in the case of extremely negative events, such that the degree to which the amygdala supports encoding of an event reduces the hippocampal contribution (Metcalfe and Jacobs, 1998; Brewin, 2001; **Figure 1C**). Indeed, functional MRI studies in PTSD patients indicate that recall of traumatic events is associated with an increased activation of the amygdala and a reduced activation of the hippocampus (Shin et al., 2005). Alternatively, it is possible that in this case the amygdala facilitates unconscious memory processes in the hippocampus (Henke et al., 2003; Degonda et al., 2005).

Although the memory distortions described in psychoanalysis (repression) and in PTSD research (memory fragmentations and intrusions due to dissociation) appear to be very different at first sight, they converge in the idea that pathological forms of unconscious memory replace declarative access to the problematic contents (**Figure 2**). In both cases, contents that are not verbally accessible (repressed conflicts or situationally accessible memories) exert a pathological influence on the subject by causing psychosomatic symptoms or intrusions and flashbacks, respectively. Thus, unconscious memories are created that do not refer to familiarity-based recognition memory, but to memories that induce intrusions and flashbacks following a trauma or neurotic symptoms following a repressed conflict. As will be elaborated below, both phenomena gain their true pathological value by a failure to integrate the problematic contents (internal conflicts or traumata) in self-referential processes. Finally, successful treatment of both includes the integration of problematic contents in a more coherent self. Therefore, we suggest that a common basis, or at least an overlap, exists for these two groups of phenomena, which might be defined in neurobiological

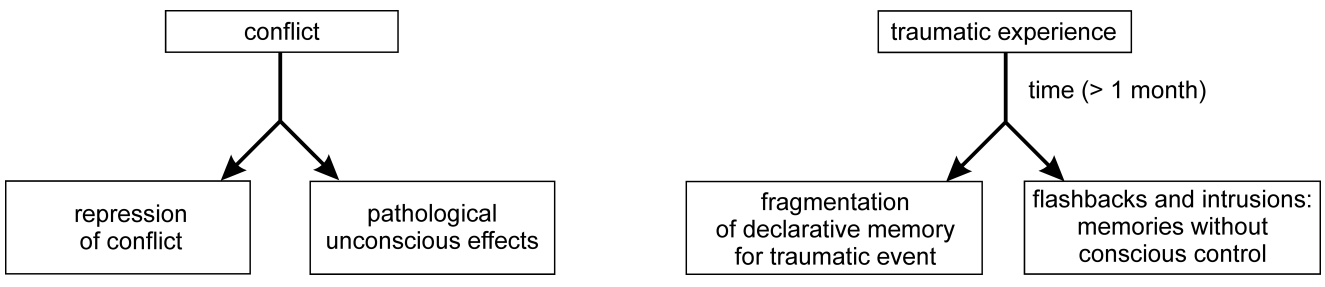



**FIGURE 2 | Memory distortions caused by repressed conflicts (left) and traumatic experiences (right).** Both psychoanalysis and PTSD theories suggest that extremely negative contents may cause memory distortions characterized by an impaired declarative memory for these events but pathological unconscious memories. The theories differ with respect to the relative contribution of the internal vs. external nature of the negative contents.

terms – possibly a modern replacement of the Freudian "metapsychology," which was deeply grounded in late nineteenth century neurophysiological knowledge (e.g., Stephan, 2002).

## EXPERIMENTAL PARADIGMS OF REPRESSION

Cognitive neuroscience studies on repression have established two main experimental paradigms which aim at investigating repression in normal healthy subjects, the "Directed forgetting" and the "Think/No-Think" paradigm (Johnson, 1994; Anderson and Green, 2001; Erdelyi, 2006). In these paradigms, forgetting is consciously and intentionally controlled by the participants, who are explicitly asked to voluntarily inhibit randomly selected subsets of items or "not to think" of them: "On some trials they (i.e., the participants) were instructed to think of the previously learned picture; on other trials they were instructed not to let the previously associated picture enter consciousness." (Think/No-Think paradigm, Depue et al., 2007). One might wonder whether these paradigms are actually aimed at investigating the psychoanalytical process of repression. However, this is explicitly stated; for example, the article by Anderson et al. (2004) using the Think/No-Think paradigm starts with the phrases: "Over a century ago, Freud proposed that unwanted memories can be excluded from awareness, a process called repression. It is unknown, however, how repression occurs in the brain." (p. 232). In both paradigms, the instructions reliably lead to a decreased proportion of stimuli in the "voluntarily forgotten" condition which can be recalled in a subsequent declarative memory test. The functional MRI results show a reduced BOLD response in the hippocampus and an increased activity in the lateral prefrontal cortex, which was interpreted as index for a recruitment of inhibitory executive control processes. However, for three reasons we argue that these paradigms do not adequately operationalize the clinical process of repression. First, repression occurs specifically in situations which overload the executive processing capacities. During real repression, voluntary control is lacking; a person represses an experience because this experience induces an unbearable conflict. Therefore, attempting to induce repression by a recruitment of executive control processes is paradoxical. The process investigated by Anderson and colleagues can more precisely be termed "suppression," the voluntary "forgetting," or "keeping down" of unwanted content. Second, if the mechanism of repression is considered in isolation, it appears as if the contents of repression are only secondary. However, repression does not occur with regard to random situations or stimuli, but only if an intense negative emotion is evoked. Thus, repression should be automatically induced by the experimental stimuli. Third, an experimental paradigm of repression should not only reduce conscious access to "repressed" stimuli, but also exclude that these stimuli are just forgotten. In fact, it should even be shown that unconscious memory for these items is enhanced: The concept of repression was introduced to explain clinical symptoms by experiences which are not consciously accessible for the patients, but which continue to exert an unconscious influence. Thus, in a non-declarative, implicit memory test, an increased proportion of these stimuli should be "remembered," as shown in PTSD patients (e.g., McNally, 1997). For example, these stimuli may be erroneously classified as "new" during conscious recollection (because no conscious memory exists for them), but this (incorrect) response may be given with a delayed reaction time as compared to stimuli which are in fact new, which could be interpreted as indirect evidence of an unconscious conflict during processing of these stimuli.

Alternative paradigms can take these critiques into account. In such paradigms, the process of repression should not be voluntarily controlled; instead, they should create a situation in which repression occurs automatically due to the emotional content of a stimulus. When the "repressed" stimuli are subsequently presented in a recognition memory test, it is hypothesized that they cannot be recalled consciously, but that they continue to exert an unconscious influence on the subject's behavior, as indicated by a measure of unconscious memory. In addition, however, there is a more general problem with the existing paradigms that are meant to operationalize repression, because they measure memory as an impairment of successful recall, but do not investigate whether memories are successfully integrated into the self. In other words, the constructive nature of successful memories is not taken into account. In the next two sections, we will explain such a constructivist concept of memory, which is particularly relevant for autobiographical memorizing (Storage and Constructive Models of Memory) and is impaired in trauma-related memory distortions (Trauma-Related Memory Distortions Due to Lack of Self-Referential Processing). Finally, we will describe paradigms which are suited to capture memory distortions related to repression of conflicts (Promising Approaches for Studying Memory Distortions Due to Repression).

## STORAGE AND CONSTRUCTIVE MODELS OF MEMORY

In the cognitive neurosciences, memory is most often conceptualized as the process of stimulus encoding, storage, and retrieval. In this framework, memory recall involves an identical repetition of the original experiences – or, as Tulving (1983) described it in his famous definition of episodic memory, as a "mental time travel" back to the original situation. Accordingly, the neural correlates of memory are usually studied by analyzing "subsequent memory"-effects, i.e., differences in brain activation patterns associated with the initial presentation of subsequently remembered as compared to forgotten items. However, this storage model of memory only captures *one* relevant dimension of memory, which relies on the identity of encoded and remembered contents. In particular, it abstracts the specific function of memory retrieval for a subject. The cultural history of memory models has always consisted in a dichotomy of such storage models of memory on the one hand, models of constructive memory on the other hand (Assmann, 2002). Constructive concepts of memory as re-interpretation of events were, e.g., investigated experimentally by Bartlett (1932). In a number of studies, he presented his British subjects a short story, "The War of the Ghosts," which contains several seemingly illogical and irrational elements. When subjects were afterward asked to recall the story in as many details as possible, they modified it according to their own cultural schemata; illogical elements were thus replaced by more coherent narratives. These studies illustrate the constructive nature of memory retrieval and suggest that memory is not only designed to retrieve events exactly as they happened, but supports specific functions in the interaction with the (internal or external) environment.

Several current memory theories have incorporated aspects of such a constructive account on memory and also highlighted the relevance of self-referential processing. According to these views, it is not exact representations of external events that are encoded and retrieved, but the results from internal processing, evaluation, and interpretation of these events. For example, the theory of transfer-appropriate processing (TAP) interprets learning as re-performing a previous act, i.e., successful learning is an appropriate transfer of underlying structures rather than an access of a memory trace (Morris et al., 1977; Bransford et al., 1979; Baddeley, 2002). According to TAP, recall is facilitated by similarity between the encoding and the retrieval state; this effect is related to both environmental contexts (e.g., if encoding occurred in the library of the Psychology Department, recall will also be easier at that place) and internal variables such as current affective states and goals (events encoded in a sad mood will also be most likely retrieved in a sad mood). In addition, many studies investigated the effects of encoding conditions related to self-referential processing, e.g., by instructing subjects to rate the pleasantness of items (Hunt and Einstein, 1981; McDaniel and Einstein, 1993). Such effects have been described, e.g., in the framework of the levels of processing theory (Craik and Lockhardt, 1972). Furthermore, several experimental paradigms have investigated why events which have never happened are "recalled" (the false memory paradigm; see, e.g., Deese, 1959; Roediger III and McDermott, 1995), and why actual events are forgotten (e.g., retrieval-induced forgetting paradigms, Anderson et al., 1994; Macrae and MacLeod, 1999; Caroll et al., 2007). None of these theories and paradigms assumes a simple storage model of memory. Instead, they capture aspects of a constructive memory theory because they emphasize that memory encoding and retrieval depend on the construction of an integrated experience and the evaluation of its personal relevance. Furthermore, recall can be considered a "reconstruction" because it depends on the situational context during which it occurs.

However, there is a fundamental difference between these accounts and constructivist memory theories in a narrower sense, which is the criterion for successful memory: In storage models of memory, recall is successful if it recapitulates crucial aspects of a previous experience; according to constructivist memory theories, however, recall is successful if it allows the recalling person to build a coherent narrative about his/her past. Imagine that two people experience the same event, but afterward report two inconsistent versions of this event. According to the storage model of memory, there are objective criteria whether the report of the first or the second person is correct. In contrast, the constructive memory theory would assume that both reports may be correct as long as subjective coherence is achieved. This view differs considerably from paradigms such as the false memory paradigm, the retrieval-induced forgetting paradigm, or paradigms of voluntary memory suppression such as the "think/no-think" or "directed forgetting" paradigms: These paradigms investigate the conditions under which memory storage fails, but not the conditions under which a successful transformation into an acceptable narrative, and integration with self-referential processes, succeed. Similarly, even though theories such as the TAP theory or models such as the Adaptive Control of Thought model (Anderson, 1976) emphasize the selective nature of encoding and retrieval, they do not require that experiences be integrated into a coherent and acceptable personal history for memory to be successful.

Recall of autobiographical memory is a prototypical example of a reconstructive memory process. Depending on situational requirements and personal aims, recollection of a personal experience serves specific functions. The main criterion for successful autobiographical memory recall is not whether a situation is exactly reproduced; instead, as it is one of the main functions of autobiographical memories to serve the construction of a coherent and acceptable self-image, these memories have to fit into a coherent construction of the past – for this reason, the subjects in Bartlett's studies modified the story according to their cultural expectations. Several more recent studies provide experimental evidence for this view. First, it has been shown in cross-cultural studies that autobiographical memory is influenced by the emphasis which is put in each culture on the self and on a unique life story (Nelson, 2003). Second, in each culture, autobiographical memory recall depends on an individual's current view of herself/himself, and serves the construction of a coherent and positive self (Wilson and Ross, 2003). For example, as improving selves are particularly gratifying (e.g., Frijda, 1988), subjects tend to view their own past abilities as inferior than their current abilities, and as lower than they viewed them before (Conway and Ross, 1984). Third, it was shown that the reported number of autobiographical memory recalls which aimed at creating self-continuity was higher in subjects with low degrees of self-concept clarity (corresponding to low levels of perceived self coherence), suggesting that autobiographical memory recall indeed served to strengthen self coherence (Bluck and Alea, 2008). Fourth, brain lesions which induce a loss of autobiographical memory may also lead to an impairment of the sense of one's self (Schacter, 1996).

The functions of autobiographical memory retrieval have been described in a review by Bluck (2003). First, memorizing serves the creation of a continuous identity: "[T]hough we often think of memory as a series of events, it is also a record of a series of selves, or a record of the self across time, an autobiography." (Bluck, 2003, p. 12). Second, memory is relevant for the creation and maintenance of social interactions: Often, personal statements are being justified by referring to autobiographical events, and a common past is created by evocation of situations that were experienced together. Finally, autobiographical memories serve a directive function, as they allow one to predict what will happen in the future and which actions will likely lead to the desired outcomes.

Cognitive neuroscience studies on the neuronal correlates of autobiographical memory recall support the view that this type of memory involves the construction of a self-image. These studies revealed an increased activation of both the hippocampus and the medial prefrontal cortex (mPFC) during processing of self-related autobiographical information (Cabeza et al., 2004; Summerfield et al., 2009). Activation of the mPFC was associated with self-referential processes, for example during presentation of photographs that subjects had taken themselves as compared to photographs taken by other subjects (Cabeza et al., 2004) or during evaluation whether adjectives could be attributed to oneself instead of to another person (Gutchess et al., 2007). The theory that autobiographical memory depends on an integration of declarative

memory and self-referential processes is depicted in **Figure 1B**. It should be noted, however, that reverse inference from activation of a given region during one experimental condition to mental processes occurring in this condition may be problematic because these activations are usually not unequivocal (for reviews on the validity of this inference see Henson, 2005; Axmacher et al., 2009).

While constructive concepts of memory exerted a strong influence on research in social psychology, most studies in the cognitive neurosciences (apart from autobiographical memory studies) followed the storage model. This is particularly problematic in the case of memory distortions due to repression and dissociation: Memory, conceptualized as a constructive process, serves to build personal identities. In contrast, repressed conflicts and traumatic events cannot become part of this personal identity; their subjective meaning is not re-evaluated and integrated into a narrative continuity with other events, but remains restricted to the situation when these events were experienced. Repressed conflicts and traumatic events remain permanently present in a pathological sense – they cannot be forgotten or temporarily dismissed. Thus, the case of memory distortions shows that constructive memory and storage memory are not only alternative concepts of memory, but are actually directly opposing each other in some respects. In the following two paragraphs, we describe further evidence for this view and present promising paradigms for the investigation of trauma-related and repression-related memory distortions.

## TRAUMA-RELATED MEMORY DISTORTIONS DUE TO LACK OF SELF-REFERENTIAL PROCESSING

Several observations suggest that what is central for trauma-related memory distortions is not the lack of a memory for a traumatizing situation *per se*, but the lack of an integration of this memory with self-referential processes. First, during traumatic events, subjects may lose their sense of agency, i.e., they do not feel themselves as autonomous human beings but just as observers. Dissociation during the traumatic event is predictive for subsequent memory impairments and also for the development of PTSD – even more than trauma severity is (Maercker et al., 2000; Ozer et al., 2003). Second, PTSD patients often describe that the traumatic scenes are remembered from a detached view outside of themselves (Brewin and Saunders, 2001; Bohleber, 2008). Thus, they also recall themselves as lacking self-referential processing during the trauma. Third, loss of agency during traumatic events may also explain why traumatic events cannot only induce PTSD, but also a complex PTSD (Herman, 1992) or developmental trauma disorder (van der Kolk, 2005) which both involve changes to the self-image and an impaired feeling of identity. Finally, flashbacks can be understood as a dissociative re-living of the traumatic situation, i.e., re-living in a state of depersonalization and derealization (Maercker, 2009). Taken together, these considerations suggest that traumatic events mainly impair the *integration* of experiences with the internal perception and valuation of these events, and that the resulting memory distortions need to be conceptualized within a constructive memory framework (**Figure 1C**).

How can the reduced self-referential processing during traumatic experiences, with the resulting impairment in the construction of memories for these events, be assessed experimentally? The effect of dissociation on the development of trauma-related memory distortions was studied in healthy control subjects by use of the *trauma film paradigm* (for an overview of this method, see Holmes and Bourne, 2008). This paradigm builds on observations that intrusive thoughts do not only occur in PTSD patients, but also in many situations of everyday life after strong emotional events (Berntsen, 1996; Mace, 2005). In the trauma film paradigm, intrusions are induced by a movie which contains emotionally disturbing scenes (e.g., scenes of victims from car accidents). This paradigm has already been used in a number of studies and reliably induces intrusive memories, which disappear after a few weeks (of course, ethical reasons prohibit the induction of an actual PTSD which is defined by the persistence of symptoms, after a trauma, for more than a month). In addition to intrusions, declarative memory for the contents of the trauma film is impaired (Brewin, 2007; Jones et al., 2007). Using this paradigm, it was shown that spontaneously occurring states of dissociation predicted subsequent intrusions (Holmes et al., 2004; Kindt et al., 2005). However, any attempts to increase intrusion incidence by experimentally induced dissociation failed (e.g., Holmes et al., 2007), suggesting that dissociation may affect memory distortions via an underlying psychological or physiological process, which is not triggered by the experimentally induced dissociation – e.g., via reduced activation of brain regions supporting self-referential processing.

The neural basis of memory effects in the trauma film paradigm was investigated in two recent functional MRI studies (Henckens et al., 2009; van Marle et al., 2009). Physiological parameters as well as subjective reports confirmed that stress was induced by segments of a distressing movie. In the first study, IAPS pictures were presented interleaved with these segments and brain activity related to declarative encoding of IAPS pictures was analyzed (Henckens et al., 2009). In contrast to the impairment of declarative memory for traumatic events observed clinically, the authors found that stress increased subsequent recollection of pictures. However, hippocampal activity was reduced during successful encoding of images under stress, suggesting that memory formation required a hippocampal-independent mechanism. These results are therefore consistent with the dual representation theory and suggest that a high stress level shifts encoding from a hippocampus-dependent to an amygdala-dependent encoding mechanism (Brewin, 2003), although no implicit memory tests were performed, and thus the exact memory processes contributing to the stress-induced memory enhancement could not be resolved. In line with this interpretation, the second study showed increased responsiveness of the amygdala to facial stimuli presented interleaved with the movie, although memory for these items was not tested (van Marle et al., 2009). Further research using the trauma film paradigm will be extremely useful to test predictions from the dual representation theory more directly: First, it will be interesting to test memory for the movie segments themselves (instead of interleaved stimuli). Second, not only declarative memory but also intrusions should be captured using the diary method (Brewin and Saunders, 2001; Bisby et al., 2009); i.e., intrusions during the weeks following the experiment should be collected by diaries given to the participants. Finally – and maybe most importantly – the effects of dissociation during the movie should be investigated. As described above, it appears to be extremely difficult to induce dissociation by experimental

modifications. Therefore, dissociation could be measured on the neural level, by a reduced activation of regions supporting self-referential processing.

## PROMISING APPROACHES FOR STUDYING MEMORY DISTORTIONS DUE TO REPRESSION

The neural signature of self-referential processing during memory recall can be used as a measure of successful reconstructive recall – or the lack thereof – not only in the trauma film paradigm, but also in paradigms investigating memory distortions due to repression. In one important paradigm (see **Figure 3**), which was initially suggested by Jung (1918), words in a list are presented consecutively and subjects are instructed to generate an associated word to each word in the list using the psychoanalytical technique of "free association" – i.e., they are asked to say the first word which comes into their mind. The idea of "free association" is rooted in Freud's psychoanalytical technique: When patients with neurotic symptoms say spontaneously what comes to their minds, they may eventually reveal material that can be linked by a skilled therapist to repressed conflicts which lead to or maintain their symptoms (Freud, 1913). The clinical usefulness of this technique has been shown many times (e.g., Person et al., 2005). According to Freud (1913), patients typically start to repress spontaneous thoughts once they may link to conflicts or memories that have to be kept unconscious. In Jung's (1918) original experiment, the process of "free association" is believed to reveal contents that are primarily unconscious but relevant for the subjects because they may be related to repressed conflicts. The behavioral result of the hypothesized repression – resistance against revelation of these contents – can be confirmed experimentally by an increase in skin conductance response (SCR) and reaction times. The link between repression and the hypothesized increase in skin conductance is an indirect one: It is known that psychological arousal (activation) leads to increased skin conductance (Lang et al., 1995). Clinical experience suggests that patients may show increased signals of stress and arousal when repressing critical contents (Person et al., 2005). Thus, it is assumed that repression may be operationalized by increases in skin conductance. Next, the same list is presented again, but now subjects are not asked to name a new word by free association, but to recall the word they have initially generated. Finally, all initially associated words are presented again, and subjects are instructed to indicate the emotional valence and intensity of these words (as an indicator of conscious emotional content). In contrast to the often described increase in declarative memory for negative emotional material (Heuer and Reisberg, 1990; Bradley et al., 1992; Christianson, 1992; Ochsner, 2000; Buchanan and Lovallo, 2001; Kensinger and Corkin, 2003), this paradigm reliably results in an *impaired* memory for words whose initial generation is associated with physiological signs of resistance (increase in skin conductance and reaction times) and which are subsequently rated as emotionally negative (Levinger and Clark, 1961; Köhler and Wilke, 1999; Köhler et al., 2002). In particular, Levinger and Clark (1961) found that high SCRs during association predicted failure of subsequent recall. Furthermore, recall was worse for words which were afterward labeled as emotional as compared to words

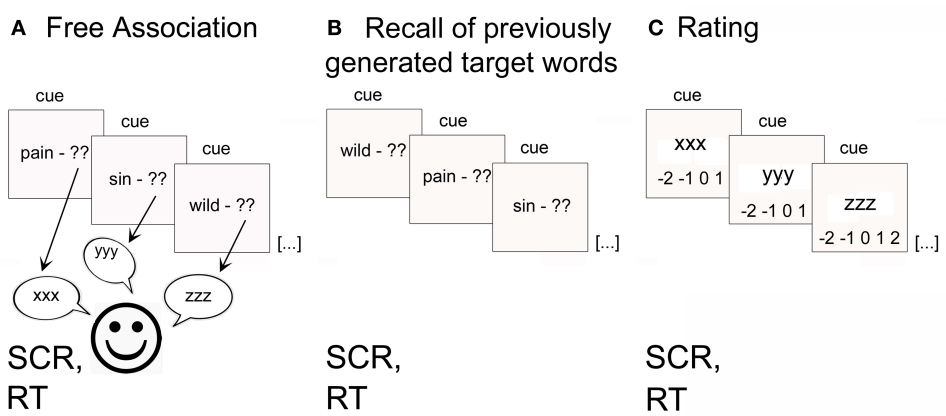



**FIGURE 3 | Repression of self-generated emotional words. (A)** Presentation of cue words, participants were instructed to speak out loudly the first word which came to their minds while skin conductance response (SCR) and reaction latencies were recorded. **(B)** At memory test, the same cue words were presented again, and participants had to recall the word which they previously associated with the cue word during phase **(A)** of the experiment. **(C)** Emotional rating of the self-generated words.

**Table 1 | Experimental paradigms for studying memory distortions.**

| Paradigm | Automatic stimulus effects? | Implicit memory enhanced? | Declarative memory impaired? | Self-referential processing considered? |
|---|---|---|---|---|
| Directed forgetting; Think/No-Think | No | ? (not tested) | Yes | No |
| Trauma film paradigm | Yes | Yes (intrusions) | Yes | Yes (dissociation) |
| Levinger/Clark/Köhler (**Figure 3**) | Yes | Yes (reaction times) | Yes | Yes (subject-specific cues) |

which were labeled as neutral. Several alternative explanations might account for these results. First, word frequency of the presented words may determine recall of associations. However, frequency was equilibrated between emotional and neutral words. Second, it is possible that recall depended on the number of possible associations to a word ("response entropy"). For words associated with high response entropy, the possible associations interfere with each other, making it more difficult to recall the initially associated word. In fact, Levinger and Clark found that forgotten words had higher response entropies than remembered words. However, partial correlations revealed that emotion and response entropy contributed independently to memory.

Several shortcomings of the study by Levinger and Clark (1961) should be noted. First, recall was only tested immediately after initial associations, leaving open the possibility that it does not induce sustained effects on memory. Second, reaction times as another possible measure of resistance toward revelation of repressed conflicts were not tested. In a follow-up study designed to overcome these shortcomings, Köhler and Wilke (1999) conduced a similar experiment, but (1) also measured reaction times during initial association, (2) asked subjects to recall the associated words not only directly after the association, but again after one week. They found that increased SCRs and reaction times during association as well as emotional ratings predicted both immediate forgetting and forgetting after 1 week.

To our knowledge, this paradigm has never been used for cognitive neuroscience experiments. However, the dual representation theory (Brewin, 2003) predicts that generation of subsequently forgotten words is associated with increased amygdala and decreased hippocampal activation, and a negative correlation of activity in these regions. Furthermore, based on the idea that repression is related to reduction in self-referential processing, we would expect that generation of these words leads to decreased activation of the mPFC, and to a reduced functional connectivity between this region and the hippocampus.

## CONCLUSIONS

To summarize, we argued that storage models of memory as usually employed in the cognitive neurosciences are unable to capture the constructive nature of autobiographical memories and the impairment of this construction due to repression and traumatic experiences, and presented a number of experimental paradigms to study memory distortions in healthy subjects (**Table 1**). It should be noted that other lines of research have started to question the storage model of memory in cognitive neuroscience as well. First, the concept of *reconsolidation* implies that memories are not encoded for permanent storage, but that each recall involves a weakening of the memory trace and its possible re-integration with new experiences (Nader et al., 2000; Nader and Hardt, 2009). Second, studies of autobiographical memory often focused on the reconstruction of experiences rather than their exact recall. Third, the "prospective memory theory" describes that similar brain regions are activated during remembering the past and imagining the future, suggesting that both processes rely on similar constructions, or simulations, of currently absent stimulus representations (Schacter and Addis, 2007; Schacter et al., 2007). Fourth, social psychological studies exploring effects of social conformity on memory retrieval (Cialdini and Goldstein, 2004; Wright et al., 2009; Axmacher et al., 2010) could, in principle, also be adapted for neuroimaging research (Berns et al., 2005; Klucharev et al., 2009). Finally, the emerging field of "neuro-psychoanalysis" attempts to bridge the traditional gap between cognitive psychology and neuroscience on the one hand, psychoanalysis on the other hand, and might contribute relevant conceptual enhancements to cognitive neuroscience (Solms and Kaplan-Solms, 2000).

## ACKNOWLEDGMENT

Anne T. A. Do Lam was supported by a grant from DFG FE 366/5-1.

## REFERENCES

American Psychiatric Association (APA). (2000). *Diagnostic and Statistical Manual of Mental Disorders – DSM-IV-TR*, 4th Edn, Text Revision. Washington, DC: American Psychiatric Association.

Anderson, J. (1976). *Language, Memory and Thought*. Hillsdale, NJ: Erlbaum Associates.

Anderson, M. C., Bjork, R. A., and Bjork, E. L. (1994). Remembering can cause forgetting: retrieval dynamics in long-term memory. *J. Exp. Psychol. Learn. Mem. Cogn.* 20, 1063–1087.

Anderson, M. C., and Green, C. (2001). Suppressing unwanted memories by executive control. *Nature* 410, 366–369.

Anderson, M. C., Ochsner, K. N., Kuhl, B., Cooper, J., Robertson, E., Gabrieli, S. W., Glover, G. H., and Gabrieli, J. D. E. (2004). Neural systems underlying the suppression of unwanted memories. *Science* 303, 232–235.

Assmann, A. (2002). *Erinnerungsräume. Formen und Wandel des kulturellen Gedächtnisses*. München, Germany: C.H. Beck.

Axmacher, N., Elger, C. E., and Fell, J. (2009). The specific contribution of neuroimaging versus neurophysiological data to understanding cognition. *Behav. Brain Res.* 200, 1–6.

Axmacher, N., Gossen, A., Elger, C. E., and Fell, J. (2010). Graded effects of social conformity on recognition memory. *PLoS ONE* 5, e9270. doi: 10.1371/journal.pone.0009270.

Baddeley, A. D. (2002). *Human Memory: Theory and Practice*. Hove, UK: Psychology Press.

Balint, M. (1969). Trauma and object relationship. *Int. J. Psychoanal.* 50, 429–435.

Bartlett, F. C. (1932). *Remembering: A Study in Experimental and Social Psychology*. Cambridge, UK: Cambridge University Press.

Bechara, A., Tranel, D., Damasio, H., Adolphs, R., Rockland, C., and Damasio, A. R. (1995). Double dissociation of conditioning and declarative knowledge relative to the amygdala and hippocampus in humans. *Science* 269, 1115–1118.

Behrendt, A., and Moritz, S. (2005). Posttraumatic stress disorder and memory problems after female genital mutilation. *Am. J. Psychiatry* 162, 1000–1001.

Berns, G. S., Chappelow, J., Zink, C. F., Pagnoni, G., Martin-Skurski, M. E., and Richards, J. (2005). Neurobiological correlates of social conformity and independence during mental rotation. *Biol. Psychiatry* 58, 245–253.

Berntsen, D. (1996). Involuntary autobiographical memory. *Appl. Cogn. Psychol.* 10, 435–454.

Bisby, J. A., Brewin, C. R., Leitz, J. R., and Curran, H. V. (2009). Acute effects of alcohol on the development of intrusive memories. *Psychopharmacology* 204, 655–666.

Bluck, S. (2003). Autobiographical memory: exploring its functions in everyday life. *Memory* 11, 113–123.

Bluck, S., and Alea, N. (2008). "Remembering being me: the self continuity function of autobiographical memory in younger and older adults," in *Self Continuity: Individual and Collective Perspectives*, ed F. Sani (New York, NY: Psychology Press), 55–70.

Bohleber, W. (2008). "Einige Probleme psychoanalytischer Traumatheorie," in *Psychoanalyse, Neurobiologie, Trauma*, eds M. Leuzinger-Bohleber, G. Roth, and A. Buchheim (Stuttgart, Germany: Schattauer), 45–54.

Bradley, M. M., Greenwald, M. K., Petry, M. C., and Lang, P. J. (1992). Remembering pictures: pleasure and arousal in memory. *J. Exp. Psychol. Learn. Mem. Cogn.* 18, 379–390.

Bransford, J. D., Franks, J. J., Morris, C. D., and Stein, B. S. (1979). "Some general constraints on learning and memory research," in *Levels of Processing in Human Memory*, eds L. S. Cermak and F. I. M. Craik (Hillsdale, NJ: Erlbaum), 331–354.

Bremner, J. D. (2002). *Does Stress Damage the Brain? Understanding Trauma-Related Disorders from a Neurological Perspective*. New York: Norton.

Bremner, J. D., and Brett, E. (1997). Trauma-related dissociative states and long-term psychopathology in post-traumatic stress disorder. *J. Trauma. Stress* 10, 37–49.

Bremner, J. D., Randall, P., Scott, T. M., Capelli, S., Delaney, R., McCarthy, G., and Charney, D. S. (1995). Deficits in short-term memory in adult survivors of childhood abuse. *Psychiatry Res.* 59, 97–107.

Bremner, J. D., Scott, T. M., Delaney, R. C., Southwick, S. M., Mason, J. W., Johnson, D. R., Innis, R. B., McCarthy, G., and Charney, D. S. (1993). Deficits in short-term memory in posttraumatic stress disorder. *Am. J. Psychiatry* 150, 1015–1019.

Bremner, J. D., Vermetten, E., Afzal, N., and Vythilingam, M. (2004). Deficits in verbal declarative memory function in women with childhood sexual abuse-related posttraumatic stress disorder. *J. Nerv. Ment. Dis.* 192, 643–649.

Breuer, J., and Freud, S. (1895). "Studies on hysteria," in *The Standard Edition of the Complete Works of Sigmund Freud*, Vol. 2, eds J. Strachey, A. Freud, A. Strachey, and A. Tyson (London: The Hogarth Press and the Institute of Psychoanalysis), 125–134.

Brewin, C. R. (2001). A cognitive neuroscience account of posttraumatic stress disorder and its treatment. *Behav. Res. Ther.* 39, 373–393.

Brewin, C. R. (2003). *Posttraumatic Stress Disorder – Malady or Myth?* New Haven: Yale University Press.

Brewin, C. R. (2007). Autobiographical memory for trauma: update on four controversies. *Memory* 15, 227–248.

Brewin, C. R., and Holmes, E. A. (2003). Psychological theories of posttraumatic stress disorder. *Clin. Psychol. Rev.* 23, 339–376.

Brewin, C. R., and Saunders, J. (2001). The effect of dissociation at encoding on intrusive memories for a stressful film. *Br. J. Med. Psychol.* 74, 467–472.

Buchanan, T. W., and Lovallo, W. R. (2001). Enhanced memory for emotional material following stress-level cortisol treatment in humans. *Psychoneuroendocrinology* 26, 307–317.

Cabeza, R., Prince, S. E., Daselaar, S. M., Greenberg, D. L., Budde, M., Dolcos, F., LaBar, K. S., and Rubin, D. C. (2004). Brain activity during episodic retrieval of autobiographical and laboratory events: an fMRI study using a novel photo paradigm. *J. Cogn. Neurosci.* 16, 1583–1594.

Caroll, M., Campbell-Ratcliffe, J., Murnane, H., and Perfect, T. J. (2007). Retrieval-induced forgetting in educational contexts: monitoring, expertise, text integration and test format. *Eur. J. Cogn. Psychol.* 19, 580–606.

Christianson, S. A. (1992). Emotional stress and eyewitness memory: a critical review. *Psychol. Bull.* 112, 284–309.

Cialdini, R. B., and Goldstein, N. J. (2004). Social influence: compliance and conformity. *Annu. Rev. Psychol.* 55, 591–621.

Clancy, S. A., Schacter, D. L., McNally, R. J., and Pitman, R. K. (2000). False recognition in women reporting recovered memories of sexual abuse. *Psychol. Sci.* 11, 26–31.

Conway, M., and Ross, M. (1984). Getting what you want by revising what you had. *J. Pers. Soc. Psychol.* 47, 738–748.

Craik, F. I. M., and Lockhardt, R. S. (1972). Levels of processing: a framework for memory research. *J. Verbal Learn. Verbal Behav.* 11, 671–684.

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *J. Exp. Psychol.* 58, 17–22.

Degonda, N., Mondadori, C. R., Bosshardt, S., Schmidt, C. F., Boesiger, P., Nitsch, R. M., Hock, C., and Henke, K. (2005). Implicit associative learning engages the hippocampus and interacts with explicit associative learning. *Neuron* 46, 505–520.

Depue, B. E., Curran, T., and Banich, M. T. (2007). Prefrontal regions orchestrate suppression of emotional memories via a two-phase process. *Science* 317, 215–219.

Dolcos, F., LaBar, K. S., and Cabeza, R. (2004). Interaction between the amygdala and the medial temporal lobe memory system predicts better memory for emotional events. *Neuron* 42, 855–863.

Elbert, T., and Schauer, M. (2002). Burnt into memory. *Nature* 419, 883.

Erdelyi, M. H. (2006). The unified theory of repression. *Behav. Brain Sci.* 29, 499–551.

Ferenczi, S. (1933). *Bausteine der Psychoanalyse*. Bern, Switzerland: Huber, 511–525.

Fonagy, P., Gergely, G., Jurist, E. J., and Target, M. (2005). *Affect regulation, Mentalization, and the Development of the Self*. (New York, NY: Other Press).

Freud, S. (1913). "On beginning the treatment (further recommendations on the technique of psycho-analysis I)," in *The Standard Edition of the Complete Works of Sigmund Freud*, Vol. 12, eds J. Strachey, A. Freud, A. Strachey, and A. Tyson (London: The Hogarth Press and the Institute of Psychoanalysis), 121–145.

Freud, S. (1915). "Repression," in *The Standard Edition of the Complete Works of Sigmund Freud*, eds J. Strachey, A. Freud, A. Strachey, and A. Tyson (London: The Hogarth Press and the Institute of Psychoanalysis), Vol. 14, pp. 143–158.

Frijda, N. H. (1988). The laws of emotion. *Am. Psychol.* 43, 349–358.

Gutchess, A. H., Kensinger, E. A., and Schacter, D. L. (2007). Aging, self-referencing, and medial prefrontal cortex. *Soc. Neurosci.* 2, 117–133.

Henckens, M. J., Hermans, E. J., Pu, Z., Joëls, M., and Fernández, G. (2009). Stressed memories: how acute stress affects memory formation in humans. *J. Neurosci.* 29, 10111–10119.

Henke, K., Mondadori, C. R., Treyer, V., Nitsch, R. M., Buck, A., and Hock, C. (2003). Nonconscious formation and reactivation of semantic associations by way of the medial temporal lobe. *Neuropsychologia* 41, 863–876.

Henson, R. (2005). What can functional neuroimaging tell the experimental psychologist? *Q. J. Exp. Psychol.* 58, 193–233.

Herbart, J. F. (1824). *Psychologie als Wissenschaft, neu gegründet auf Erfahrung, Metaphysik und Mathematik*. Königsberg: A.W. Unzer.

Herman, J. (1992). Complex PTB: a syndrome in survivors of prolonged and repeated trauma. *J. Trauma. Stress* 5, 377–391.

Heuer, F., and Reisberg, D. (1990). Vivid memories of emotional events: the accuracy of remembered minutiae. *Mem. Cogn.* 18, 496–506.

Holmes, E. A., and Bourne, C. (2008). Inducing and modulating intrusive emotional memories: a review of the trauma film paradigm. *Acta Psychol.* 127, 553–566.

Holmes, E. A., Brewin, C. R., and Hennessy, R. G. (2004). Trauma films, information processing, and intrusive memory development. *J. Exp. Psychol. Gen.* 133, 3–22.

Holmes, E. A., Oakley, D. A., Stuart, A., and Brewin, C.R. (2007). Investigating peri-traumatic dissociation using hypnosis during a traumatic film. *J. Trauma Dissociation* 7, 91–111.

Hunt, R. R., and Einstein, G. O. (1981). Relational and item-specific information in memory. *J. Verbal Learn. Verbal Behav.* 20, 497–514.

Jelinek, L., Jacobson, D., Kellner, M., Larbig, F., Biesold, K.-H., Barre, K., and Moritz, S. (2006). Verbal and nonverbal memory functioning in posttraumatic stress disorder (PTSD). *J. Clin. Exp. Neuropsychol.* 28, 940–948.

Johnson, H. M. (1994). Processes of successful intentional forgetting. *Psychol. Bull.* 116, 274–292.

Johnson, M. K., Hashtroudi, S., and Lindsay, D. S. (1993). Source monitoring. *Psychol. Bull.* 114, 3–28.

Jones, C., Harvey, A. G., and Brewin, C. R. (2007). The organisation and content of trauma memories in survivors of road traffic accidents. *Behav. Res. Ther.* 45, 151–162.

Jung, C. G. (1918). *Studies in Word-Association*. London: Heinemann.

Kensinger, E. A., and Corkin, S. (2003). Effect of negative emotional content on working memory and long-term memory. *Emotion* 3, 378–393.

Kensinger, E. A., and Corkin, S. (2004). Two routes to emotional memory: distinct neural processes for valence and arousal. *Proc. Natl. Acad. Sci. U.S.A.* 101, 3310–3315.

Kim, J. J., and Diamond, D. M. (2002). The stressed hippocampus, synaptic plasticity and lost memories. *Nat. Rev. Neurosci.* 3, 453–462.

Kindt, M., Van den Hout, M., and Buck, N. (2005). Dissociation related to subjective memory fragmentation and intrusions but not to objective memory disturbances. *J. Behav. Ther. Exp. Psychiatry* 36, 43–59.

Klucharev, V., Hytönen, K., Rijpkema, M., Smidts, A., and Fernández, G. (2009). Reinforcement learning signal predicts social conformity. *Neuron* 61, 140–151.

Köhler, T., Thiede, G., and Thöns, M. (2002). Kurz- und längerfristiges Vergessen von Wortassoziationen – eine experimentelle Studie zur Freudschen Lehre von Widerstand und Verdrängung. *Z. Klin. Psychol. Psychiatr. Psychother.* 50, 328–333.

Köhler, T., and Wilke, W. (1999). Das Vergessen von Wortassoziationen in Abhängigkeit von Indikatoren ihrer Emotionalität: Eine Möglichkeit der Überprüfung des Freudschen Verdrängungskonzepts? *Psychother. Psychosom. Med. Psychol.* 49, 64–67.

LaBar, K. S., LeDoux, J. E., Spencer, D. D., and Phelps, E. A. (1995). Impaired fear conditioning following unilateral temporal lobectomy in humans. *J. Neurosci.* 15, 6846–6855.

Lang, P. J., Bradley, M. M., and Cuthbert, B. N. (1995). *International Affective Picture System (IAPS): Technical Manual and Affective Ratings*. Gainesville, FL: The Center for Research in Psychophysiology, University of Florida.

Lanius, R. A., Williamson, P. C., Boksman, K., Densmore, M., Gupta, M., Neufeld, R. W. J., Gati, J. S., and Menon, R. S. (2002). Brain activation during script-driven imagery induced dissociative responses in PTSD: a functional magnetic resonance imaging investigation. *Biol. Psychiatry* 52, 305–311.

Leichsenring, F., and Rabung, S. (2008). Effectiveness of long-term

psychodynamic psychotherapy: a meta-analysis. *JAMA* 300, 1551–1565.

Levinger, G., and Clark, J. (1961). Emotional factor in the forgetting of word associations. *J. Abnorm. Soc. Psychol.* 62, 99–105.

Mace, J. H. (2005). Priming involuntary autobiographical memories. *Memory* 13, 874–884.

Macrae, C. N., and MacLeod, M. D. (1999). On recollections lost: when practice makes imperfect. *J. Pers. Soc. Psychol.* 77, 463–473.

Maercker, A. (2009). *Posttraumatische Belastungsstörungen*. Heidelberg, Germany: Springer.

Maercker, A., Schützwohl, M., and Beauducel, A. (2000). Trauma severity and initial reactions as precipitating factors for posttraumatic stress disorder and chronic dissociation. *J. Trauma. Stress* 13, 651–660.

McDaniel, M.A., and Einstein, G. O. (1993). The importance of cue familiarity and cue distinctiveness in prospective memory. *Memory* 1, 23–41.

McNally, R. J. (1997). Implicit and explicit memory for trauma-related information in PTSD. *Ann. N. Y. Acad. Sci.* 821, 219–224.

Metcalfe, J., and Jacobs, W. J. (1998). "Emotional memory: the effects of stress on "cool" and "hot" memory systems," in *The Psychology of Learning and Motivation: Advances in Research and Theory*, ed L. D. Medin (San Diego: Academic Press), 187–222.

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.

Nader, K., and Hardt, O. (2009). A single standard for memory: the case for reconsolidation. *Nat. Rev. Neurosci.* 10, 224–234.

Nader, K., Schafe, G. E., and Le Doux, J. E. (2000). Fear memories require protein synthesis in the amygdala for reconsolidation after retrieval. *Nature* 406, 722–726.

Nelson, K. (2003). Self and social functions: individual autobiographical memory and collective narrative. *Memory* 11, 125–136.

Neuner, F., Schauer, M., Klaschik, C., Karunakara, U., and Elbert, T. (2004). A comparison of narrative exposure therapy, supportive counseling, and psychoeducation for treating posttraumatic stress disorder in an African refugee settlement. *J. Consult. Clin. Psychol.* 72, 579–587.

Ochsner, K. N. (2000). Are affective events richly recollected or simply familiar? The experience and process of recognizing feelings past. *J. Exp. Psychol. Gen.* 129, 242–261.

Ozer, E. J., Best, S. R., Lipsey, T. L., and Weiss, D. S. (2003). Predictors of posttraumatic stress disorder and symptoms in adults: a meta-analysis. *Psychol. Bull.* 129, 52–73.

Person, E. S., Cooper, A. M., and Gabbard, G. O. (eds). (2005). *The American Psychiatric Publishing Textbook of Psychoanalysis*. Washington, DC: American Psychiatric Publishing, 602.

Phelps, E. A. (2004). Human emotion and memory: interactions of the amygdala and hippocampal complex. *Curr. Opin. Neurobiol.* 14, 198–202.

Rockstroh, B., and Elbert, T. (2010). Traces of fear in the neural web – magnetoencephalographic responding to arousing pictorial stimuli. *Int. J. Psychophysiol.* 78, 14–19.

Roediger, H. L. III, and McDermott, K. B. (1995). Creating false memories: remembering words not presented in lists. *J. Exp. Psychol. Learn. Mem. Cogn.* 21, 803–814.

Sapolsky, R. M. (2003). Stress and plasticity in the limbic system. *Neurochem. Res.* 28, 1735–1742.

Schacter, D. L. (1996). *Searching for Memory: The Brain, the Mind, and the Past*. New York: Basic Books Inc.

Schacter, D. L., and Addis, D. R. (2007). Constructive memory: the ghosts of past and future. *Nature* 445, 27.

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

Seidenbecher, T., Laxmi, T. R., Stork, O., and Pape, H. C. (2003). Amygdalar and hippocampal theta rhythm synchronization during fear memory retrieval. *Science* 301, 846–850.

Shin, L. M., Rauch, S. L., and Pitman, R. K. (2005). "Structural and functional anatomy of PTSD: findings from neuroimaging research," in *Neuropsychology of PTSD: Biological, Cognitive, and Clinical Perspectives*, eds J. J. Vasterling and C. R. Brewin (New York: The Guilford Press), 59–82.

Smith, A. P., Stephan, K. E., Rugg, M. D., and Dolan, R. J. (2006). Task and content modulate amygdala-hippocampal connectivity in emotional retrieval. *Neuron* 49, 631–638.

Solms, M., and Kaplan-Solms, K. (2000). *Clinical Studies in Neuro-Psychoanalysis: Introduction to a Depth Neuropsychology*. London, UK: Karnac Books.

Stephan, A. (2002). "Was die Psychoanalyse im Lichte interdisziplinärer Forschung erwartet," in *Psychoanalse im Dialog der Wissenschaften. Band 1: Europäische Perspektiven*, ed P. Giamperi-Deutsch (Stuttgart, Germany: Kohlhammer), 76–87.

Stern, D. N. (1985). *The Interpersonal World of the Infant: A View from Psychoanalysis and Developmental Psychology*. New York, NY: Basic Books.

Summerfield, J. J., Hassabis, D., and Maguire, E. A. (2009). Cortical midline involvement in autobiographical memory. *Neuroimage* 44, 1188–1200.

Tulving, E. (1983). *Elements of Episodic Memory*. Oxford, UK: Clarendon.

van der Kolk, B. (2005). Developmental trauma disorder. Towards a rational diagnosis for children with complex trauma histories. *Psychiatr. Ann.* 35, 979–987.

van Marle, H. J., Hermans, E. J., Qin, S., and Fernández, G. (2009). From specificity to sensitivity: how acute stress affects amygdala processing of biologically salient stimuli. *Biol. Psychiatry* 66, 649–655.

Wilson, A. E., and Ross, M. (2003). The identity function of autobiographical memory: time is on our side. *Memory* 11, 137–149.

Wöller, W., and Kruse, J. (2010). *Tiefenpsychologisch fundierte Psychotherapie, Basisbuch und Praxisleitfaden*. Stuttgart, Germany: Schattauer.

World Health Organization (WHO). (1992). *The ICD-10 Classification of Mental and Behavioral Disorders: Clinical Descriptions and Diagnostic Guidelines*. Geneva: World Health Organization.

Wright, D. B., Memon, A., Skagerberg, E. M., and Gabbert, F. (2009). When eyewitnesses talk. *Curr. Dir. Psychol. Sci.* 18, 174–178.

Yehuda, R., Keefe, R. S., Harvey, P. D., Levengood, R. A., Gerber, D. K., Geni, J., and Siever, L. J. (1995). Learning and memory in combat veterans with posttraumatic stress disorder. *Am. J. Psychiatry* 152, 137–139.

**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 2010; accepted: 11 October 2010; published online: 22 November 2010.*

*Citation: Axmacher N, Do Lam ATA, Kessler H and Fell J (2010) Natural memory beyond the storage model: repression, trauma, and the construction of a personal past. Front. Hum. Neurosci.* ***4****:211. doi: 10.3389/fnhum.2010.00211*

frontiers in
HUMAN NEUROSCIENCE

REVIEW ARTICLE
published: 13 December 2010
doi: 10.3389/fnhum.2010.00222

# To head or to heed? Beyond the surface of selective action inhibition: a review

**Wery P.M. van den Wildenberg[1]*, Scott A. Wylie[2], Birte U. Forstmann[3], Borís Burle[4], Thierry Hasbroucq[4] and K. Richard Ridderinkhof[1,3]**

[1] Department of Psychology, Amsterdam Center for the Study of Adaptive Control in Brain and Behavior, University of Amsterdam, Amsterdam, Netherlands
[2] Neurology Department, University of Virginia Health Systems, Charlottesville, VA, USA
[3] Spinoza Center for Neuroimaging, University of Amsterdam, Amsterdam, Netherlands
[4] Laboratoire de Neurobiologie de la Cognition, Centre National de la Recherche Scientifique, Aix-Marseille Université, Marseille, France

**Edited by:**
Michael X. Cohen, University of Amsterdam, Netherlands

**Reviewed by:**
Tobias Egner, Duke University, USA
Birgit Stürmer, Humboldt Universität Berlin, Germany

***Correspondence:**
Wery P.M. van den Wildenberg, Department of Psychology, University of Amsterdam, Roetersstraat 15, 1018 WB Amsterdam, Netherlands.
e-mail: w.p.m.vandenwildenberg@uva.nl

To head rather than heed to temptations is easier said than done. Since tempting actions are often contextually inappropriate, selective suppression is invoked to inhibit such actions. Thus far, laboratory tasks have not been very successful in highlighting these processes. We suggest that this is for three reasons. First, it is important to dissociate between an early susceptibility to making stimulus-driven impulsive but erroneous actions, and the subsequent selective suppression of these impulses that facilitates the selection of the correct action. Second, studies have focused on mean or median reaction times (RT), which conceals the temporal dynamics of action control. Third, studies have focused on group means, while considering individual differences as a source of error variance. Here, we present an overview of recent behavioral and imaging studies that overcame these limitations by analyzing RT distributions. As will become clear, this approach has revealed variations in inhibitory control over impulsive actions as a function of task instructions, conflict probability, and between-trial adjustments (following conflict or following an error trial) that are hidden if mean RTs are analyzed. Next, we discuss a selection of behavioral as well as imaging studies to illustrate that individual differences are meaningful and help understand selective suppression during action selection within samples of young and healthy individuals, but also within clinical samples of patients diagnosed with attention deficit/hyperactivity disorder or Parkinson's disease.

**Keywords: action control, response inhibition, prefrontal cortex, basal ganglia, interference control**

## IMPULSIVE ACTIVATION AND SELECTIVE SUPPRESSION OF ACTIONS DURING CONFLICT

"Look before you leap" and "haste makes waste" are just two examples of every-day expressions – not to say clichés – that point to the precarious balance between reacting impulsively on the one hand versus carefully weighing your response options before acting. The need for behavioral restraint during decision making is particularly relevant when confronted with irrelevant attributes or changes in the environment that activate unintentional response tendencies that conflict with the desired appropriate behavior. This review first introduces a theoretical framework, *the activation–suppression model* that describes the effects of processing conflicting information on behavior. Importantly, the activation–suppression model generates two specific individual parameters that represent the temporal aspects of (1) the susceptibility to make fast impulsive reactions, and (2) the proficiency of selective inhibitory control over these unwanted actions to facilitate the selection of the appropriate response. This useful dissociation is illustrated by an overview of behavioral studies that helped identify experimental variables (e.g., conflict probability and task instructions) as well as differences between groups of individuals (e.g., age, and clinical diagnoses like attention deficit/hyperactivity disorder [AD/HD] and Parkinson's disease) that affect the initial activation and selective suppression of unwanted action tendencies. Finally, we present an overview of recent brain imaging studies that focus on the underlying neural mechanisms. As will become clear, activation and effective selective suppression of unwanted actions is expressed by subthreshold muscle activity that can be measured using electromyography (EMG) recorded from the incorrect response hand. Application of brain imaging techniques advanced our understanding of the cortical activation patterns, structural connectivity, and subcortical contributions that are associated with the voluntary selection of actions in the face of conflicting response impulses.

**Abbreviations:** AD/HD, attention deficit/hyperactivity disorder; CAF, conditional accuracy function; EEG, electroencephalography; EMG, electromyography; IFC, inferior frontal cortex; pre-SMA, pre-supplementary motor area; RT, reaction time; STN, subthalamic nucleus.

### PARADIGMS THAT INDUCE CONFLICT

Over the years, several reaction time (RT) paradigms have been developed to study cognitive processing in conflicting situations in which competing response tendencies are simultaneously active. Widely used variants of such conflict paradigms are the Stroop task (Stroop, 1935) and the Eriksen flanker task (Eriksen and Eriksen, 1974), that all share a common cognitive processing architecture, *the dual-route model* (Kornblum et al., 1990; Kornblum, 1994). Although these different tasks clearly induce interference produced

by response conflict, the source that triggers interference may vary across conflict tasks. For example, the Stroop interference involves a number of potential sources of variance, such as response conflict, perceptual conflict (as in the Flanker task), and semantic interference. Alternatively, the Simon task (Simon, 1967, 1990) induces response conflict and inherently avoids interference associated with perceptual or semantic conflict. The Simon task (Simon, 1967, 1990) in particular provides a context in which irrelevant stimulus information can elicit a strong response impulse that interferes with goal-directed action (see **Figure 1**). The task typically requires a fast button-press to a goal-relevant stimulus aspect embedded in a goal-irrelevant stimulus dimension. For example, participants may be instructed to issue a discriminative response according to the color of a stimulus; to press the left response button when the circle is blue and to issue a right-hand button press to a green circle. Notably, the colored circles are presented to the left or right of a fixation point.

According to dual-route models of information processing, the spatial location of the stimulus, despite being irrelevant, automatically and rapidly activates the spatially corresponding response via a direct processing route (Kornblum et al., 1990; de Jong et al., 1994; Eimer et al., 1995; Ridderinkhof, 2002a; Ridderinkhof et al., 2004c). In contrast, the relevant stimulus feature engages a deliberate processing route that utilizes a slower controlled translation of relevant stimulus features into a correct response according to task instructions. On *compatible* trials, the direct route and the deliberate processing route converge on activation of the same response and, in so doing, facilitate both RT and accuracy. In contrast, on *incompatible* trials, the two responses, that is, the one activated by the automatic processing of the circle's spatial location and the one activated according to the deliberate processing of the circle's color, are incompatible, thereby slowing RT and increasing error rates. The mean *compatibility* or interference effect is taken to reflect the extra demands and time required to overcome the interference caused by the incorrect response activation produced on incompatible trials that are absent in compatible trials due to the facilitation from direct-route processing (Ridderinkhof et al., 2004c). However, as will become clear in the next section, essential information about the temporal aspects of information processing in conflict situations is lost if behavioral analyses are restricted to the mean interference effect.

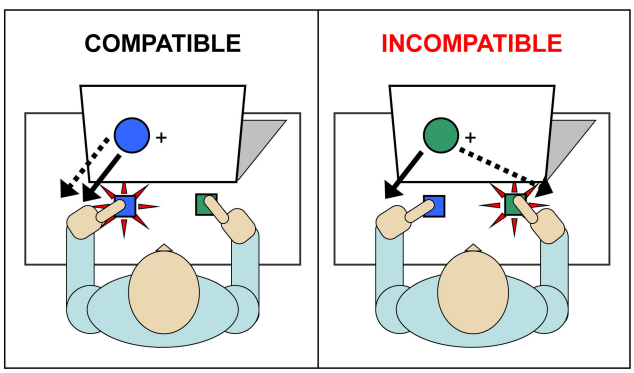



**FIGURE 1 | Simon task.** Participants press the left button to a blue circle and a right button to a green circle (dashed line). Responses are also driven by an irrelevant stimulus dimension, circle location, as indicated by the solid line. For compatible trials, both relevant (color) and irrelevant (location) stimulus dimensions activate the correct action. On incompatible trials, the irrelevant dimension activates an incorrect response tendency, which interferes with selection of the correct response.

## THE ACTIVATION–SUPPRESSION MODEL

Interestingly, many Simon studies that compared fast versus slower responses reported a reduction of the Simon effect if RT is relatively long. These essential temporal dynamics are revealed if the Simon effect is computed as a function of response time, but easily missed if the overall mean Simon effect is taken as the dependent measure of interference control. *The activation–suppression model* (see **Figure 2**), a recent extension of the dual-route architecture, and a related analytic technique were developed to incorporate the temporal dynamics that underlie the expression of impulsive errors followed by a gradual build-up of selective suppression as an act of cognitive control (Ridderinkhof, 2002b; Ridderinkhof et al., 2004c). Based on these temporal aspects, the model predicts that faster reactions on conflict trials should be more vulnerable to impulsive actions that are captured by the irrelevant stimulus dimension. Conversely, the model asserts that slower reactions on incompatible trials are less likely to be negatively impacted by incorrect action impulses because selective suppression had time to accrue to counteract these involuntary impulses. Given that the suppression mechanism needs time to become effective, slower responses benefit more from the effect of selective suppression than faster responses. Therefore, slow responses show a relatively less pronounced interference effect. For example, in the Eriksen flanker task, the interference effect increases with RT, but to a lesser extent for slower RT segments compared to faster segments (Wylie et al., 2007, 2009a,b). In the Simon task, the interference effect reduces in absolute terms, and can even reverse (i.e., go below 0) over time. Here, the slowest RT segment is associated with the smallest Simon effect in absolute value.

Thus, the activation–suppression model refines dual-route models of interference effects by incorporating specific hypotheses about the temporal dynamics of incorrect response activation followed by top-down suppression of unwanted impulsive actions. As illustrated next, the time-courses of these dissociable processes are

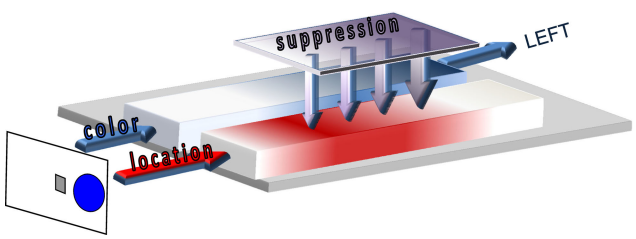



**FIGURE 2 | Activation–suppression model.** The relevant stimulus dimension (color blue) is processed by the slow deliberate route (represented in blue) while the irrelevant location dimension (right location activating the right hand) is processed by the fast direct route (in red). Selective suppression of the location-based activation by the inhibition module (represented in purple) needs time to build up, and facilitates the selection and execution of the correct left-hand response.

masked using overall mean interference effects, but exposed using RT distributional analyses of fast errors and interference effects, respectively. These analyses empirically show that incompatible trials yield an early automatic response impulse that is distinct from a later controlled top-down response suppression mechanism.

### CONDITIONAL ACCURACY FUNCTIONS REVEAL IMPULSIVE RESPONSE ACTIVATION

According to the activation–suppression model, the susceptibility to react impulsively is revealed by the relation between fast errors and response speed. Stronger initial impulsivity on incompatible trials is expressed by an increase in the proportion of fast errors as less time is available for the build-up of suppression to counter this incorrect activation (Kornblum et al., 1990). Thus, plotting accuracy rates for incompatible trials as a function of RT, using the *conditional accuracy function* or CAF, provides a means for studying the strength of automatic responses that are captured in conflicting situations, with stronger capture associated with a higher frequency of fast response errors.

Notably, the strength of automatic response capture is sensitive to various experimental factors that directly affect conflict processing. Wylie et al. (2009b) used an Eriksen flanker paradigm that required participants to respond to a central arrow stimulus while ignoring flanking arrow distracters. These flanking arrows signaled either the same (i.e., are compatible $\rightarrow \rightarrow \underline{\rightarrow} \rightarrow \rightarrow$) or signaled the opposite response as the target arrow (i.e., are incompatible $\rightarrow \rightarrow \underline{\leftarrow} \rightarrow \rightarrow$). Responses are typically slowed and less accurate when the flanking arrows point to the opposite direction as the central arrow, thus inducing conflict. Participants performed the Flanker task under instructions that either emphasized speed or accuracy of responses. Distributional analyses were used to investigate the temporal dynamics of direct response activation on conflict trials on which target and flankers signal opposite responses. According to the activation–suppression hypothesis, the proportion of fast errors reflects the strength of initial response capture by the incompatible flankers, quantified by the slope of the first segment of the CAF connecting the first two points (see **Figure 3**). Steeper CAF slopes reflect a higher proportion of fast errors, suggestive of stronger initial activation of the incorrect response (Ridderinkhof, 2002b; Wylie et al., 2009b). As can be seen in **Figure 3**, responses on compatible trials are associated with near perfect accuracy. In contrast, incompatible flanker trials produced significantly more errors (i.e., steeper positive-going slope) in the fast proportion of the RT distribution. Interestingly, the effect of flanker incompatibility on making impulsive errors was modified by task instructions. An emphasis on the importance of response speed greatly increased errors on conflict trials compared to the instruction condition that stressed accuracy. This was especially true for the fast proportion of responses on incompatible trials (Band et al., 2003; Wylie et al., 2009b).

The speed-accuracy study by Wylie et al. (2009b) showed that response strategies affect the susceptibility to making fast impulsive errors in conflict situations. In addition to these macro-adjustments that involve long-term modifications in speed-accuracy strategy, CAFs also revealed micro-adjustments on a trial-by-trial level. These between-trial adjustments in interference control on trial N are triggered by incompatibility on trial N-1. Ridderinkhof (2002b) administered a version of the Simon task to a sample of healthy participants and focused on accuracy rates on trials that were preceded either by a correct response or by a response error. Between-trial adjustments were evidenced by an increase in accuracy for fast responses from chance level on post-correct trials to about 75% on trials that followed an error. This increase in accuracy on post-error trials relative to post-correct trials might be related to a top-down controlled shift in speed-accuracy trade-off because of post-error slowing. Another example of between-trial control is presented by Stins et al. (2007) who also plotted accuracy as a function of response speed. They observed that the accuracy of fast responses to incompatible flanker trials that were preceded by compatible trials was as low as 26%, which is well below chance

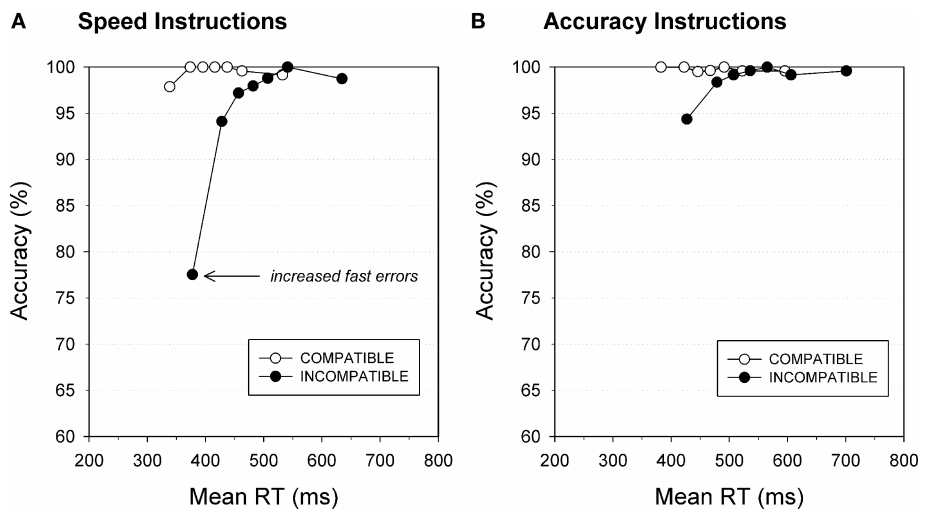



**FIGURE 3 | Task instructions affect fast errors.** Conditional accuracy functions illustrate the increased tendency to commit fast errors to incompatible flanker (black circles) compared to compatible flanker trials (white circles). The proportion of fast errors on conflict trials increases under instructions that emphasize speed **(A)** compared to accuracy of responding **(B)**. Figure modified after Wylie et al. (2009b).

level. In contrast, accuracy for fast incompatible trials following incompatible trials increased to up to 50%. Wylie et al. (2010) confirmed that this between-trial conflict adaptation is preserved in individuals diagnosed with Parkinson's disease. **Figure 4** illustrates that healthy participants as well as Parkinson's patients performing on a Simon conflict task essentially show a comparable pattern of reduced direct response activation on incompatible trials following an incompatible trial (i.e., after a conflict situation, **Figure 4B**) compared to incompatible trials preceded by a compatible trial (**Figure 4A**).

That same clinical study also revealed individual differences in the susceptibility to making fast impulsive errors within the group of Parkinson's patients (Wylie et al., 2010). Forty-five patients with mild to moderate motor symptoms performed the Simon task and were divided into three subgroups that reflected relatively less severe, moderately severe, and most severe motor symptoms according to a clinical motor rating scale. CAFs were computed to study impulsive errors on incompatible trials. Group analyses revealed that patients with the most severe motor symptoms committed a significantly higher proportion of fast impulsive errors compared to the two subgroups with less severe motor symptoms. Thus, the activation of unwanted response impulses, as measured by CAFs in the Simon task, is sensitive to individual as well as group differences in the susceptibility to impulsive action selection (Wylie et al., 2010).

The above studies illustrate the usefulness of distributional analyses to reveal a pattern of within- and between-trial adjustments that affect the susceptibility of making fast impulsive errors that are driven by irrelevant information. Note that analyses of overall accuracy performance do not reveal these temporal effects on behavior. The next section provides an overview of studies that used distributional analyses to quantify the temporal aspects of selective suppression to counteract the activation of unwanted response tendencies in an attempt to resolve response conflict.

## DELTA PLOTS REVEAL THE DYNAMICS OF SELECTIVE ACTION INHIBITION

A delta plot is a graphical representation that displays the temporal dynamics of the RT difference between two experimental conditions (the *delta* value) as a function of response speed (see **Figure 5A**; (de Jong et al., 1994; Ridderinkhof, 2002a,b). Take two hypothetical experimental conditions, X and Y, where one condition is associated with slower responses than the other. The typical pattern is that the proportional difference in RT between the two conditions is similar across RT segments (cf. Luce, 1986). Consequently, the absolute RT difference between conditions X and Y increases from fast to slower segments and the slope values of the lines connecting the delta values are positive going (see Ridderinkhof et al., 2005; Wagenmakers et al., 2005; Speckman et al., 2008). If the two RT distributions belong to the same family, the delta plot should be linear. In contrast, conflict paradigms like the Eriksen flanker task and the Simon task typically yield a different delta-plot pattern. Here, delta slopes in the slower RT segments tend to level off or even go negative, indicating that the differences between response latencies in the X and Y conditions decrease instead of increase as a function of RT.

The activation–suppression hypothesis postulates that this leveling-off is indicative of selective suppression by a top-down control mechanism. This selective suppression acts to resolve response interference by counteracting the initial (incorrect) response activation that is activated by the direct processing route. In contrast to the rapid engagement of the response capture mechanism, top-down suppression takes time to build-up and, therefore, is most effective for responses that are relatively slow (Burle et al., 2002; Ridderinkhof, 2002a; Ridderinkhof et al., 2004c; Forstmann et al., 2008a,b; Wylie et al., 2010). For instance, the faster one responds, the less likely it is that suppression will have accrued to a level that is sufficient to counteract response capture. Rather, slower responses are more likely to benefit from the build-up of suppression to resolve interference. Consequently, correct slow responses to compatible

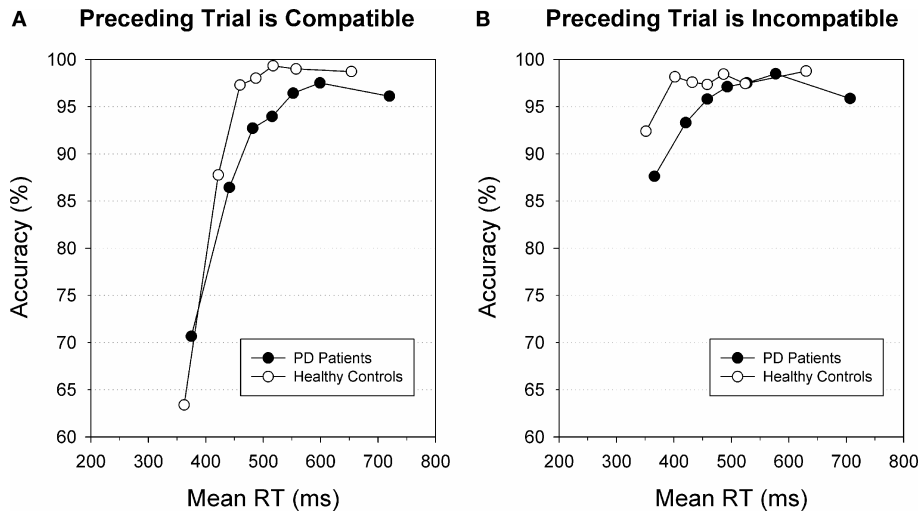



**FIGURE 4 | Effect of preceding trial on fast errors.** Conditional accuracy functions for incompatible trials preceded by compatible trials **(A)** or by incompatible trials **(B)** for individuals diagnosed with Parkinson's disease (black circles) and healthy controls (white circles). For both groups, fast responses on incompatible trials are associated with increased error rates, notably on trials that follow compatible trials. Both groups displayed these between-trial adjustments. Figure modified after Wylie et al. (2010).

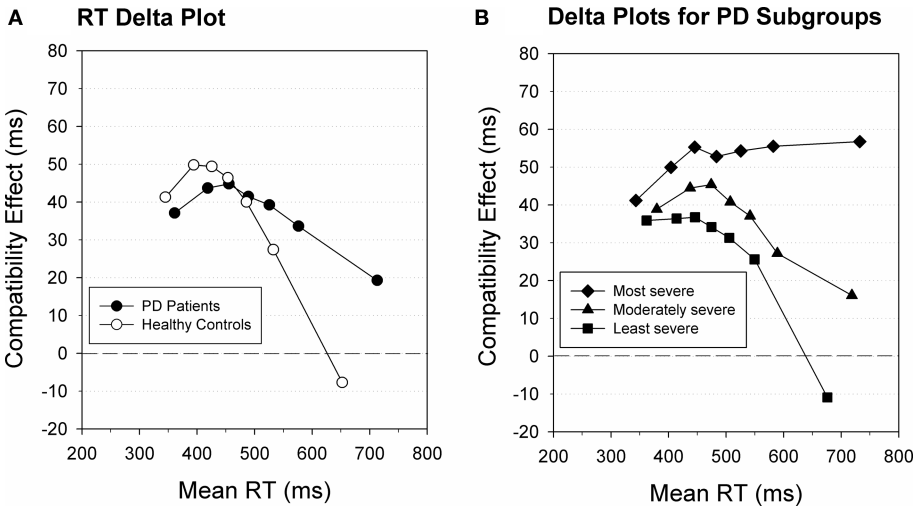



**FIGURE 5 | Delta plot for RT. (A)** Delta plot illustrating impaired selective suppression in patients diagnosed with Parkinson's disease (black) compared to age-matched healthy controls (white). **(B)** Delta plot illustrating impaired selective suppression in mild to moderate PD patients associated with relatively more severe clinical symptoms compared to patients with less severe symptoms. Figure modified after Wylie et al. (2010).

stimuli will be less facilitated by position-driven response capture, whereas correct slow responses to incompatible stimuli will be less delayed. Thus, given these dynamics over the course of a trial, interference effects are reduced by selective response inhibition more in slow than in fast responses. Consequently, Simon delta plots should reveal a pattern of reduced interference in slower segments of the RT distribution as the suppression mechanism becomes more fully engaged (de Jong et al., 1994; Ridderinkhof, 2002a,b).

Given that effective selective inhibition results in a more pronounced reduction of interference effects for slow responses, as argued above, then the attenuation of interference effects (i.e., the leveling-off of the delta-plot slope) should be more pronounced in experimental conditions that require selective suppression (Burle et al., 2002; Ridderinkhof, 2002b; Wijnen and Ridderinkhof, 2007; Wylie et al., 2009b). Moreover, the slope value that quantifies the reduction of interference at the slowest segment of the RT distribution serves as a sensitive metric of the proficiency of selective response inhibition (Burle et al., 2002; Ridderinkhof, 2002b; Ridderinkhof et al., 2004c, 2005). On the basis of the assumptions of the temporal dynamics of online control, derived from the activation–suppression model, we take the slope value of the slowest delta-plot segment to best capture the proficiency of selective suppression. This does not mean that the effects of selective suppression may not become evident in earlier RT segments, but this depends on several factors like the strength of initial response capture, the number of RT segments and the strength of the suppression mechanism.

First, we focus on studies that establish the basic validity of interpreting these delta patterns as reflecting selective suppression, and then cover studies that have used delta-plot measures as a probing tool. Several behavioral studies provided proof of principle for the notion that the delta-slope value at the slowest portions of the delta plot varies as a function of experimental manipulations of inhibitory demands. Increasing the relative probability of incompatible trials generally reduces the overall mean Simon effect (Stürmer et al., 2002) and yields stronger selective suppression of location-driven activation for slower responses, reflected by a steeper negative-going delta-plot slope as RT slows (Ridderinkhof, 2002b). Similarly, a dramatic negative-going delta slope occurs much earlier in the RT distribution on trials that follow incompatible trials compared to trials following compatible trials (Ridderinkhof, 2002b; Wylie et al., 2010). A compelling illustration that confirms the link between delta-slope values and inhibitory control is provided by a study comparing Flanker effects in children diagnosed with AD/HD and matched controls (Ridderinkhof et al., 2005). Although overall group differences in mean flanker effects were marginal, delta-plot analyses uncovered a pattern of less proficient interference control among children with AD/HD. Moreover, among children with AD/HD, delta slopes were more negative-going after administration of methylphenidate, suggesting an ameliorative effect on the response inhibition deficit. These findings support current theories that emphasize a deficit in response inhibition as a fundamental neurocognitive impairment in AD/HD (Nigg, 2001).

The studies above confirmed the notion that variations in the temporal dynamics of interference effects reflect demands on selective suppression that are sensitive to experimental manipulations as well as group differences in inhibitory proficiency. The activation–suppression model inspired several experiments that used the delta-plot technique to further probe variations in selective suppression. For example, Burle et al. (2005) explored the time-course of selective suppression by systematically varying the temporal overlap between the onsets of the relevant and irrelevant stimulus dimensions. They revealed a clear reversal of the interference effect (i.e., negative slope values) if the irrelevant information was presented before the relevant information; long RTs were associated with a negative interference effect, consistent with more effective suppression. Furthermore, manipulation of the spatial configuration of stimuli in a Simon task (e.g., vertical versus

horizontal displacement) yields a negative-going delta-plot pattern for horizontally displaced stimuli, but a positive slope for vertically displaced stimuli (Wiegand and Wascher, 2005, 2007). This suggested a stronger need for selective suppression when stimuli are presented in the left or right visual field that is directly corresponding to the location of the required response hand. Furthermore, compared to a rest condition, physical exercise reduced the proficiency of selective suppression (Davranche and McMorris, 2009). Finally, selective suppression, as revealed by negative-going delta slopes, is not restricted to manual actions since a similar pattern has been obtained for foot responses (Davranche et al., 2009) and for eye movements (Wijnen and Ridderinkhof, 2007).

Delta-plot analyses are not only sensitive to experimental factors, but also provide a sensitive quantitative metric to study individual differences in the proficiency to suppress interference that arises from the activation of conflicting responses. In their developmental study of participants ranging in age between 19 and 82 years, Juncos-Rabadán et al. (2008) applied the delta-plot technique on data obtained with a variant of the Simon task. They observed that for 19- to 26-year-old participants the magnitude of the Simon interference effect decreased with response speed (i.e., showing a negative-going delta slope). However, interference increased as a function of RT among 70- to 82-year-olds. This developmental pattern supports the conclusion that selective suppression declines with older age. Focusing on children, Bub et al. (2006) presented the standard Stroop task (Stroop, 1935) to a group of children between 7 and 11 years of age. Interference was induced by having the children name the print color of a (non-color) word while resisting the prepotent tendency to read the word aloud. Younger children displayed increased mean interference on conflict trials, both in terms of increased error rates and RT slowing. However, plotting Stroop interference effects as a function of response speed revealed a reduction of the interference effect for slower responses that was at least as pronounced for younger as for older children (Bub et al., 2006). Taken together, these developmental studies indicate that the proficiency of selective suppression emerges across the early years of development but is vulnerable to decline with advanced age. In addition, several clinical studies that employed the delta-plot technique have pointed to increased difficulty in resolving interference in several populations, such as patients diagnosed with mild cognitive impairment (Wylie et al., 2007) and Parkinson's disease (see **Figures 5A,B**; Wylie et al., 2009a,b, 2010).

### ALTERNATIVE ACCOUNTS

The behavioral studies discussed above illustrate the effectiveness of distributional analyses of response errors and response speed to quantify the temporal aspects and to dissociate the activation of incorrect responses and its subsequent selective suppression. The activation–suppression model provides a powerful framework to interpret the present results; the patterns derived from distributional analyses conform accurately to the model's predictions and the results point to the hypothesized variations in inhibitory control as a function of experimental factors. However, we recognize that accounts other than the activation–suppression model may potentially explain the reduction of interference effects over time. For example, the leveling-off of interference effects toward the slow end of the RT distribution might result from a process of passive decay of the initial response activation by the direct route, instead of the activation of a top-down suppression mechanism (Hommel, 1994). Although passive decay might potentially result in an attenuation of interference effects, mere decay is unlikely to account for reversals of the Simon effect. Such negative Simon interference effects are evident in the delta plot for a group of healthy control participants (see **Figure 5A**), and for both controls and Parkinson patients when plotting the Simon delta values for trials that are preceded by incompatible trials (e.g., Wylie et al., 2010). An interference effect below 0 would require an active process of suppression of the incorrect response activation during incompatible trials that leads to faster response selection on these trials. Similarly, computational modeling architectures that do not take into account active suppression units have not been very successful in describing negative interference effects (e.g., Davelaar, 2008).

Support for the involvement of active top-down suppression to resolve conflict is provided by psychophysiological experiments that focused on neurocognitive mechanisms that underlie the expression and suppression of impulsive actions. The next sections provide a selective review of studies that used individual differences in RT-distribution parameters to probe the efficiency and temporal dynamics of selective response inhibition and to guide the detailed analysis of neuroimaging data.

## EMG ERRORS REPRESENT DETECTED AND CORRECTED IMPULSIVE ACTIVATION

The notion of incorrect response activation refers to a covert process that is not directly observable in behavior. Electrophysiological measures provide a useful tool to reveal such covert processes. In this respect, recording the electromyographic activity of the muscles involved in response execution has proved useful (Coles et al., 1985; Smid et al., 1990; Burle et al., 2002). Compared to other potential measures of response activation like electroencephalography (EEG), EMG provides several major advantages. First, EMG activity can unequivocally be associated with motor responses, whereas it is harder to map EEG measures directly to motor responses because of volume conduction. Second, thanks to its excellent signal-to-noise ratio, EMG activity can be detected on a trial-by-trial basis, hence avoiding spurious effects of averaging (see Burle et al., 2008; **Figure 6** for an example).

### ACTIVATION OF INCORRECT ACTIONS: PARTIAL EMG ERRORS

In the context of conflict tasks, EMG recordings have revealed that on a proportion of trials in which the correct response was issued, phasic EMG activities occurred in both the correct and incorrect response agonist muscles (Coles et al., 1985; Eriksen et al., 1985). Following these seminal papers, incorrect EMG was studied in more detail by Smid et al. (1990) who distinguished trials on which the incorrect EMG occurred *before* the correct response, from trials on which the incorrect EMG occurred *after* it (see **Figure 6A**). They reported that only the first trial category was contributing to conflict effects, making this category of special interest, whereas the second category likely reflects mere motor noise. Follow-up studies, although sparse, have largely confirmed the particular interest of such "partial error" trials. **Table 1** presents the percentage of partial errors on compatible and incompatible trials observed across different tasks and studies. Although the absolute percentage

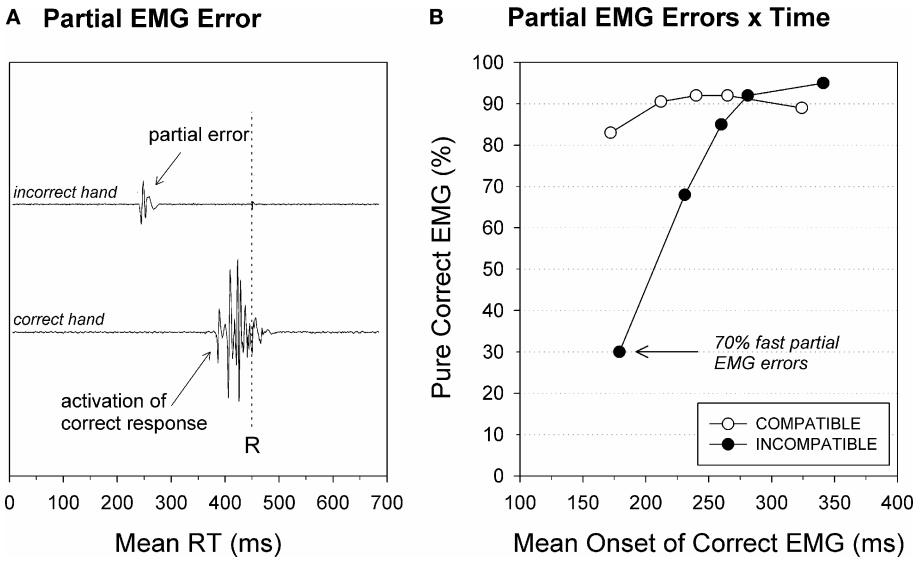



**FIGURE 6 | Partial EMG errors. (A)** Electromyogram (EMG) from muscles controlling the incorrect (upper trace) and the correct response (lower trace). The vertical dashed line indicates the mechanical button-press response. The correct overt response was preceded by partial EMG activity in the muscle that controls the incorrect response that is too weak to trigger an overt error. **(B)** Fast responses on incompatible trials yield a high percentage of stimulus-driven partial EMG errors of the incorrect hand before the correct button press is emitted. Re-analyzed data published by Burle et al. (2002).

**Table 1 | Percentage of partial EMG errors for compatible and incompatible trials reported in the literature.**

| | Compatible (%) | Incompatible (%) |
|---|---|---|
| **SIMON TASK** | | |
| Burle and Bonnet (1999) | 15 | 28 |
| Hasbroucq et al. (1999) | 10 | 17 |
| Burle et al. (2002, exp. 1) | 4 | 7 |
| Burle et al. (2002, exp. 2) | 5 | 13 |
| Hasbroucq et al. (2009, exp. 1) | 15 | 26 |
| Hasbroucq et al. (2009, exp. 2) | 17 | 23 |
| **ERIKSEN TASK** | | |
| Smid et al. (1990) | 8 | 23 |
| Burle et al. (2008) | 14 | 22 |
| Burle and Hasbroucq (submitted, exp. 1) | 12 | 27 |
| Burle and Hasbroucq (submitted, exp. 2) | 20 | 26 |
| **FITTS TASK** | | |
| Hasbroucq et al. (2001) | 11 | 19 |

*A few other studies recorded partial errors in conflict tasks but did not report percentages.*

may vary from study to study, it is clear that the percentage of partial EMG errors is always larger on incompatible compared to compatible trials. Such an increase in incorrect response activation on incompatible trials directly supports a key assumption of the dual-route model (Kornblum et al., 1990), namely that the irrelevant aspect (position, flankers, etc.), more or less automatically, activates the response to which it is associated, either by instruction (e.g., in the Eriksen task) or by habit strength (e.g., in the Simon and Fitts tasks).

Further arguments supporting the idea that those incorrect response activations are true partial errors come from EEG and sequential effect analysis. First, incorrect EMG occurring before the response (i.e., partial errors) are followed by a small but robust error negativity (Scheffers et al., 1996; Vidal et al., 2000; Masaki et al., 2007; Burle et al., 2008; Roger et al., 2010), whereas incorrect EMG occurring after the correct response are not (Allain et al., 2004). Furthermore, we recently showed that, like overt errors, partial errors are followed by response slowing, expressed as a RT lengthening after a trial containing a partial error (although this post-error slowing is much smaller than after an overt error; Allain et al., 2009).

Besides this rather "static" view of incorrect response activation, a more dynamic view is obtained by the CAF that plots the percentage of correct responses as a function of the latency of the response. Instead of taking into account only overt behavior (i.e., RT), one can plot the frequency of incorrect response activation (EMG measure), as a function of the activation latency. An example of such an analysis in a Simon task is presented in **Figure 6B**. As for "traditional" CAFs, it appears that relatively slow response activations are mainly correct, with an asymptote around 5% errors, but that fast response activations are much less accurate. Interestingly, CAFs based on partial-errors reveal that fast incorrect response activations are clearly above chance level on incompatible trials (70% partial EMG errors), whereas classical CAFs of RT often indicate chance level accuracy. Such a high percentage of incorrect response activation discards the possibility that partial errors simply reflect guesses, since guesses should yield accuracy rates around chance level. To the contrary, such a pattern clearly indicates that these incorrect EMGs are mainly stimulus driven, and hence reflect the activation of response by the irrelevant dimension of the stimulation display.

### SELECTIVE SUPPRESSION OF INCORRECTLY ACTIVATED EMG RESPONSES

Since partial EMG errors truly reflect stimulus-driven incorrect response activation, as argued above, they offer an opportunity to directly test the response suppression hypothesis. Indeed, on trials with a partial EMG error, the incorrect response has been largely activated (as it reached late stages of the reaction process) but was nonetheless appropriately suppressed, since the correct response was finally issued. According to the activation–suppression model (Ridderinkhof, 2002a), one should thus predict steeper negative-going delta-plots for trials containing partial EMG errors. This prediction was tested directly (Burle et al., 2002), and the results are presented in **Figure 7**. As one can see, when all trials are taken into account, that is pooled across trials with and trials without partial EMG, the delta-plot slopes are negative-going (black circles) as usually reported in the literature. Interestingly, when excluding trials with partial EMG activation of the incorrect response, the last delta-plot slope is much less negative going (white circles). This finding supports the idea that the reduction of the interference effect, reflected by the delta slope, is related to the detection and suppression of incorrect responses that are activated up to the level of the incorrect muscles.

In sum, analysis of EMG patterns has provided direct evidence for the subthreshold activation of impulsive actions up to the level of the muscles that control the motor response and its subsequent selective suppression. The following section presents an overview of functional as well as structural imaging studies that used the parameters generated by the activation–suppression model in search of brain areas that are associated with activation and selective suppression.

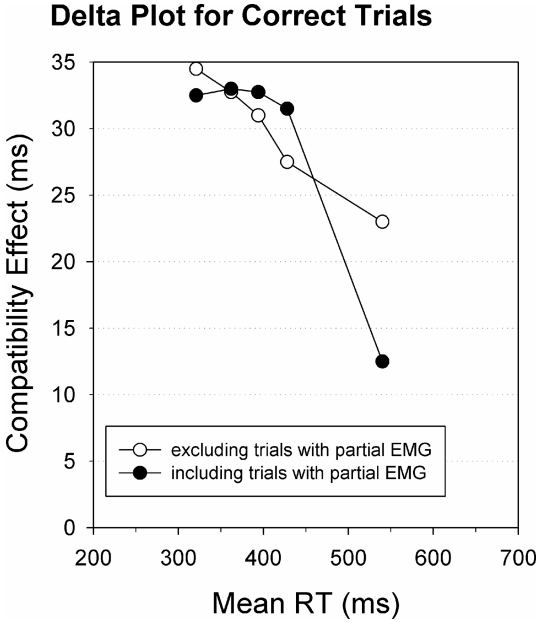



**FIGURE 7 | RT delta plots including and excluding partial error trials.** RT delta plots for trials with a correct button-press response. Including trials with partial EMG errors yields a steeper negative slope value (black circles) compared to excluding EMG error trials (white circles). Re-analyzed data published by Burle et al. (2002).

## NEURAL MECHANISMS UNDERLYING ACTIVATION AND INTERFERENCE CONTROL

Despite the wealth of literature on conflict tasks such as the Stroop task and Eriksen flanker variants, neuroimaging studies using the Simon task are still relatively rare. The activation pattern most commonly reported in conflict tasks consists of a fronto-parietal and fronto-striatal network, including the medial frontal cortex, the lateral prefrontal cortex, the posterior parietal cortex, as well as the dorsal striatum and the subthalamic nucleus (STN) (e.g., Pardo et al., 1990; Peterson et al., 2002; Hazeltine et al., 2003; Schumacher et al., 2007; for a meta-analysis see Nee et al., 2007; for a reviews see Ridderinkhof et al., 2004b, 2010). More specifically, processes like monitoring of response conflict and response selection are believed to play a crucial role in task performance and have been associated with functioning of the medial frontal cortex (Ridderinkhof et al., 2004a). Others have pointed to the essential role of the lateral prefrontal cortex in implementing top-down control after experiencing response conflict or suppressing unwanted response tendencies (Kerns et al., 2004). Finally, the posterior parietal cortex subserves visuo-motor transformations via middle occipital and inferior temporal cortices important for attentional control (Kastner and Ungerleider, 2000).

A central question is to what extent this general network reveals task-specific or task-independent selection, control, and suppression of actions (see also Schumacher et al., 2003). The Stroop task engages several putative control processes, including monitoring of response conflict, selection between competing responses, and selective response inhibition, but the relevant contrast in this task may capture other processes as well (e.g., conflict at the semantic level). A recent quantitative meta-analysis by Nee et al. (2007) that included fMRI data from several Stroop tasks revealed a strong left lateralized network including the left middle frontal gyrus, the insula, as well as the left posterior parietal cortex. Contrary to Stroop stimuli, the target and flanker stimuli in the Eriksen flanker paradigm are presented within the same stimulus dimension without requiring semantic processes. With respect to Flanker task performance, meta-analyses shows mainly right lateralized activation clusters in the middle frontal gyrus and the insula (Nee et al., 2007). Interestingly, these patterns are somewhat comparable to activations found with stimulus–response compatibility tasks such as the Simon task (Peterson et al., 2002; Liu et al., 2004).

### DATA-DRIVEN VERSUS COVARIATE-BASED fMRI APPROACH

The heterogeneity of the activation patterns on conflict tasks reported above might be driven by two underlying sources of variance. First, most studies do not take into account considerable differences in interference control across individuals. When activation levels are pooled over participants, the variance related to individual differences may mask the detection of activation patterns that emerge when contrasting incompatible and compatible trial types. The second source of variance that might have contributed to inconsistent findings across fMRI reports relates to the calculation contrasting brain activation derived from incompatible trials and compatible trials, i.e., the mean interference effect on brain activity. As outlined in the previous sections, overall mean interference effects mask the vital dynamic aspects of response impulse activation followed by selective suppression of this activation. A

potentially more effective approach is to relate brain activation with specific behavioral parameters as specified by the activation–suppression model and as revealed by distributional analyses. Recent neuroimaging studies by Forstmann et al. (2008a,b) adopted this approach by focusing on individual differences in interference control, taking into account the temporal dynamics of activation and subsequent suppression of incorrect response activation. Twenty-four participants performed a variant of the Simon task (see **Figure 1**) in which the designated response is indicated by an aspect of the imperative signal (its color), but competing response impulses were elicited by the task-irrelevant spatial location of the stimulus. In a separate neutral block without spatial interference, the participants responded to the color of a circle presented in the middle of the screen. The following sections provide answers to questions like, "does the activation of certain brain areas covary with the strength of initial capture by an incorrect response impulse?" and "are there dissociable brain areas whose activation is linked to the proficiency of selective suppression as revealed by negative-going delta slopes?"

#### IMPULSIVE ERRORS: ACTIVATION IN THE RIGHT PRE-SMA

In line with the predictions of the activation–suppression model, fast responses on conflict trials were considerably more prone to errors than slower incompatible responses. The error slope value derived from the fast RT segment for incompatible trials was used to infer the individual's susceptibility to making fast, impulsive errors (red slope depicted in **Figure 8**, upper panel). Individuals showing a relatively steeper positive-going error slope value for the fast RT segment had a stronger tendency to emit fast and unintentional impulsive responses that are driven by the irrelevant stimulus dimension. Analyses of brain activation focused on the contrast between incompatible versus neutral trials. To study the neural substrates of impulsive action selection, individual slope values derived from the fast segment of the CAF for incompatible responses were entered as covariates in the fMRI regression analyses. Steeper positive-going error slopes were associated with enhanced activation in the right pre-supplementary motor area (pre-SMA, see **Figure 8**, middle panel). The pre-SMA was found to be active on correct incompatible trials that are associated with increased error rates (Forstmann et al., 2008b). This pattern indicates that individuals who are highly susceptible to making fast impulsive errors show increased pre-SMA activity reflecting the elevated need to select a correct response in conflicting situations. Notably, the relation between errors and percent signal change derived from the pre-SMA was evident only for the fast RT segment, and was absent for slower responses, a dynamic pattern that is in line with the tenets of the activation–suppression model.

#### SELECTIVE SUPPRESSION: ACTIVATION IN THE RIGHT IFC

Next, the covariate-based fMRI approach was used to identify the neurocognitive correlates of selective response inhibition to resolve conflict arising from the activation of involuntary action impulses. To this end, the individual reduction of the Simon effect was quantified as a function of RT; the value of the delta slope derived from the slow segment of the RT distribution (red slope depicted in **Figure 8**, upper-right panel). Entering individual delta-slope values into the fMRI regression analyses revealed significant activations in the right inferior frontal cortex (IFC, BA 44) bordering the right anterior insula (Forstmann et al., 2008b). This pattern reveals that individuals showing relatively negative-going delta-slope values derived from the slowest segment of the RT distribution present enhanced activation in the right IFC. The specificity of taking into account the temporal aspects of interference control is underscored by the null-results of two additional analyses. First, the two delta-slope values derived from faster RT segments did not co-vary significantly with brain activation. A second null finding was obtained when individual mean interference effects (i.e., overall average Simon effects) were entered as covariates in the fMRI regression model. These findings provide direct evidence that activation in the right IFC varies as a function of individual efficacy to resolve response conflict that arises from involuntary action tendencies.

#### STRUCTURAL CONTRIBUTIONS TO SELECTIVE SUPPRESSION

The fMRI studies discussed above capitalized on individual differences in the activation and selective suppression of incorrect responses and identified the involvement of the pre-SMA and the right IFC, respectively. This section highlights a study that has specified the relationship between behavioral parameters derived from the RT distributions and the structural integrity of the brain. The specific fMRI activation patterns in the pre-SMA and right IFC were replicated and extended in a second covariate-based study using a modified version of the Simon task (Forstmann et al., 2008a). Again, individual distribution parameters (i.e., the delta-slope value of the slow RT segment) predicted activation in the right IFC, confirming the close match between the proficiency of selective suppression and right IFC function. As a novel approach, structural data on the density of coherent white matter tracts were obtained using diffusion tensor imaging (DTI). The dependent measure of interest was fractional anisotropy (FA), reflecting the degree of diffusion anisotropy within a voxel that depends on microstructural features of the tissue that includes fiber properties and tract coherence. First, individual RT distribution parameters were used to classify subgroups of good and poor inhibitors based on a median split of the slopes from the slowest segment of the delta plot. The connectivity data revealed higher FA values in the right anterior part of the inferior fronto-occipital fasciculus (FOF) and in the precuneus for good versus poor performers (see **Figure 9**). This pattern reveals a tight linkage between brain structure (FA values in the anterior part of the FOF) and the proficiency of selective response inhibition. This pattern of connectivity differences in the anterior part of the FOF closely resembled the corresponding pattern for functional activation of the right IFC. Finally, our third hypothesis stated that the right IFC BOLD activation and the right IFC local structural differences should correlate positively. This hypothesis was also confirmed by the present data pattern (**Figure 9**). These finding reflect a systematic linkage of individual differences in selective response inhibition at the behavioral, functional, and structural level, as supported by independent techniques (Aron et al., 2007; Forstmann et al., 2008a,b; Neubert et al., 2010).

#### THE SUBTHALAMIC NUCLEUS: DEEP-BRAIN STIMULATION

According to many contemporary views, the implementation of action selection and inhibition processes is accomplished by cortical–subcortical interactions. To this point, the reviewed studies

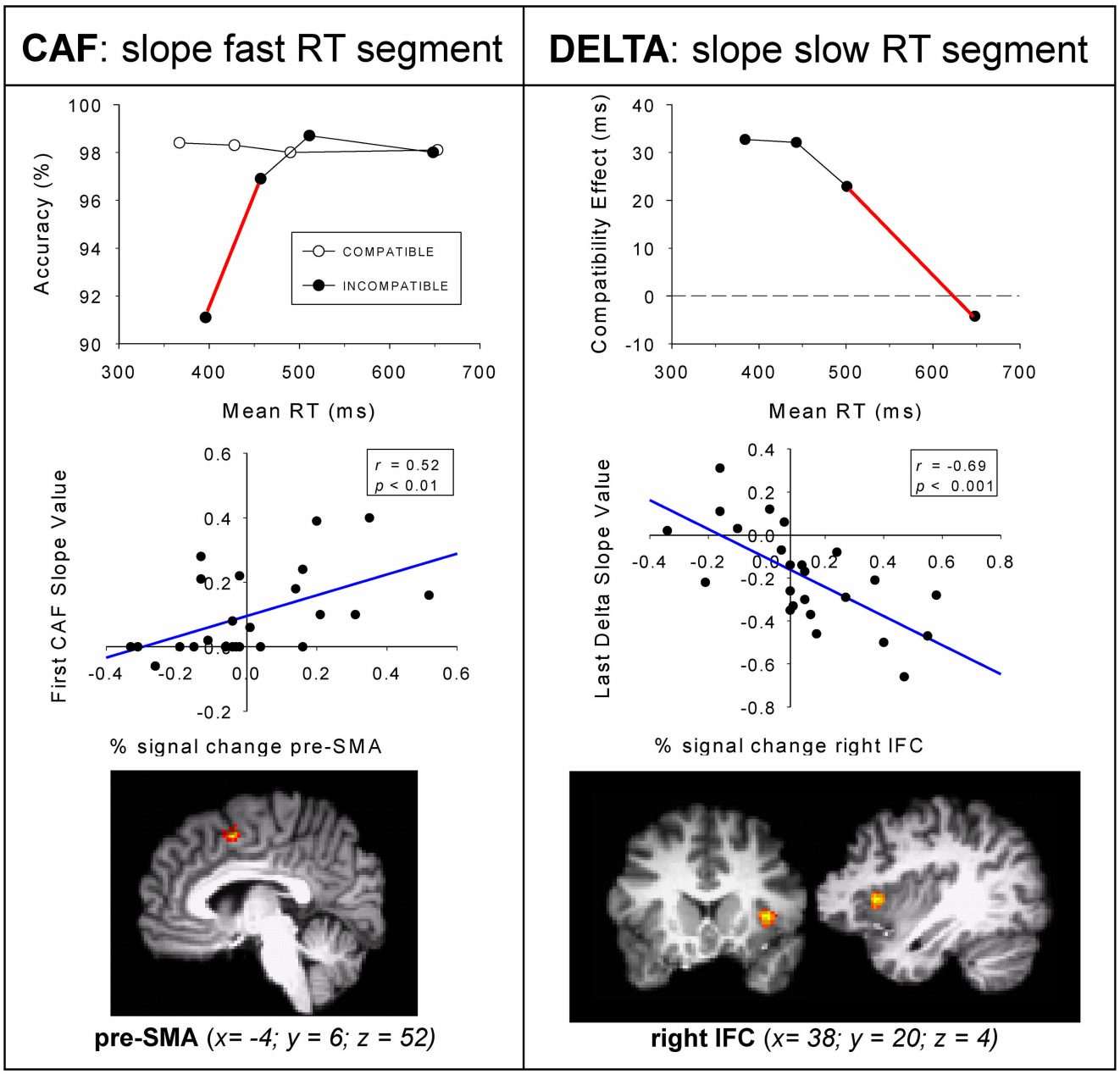



**FIGURE 8 | Slope values predict brain activation.** Covariate analyses with individual behavioral parameters from 24 participants. Top row displays the average conditional accuracy function (CAF) and average delta slopes separated in fast, middle, and slow RT segments. Middle row left shows Pearson correlations between the first CAF slope for incompatible trials and the % signal change (*x*-axis) derived from the pre-SMA. Middle row right shows Pearson correlations between the last delta slope (*y*-axis) and % signal change (*x*-axis) derived from the right IFC. Figure modified after Forstmann et al. (2008b).

indicate key involvement of cortical activity in pre-SMA and right IFC in action selection and selective inhibition of action, respectively. Both of these cortical areas project directly to the basal ganglia, a network of subcortical structures whose involvement in releasing or preventing movement has been firmly established (see Mink, 1996, for a review). Of particular interest in the present review are so-called hyper-direct pathways connecting pre-SMA and right IFC to the STN that have been postulated to play an important role in action control processes. As a result, the STN may be unique among basal ganglia structures in its involvement in both rapid action selection and top-down selective suppression processes (Frank, 2006; van den Wildenberg et al., 2006).

Studies of patients diagnosed with Parkinson's disease who have been surgically implanted with deep-brain stimulation (DBS) electrodes in the STN offer a unique opportunity to examine the role of STN modulation on action selection and inhibition processes. In a recent study, we used distributional analyses of the Simon effect to investigate the role of STN stimulation in medicated Parkinson's

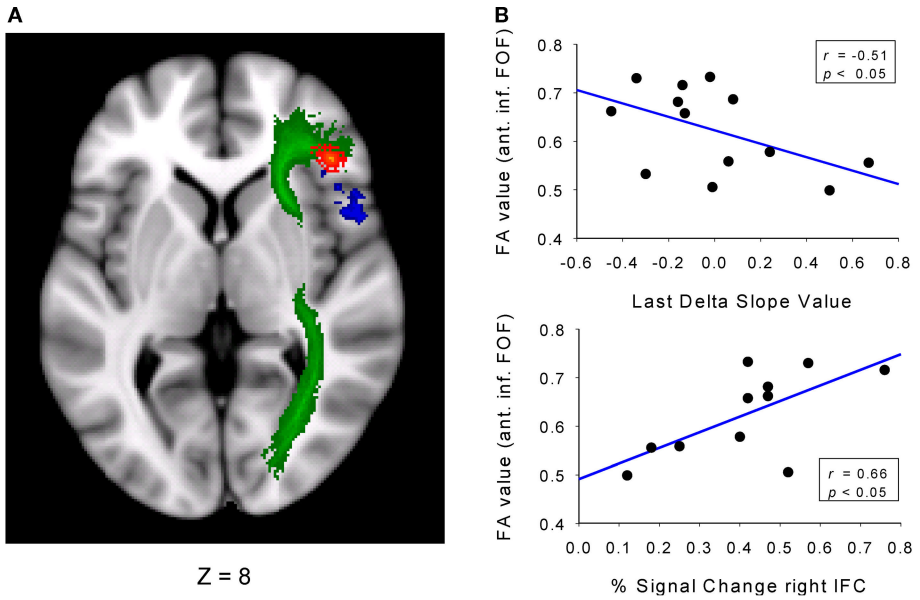



**FIGURE 9 | Relation between delta slope, rIFC activation, and structural measures. (A)** Axial view showing the inferior fronto-occipital fasciculus (FOF, in green). Contrasting good versus poor inhibitors shows increased connectivity in the anterior FOF (in red) and increased BOLD activation in the right IFC (blue). (**B**, upper plot) Pearson correlation between connectivity in the anterior part of the FOF and slope values derived from the delta plot of the last RT segment. (**B**, lower plot) Pearson correlation between connectivity in the anterior part of the FOF and percentage BOLD signal change derived from the right IFC. Figure modified after Forstmann et al. (2008a).

patients on the strength of capture by response impulses and the proficiency of selective suppression (Wylie et al., 2010). Parkinson's patients performed the Simon task under conditions in which their STN was and was not being stimulated. Whereas mean values did not disclose differences in interference effects attributable to STN stimulation, distributional analyses revealed important effects on response capture and selective suppression dynamics. When the STN was not being stimulated, patients showed similar rates of fast errors as healthy controls, suggesting normal susceptibility to impulsive behavior. However, Parkinson's patients were less proficient at suppressing the interference with goal-directed action caused by the activation of the incorrect response impulse (i.e., they showed less reduction of the Simon effect as RT slowed), a finding that replicated previous results in medicated Parkinson's disease patients who were not treated with STN DBS (Wylie et al., 2010). Turning on stimulation to the STN produced two significant effects on performance. First, patients reacted more impulsively compared to when their STN stimulation was turned off as well as compared to healthy controls. That is, they showed an appreciable increase in fast errors on conflict trials. Second, on conflict trials that were responded to correctly, STN stimulation improved the patients' ability to suppress the interference that was generated by the activation of the impulsive action. Thus, STN stimulation both increased impulsive action selection yet improved inhibitory control over the conflict produced by impulsive action tendencies that did not lead to overt response errors. These seemingly paradoxical effects can be understood within the activation–suppression model by asserting that STN stimulation influences two temporally and neurally dissociable aspects of conflict processing, an initial phase related to an individual's susceptibility to response capture by incorrect response impulses and a later phase associated with the engagement and build-up of top-down, selective inhibitory control over conflicting responses. These results fit well with human and animal studies that have linked STN activity and its manipulation to changes in both impulsive behavior and response suppression across a variety of experimental tasks (Eagle et al., 2008; Ballanger et al., 2009). It will be important for future imaging studies to investigate the potential role of pre-SMA and right IFC projections to STN as mediators of action selection and inhibition circuits, respectively.

## CONCLUSION

This selective review of behavioral and psychophysiological correlates of performance on conflict tasks shows that distributional analyses can isolate vital processes that are involved in the cognitive control of interference effects with greater precision than can mean performance measures. A merit of the activation–suppression model is the functional dissociation between the susceptibility of making fast impulsive actions (response capture), and subsequent cognitive control that is engaged to keep unwanted action activation in check to facilitate the selection of the correct action (selective action inhibition). The first process is indexed by a pattern of fast errors on incompatible trials (depicted in a CAF). The second process is reflected by a decrease in the interference effect for slower responses (depicted in a delta plot). Importantly, the activation–suppression model yields behavioral parameters that are sensitive to these temporal dynamics and, as such, provided a useful tool to study the impact of experimental factors and individual differences in action control during conflict.

Model-based analyses that capitalize on individual differences in response capture and subsequent suppression revealed structural as well as functional differences related to specific brain

regions. Parameters indicating response capture and selective inhibition in the Simon task have guided neurocognitive studies highlighting the pre-SMA and right IFC as key nodes for action selection. Direct connections between the pre-SMA and basal ganglia structures (most prominently the anterior dorsal striatum and the STN) serve to keep basal ganglia output in check until intention-driven action selection has completed. Extraneous, impulsive action affordances may capture the action system non-deliberately. When multiple stimulus-action association alternatives compete for activation, the demands on action control are highest, and selecting the appropriate action engages stronger activation of the pre-SMA compared to when response capture is absent. The pre-SMA may act as an action-selection director, modulating the action-selection gate through which the available action affordances are translated into actual actions (Mars et al., 2009; Ridderinkhof et al., 2010). The right IFC is recruited to signal the need for implementation of selective response inhibition (Forstmann et al., 2008a,b, see also Aron et al., 2007; Verbruggen et al., 2010). In particular, individual differences in the efficiency to implement inhibitory control in humans are associated consistently with functional and structural differences in the right IFC. In turn, the basal ganglia serve to keep all responses in check until the final green signal is received from upstream (Frank, 2006; van den Wildenberg et al., 2006; Wylie et al., 2010). In later processing stages, suppression of unwanted responses may be evident at the level of the primary motor cortex (van den Wildenberg et al., 2010) and even shows in suppression of subthreshold EMG activity in incorrectly activated response muscles (Burle et al., 2002).

## ACKNOWLEDGMENTS

Wery P. M. van den Wildenberg, Birte U. Forstmann, and K. Richard Ridderinkhof were supported by grants from the Netherlands Organization for Scientific Research (NWO). This work was supported by a grant (K23AG028750) awarded to Scott A. Wylie by the National Institute on Aging (the content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute On Aging or the National Institutes of Health). Borís Burle was supported by a grant from the European Research Council under the European Community's Seventh Framework Program (FP7/2007-2013 Grant Agreement no. 241077).

## REFERENCES

Allain, S., Burle, B., Hasbroucq, T., and Vidal, F. (2009). Sequential adjustments before and after partial errors. *Psychon. Bull. Rev.* 16, 356–362.

Allain, S., Carbonnell, L., Burle, B., Hasbroucq, T., and Vidal, F. (2004). "Electromyographic activities and the Ne," in *Errors, Conflicts, and the Brain: Current Opinions on Performance Monitoring*, eds M. Ullsperger and M. Falkenstein (Leipzig: MPI of Cognitive Neuroscience), 21–27.

Aron, A. R., Behrens, T. E., Smith, S., Frank, M. J., and Poldrack, R. A. (2007). Triangulating a cognitive control network using diffusion-weighted magnetic resonance imaging (MRI) and functional MRI. *J. Neurosci.* 27, 3743–3752.

Ballanger, B., van Eimeren, T., Moro, E., Lozano, A. M., Hamani, C., Boulinguez, P., Pellecchia, G., Houle. S., Poon, Y. Y., Lang, A. E., and Strafella, A. P. (2009). Stimulation of the subthalamic nucleus and impulsivity: release your horses. *Ann. Neurol.* 66, 817–824.

Band, G. P. H., Ridderinkhof, K. R., and van der Molen, M. W. (2003). Speed-accuracy modulation in case of conflict: the roles of activation and inhibition. *Psychol. Res.* 67, 266–279.

Bub, D. N., Masson, M. E. J., and Lalonde, C. E. (2006). Cognitive control in children: Stroop interference and suppression of word reading. *Psychol. Sci.* 17, 351–357.

Burle, B., and Bonnet, M. (1999). What's an internal clock for? From temporal information processing to temporal processing of information. *Behav. Processes* 45, 59–72.

Burle, B., Possamaï, C. A., Vidal, F., Bonnet, M., and Hasbroucq, T. (2002). Executive control in the Simon effect: an electromyographic and distributional analysis. *Psychol. Res.* 66, 324–336.

Burle, B., Roger, C., Allain, S., Vidal, F., and Hasbroucq, T. (2008). Error negativity does not reflect conflict: a re-appraisal of conflict monitoring and anterior cingulate cortex activity. *J. Cogn. Neurosci.* 20, 1637–1655.

Burle, B., van den Wildenberg, W. P. M., and Ridderinkhof, K. R. (2005). Dynamics of facilitation and interference in cue-priming and Simon tasks. *Eur. J. Cogn. Psychol.* 17, 619–641.

Coles, M. G. H., Gratton, G., Bashore, T. R., Eriksen, C. W., and Donchin, E. (1985). A psychophysiological investigation of the continuous flow of human information processing. *J. Exp. Psychol. Hum. Percept. Perform.* 11, 529–552.

Davelaar, E. J. (2008). A computational study of conflict-monitoring at two levels of processing: reaction time distributional analyses and hemodynamic responses. *Brain Res.* 1202, 109–119.

Davranche, K., and McMorris T. (2009). Specific effects of acute moderate exercise on cognitive control. *Brain Cogn.* 69, 565–570.

Davranche, K., Paleresompoulle, D., Pernaud, R., Labarelle, J., and Hasbroucq, T. (2009). Decision making in elite white-water athletes paddling on a kayak ergometer. *J. Sport Exerc. Psychol.* 31, 554–565.

de Jong, R., Liang, C. C., and Lauber, E. (1994). Conditional and unconditional automaticity: a dual-process model of effects of spatial stimulus–response correspondence. *J. Exp. Psychol. Hum. Percept. Perform.* 20, 731–750.

Eagle, D. M., Baunez, C., Hutcheson, D. M., Lehmann, O., Shah, A. P., and Robbins, T. W. (2008). Stop-signal reaction-time task performance: role of prefrontal cortex and subthalamic nucleus. *Cereb. Cortex* 18, 178–188.

Eimer, M., Hommel, B., and Prinz, W. (1995). S-R compatibility and response selection. *Acta Psychol.* 90, 301–313.

Eriksen, B. A., and Eriksen, C. W. (1974). Effects of noise letters upon the identification of target letters in a non-search task. *Percept. Psychophys.* 16, 143–149.

Eriksen, C. W., Coles, M. G. H., Morris, L. R., and O'Hara, W. P. (1985). An electromyographic examination of response competition. *Bull. Psychon. Soc.* 23, 165–168.

Forstmann, B. U., Jahfari, S., Scholte, H. S., Wolfensteller, U., van den Wildenberg, W. P. M., and Ridderinkhof, K. R. (2008a). Function and structure of the right inferior frontal cortex predict individual differences in response inhibition: a model-based approach. *J. Neurosci.* 28, 9790–9796.

Forstmann, B. U., van den Wildenberg, W. P. M., and Ridderinkhof, K. R. (2008b). Neural mechanisms, temporal dynamics, and individual differences in interference control. *J. Cogn. Neurosci.* 20, 1854–1865.

Frank, M. J. (2006). Hold your horses: a dynamic computational role for the subthalamic nucleus in decision making. *Neural Netw.* 19, 1120–1136.

Hasbroucq, T., Burle, B., Akamatsu, M., Vidal, and F., and Possamaï, C. A. (2001). An electromyographic investigation of the effect of stimulus–response mapping on choice reaction time. *Psychophysiology* 38, 157–162.

Hasbroucq, T., Burle, B., Vidal, F., and Possamaï, C. A. (2009). Stimulus-hand correspondence and direct response activation: an electromyographic analysis. *Psychophysiology* 46, 1160–1169.

Hasbroucq, T., Possamaï, C. A., Bonnet, M., and Vidal, F. (1999). Effect of the irrelevant location of the response signal on choice reaction time: an electromyographic study in humans. *Psychophysiology* 36, 522–526.

Hazeltine, E., Bunge, S. A., Scanlon, M. D., and Gabrieli, J. D. E. (2003). Material-dependent and material-independent selection processes in the frontal and parietal lobes: an event-related fMRI investigation of response competition. *Neuropsychologia* 41, 1208–1217.

Hommel, B. (1994). Spontaneous decay of response code activation. *Psychol. Res.* 56, 261–268.

Juncos-Rabadán, O. J., Pereiro, A. X., and Facal, D. (2008). Cognitive interference and aging: insight from a spatial stimulus–response consistency task. *Acta Psychol.* 127, 237–246.

Kastner, S., and Ungerleider, L. G. (2000). Mechanisms of visual attention in the human cortex. *Annu. Rev. Neurosci.* 23, 315–341.

Kerns, J. G., Cohen, J. D., MacDonald, A. W. III, Cho, R. Y., Stenger, V. A., and Carter, C. S. (2004). Anterior cingulate conflict monitoring and adjustments in control. *Science* 303, 1023–1026.

Kornblum, S. (1994). The way irrelevant dimensions are processed depends

on what they overlap with: the case of Stroop- and Simon-like stimuli. *Psychol. Res.* 56, 130–135.

Kornblum, S., Hasbroucq, T., and Osman, A. (1990). Dimensional overlap: cognitive basis for stimulus–response compatibility: a model and taxonomy. *Psychol. Rev.* 97, 253–270.

Liu, X., Banich, M. T., Jacobson, B. L., and Tanabe, J. L. (2004). Common and distinct neural substrates of attentional control in an integrated Simon and spatial Stroop task as assessed by event-related fMRI. *Neuroimage* 22, 1097–1106.

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

Mars, R. B., Klein, M. C., Neubert, F. X., Olivier, E., Buch, E. R., Boorman, E. D., and Rushworth, M. F. (2009). Short-latency influence of medial frontal cortex on primary motor cortex during action selection under conflict. *J. Neurosci.* 29, 6926–6931.

Masaki, H., Falkenstein, M., Stürmer, B., Pinkpank, T., and Sommer, W. (2007). Does the error negativity reflect response conflict strength? evidence from a Simon task. *Psychophysiology* 44, 579–585.

Mink, J. (1996). The basal ganglia: focused selection and inhibition of competing motor programs. *Prog. Neurobiol.* 50, 381–425.

Nee, D. E., Wager, T. D., and Jonides, J. (2007). Interference resolution: insights from a meta-analysis of neuroimaging tasks. *Cogn. Affect. Behav. Neurosci.* 7, 1–17.

Neubert, F., Mars, R. B., Buch, E. R., Olivier, E., and Rushworth, M. F. S. (2010). Cortical and subcortical interactions during action reprogramming and their related white matter pathways. *Proc. Natl. Acad. Sci. U.S.A.* 107, 13240–13245.

Nigg, J. R. (2001). Is AD/HD a disinhibitory disorder? *Psychol. Bull.* 127, 571–598.

Pardo, J. V., Pardo, P. J., Janer, K. W., and Raichle, M. E. (1990). The anterior cingulate cortex mediates processing selection in the Stroop attentional conflict paradigm. *Proc. Natl. Acad. Sci. U.S.A.* 87, 256–259.

Peterson, B. S., Kane, M. J., Alexander, G. M., Lacadie, C., Skudlarski, P., Leung, H. C., May, J., and Gore, J. C. (2002). An event-related functional MRI study comparing interference effects in the Simon and Stroop tasks. *Cogn. Brain Res.* 13, 427–240.

Ridderinkhof, K. R. (2002a). "Activation and suppression in conflict tasks: empirical clarification through distributional analyses," in *Attention and Performance*, Vol. XIX, *Common Mechanisms in Perception and Action*, eds W. Prinz and B. Hommel (Oxford: Oxford University Press), 494–519.

Ridderinkhof, K. R. (2002b). Micro- and macro-adjustments of task set: activation and suppression in conflict tasks. *Psychol. Res.* 66, 312–323.

Ridderinkhof, K. R., Forstmann, B. U., Wylie, S. A., Burle, B., and van den Wildenberg, W. P. M. (2010). Neurocognitive mechanisms of action control: resisting the call of the sirens. *Wiley Interdiscip. Rev. Cogn. Science.* doi: 10.1002/wcs.99.

Ridderinkhof, K. R., Scheres, A., Oosterlaan, J., and Sergeant, J. (2005). Distribution-analytical techniques in the study of AD/HD: delta plot analyses reveal deficits in response suppression that are eliminated by methylphenidate treatment. *J. Abnorm. Psychol.* 114, 197–215.

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

Ridderinkhof, K. R., van den Wildenberg, W. P. M., Segalowitz, S. J., and Carter, C. S. (2004b). Neurocognitive mechanisms of cognitive control: the role of prefrontal cortex in action selection, response inhibition, performance monitoring, and reward-based learning. *Brain Cogn.* 56, 129–140.

Ridderinkhof, K. R., van den Wildenberg, W. P. M., Wijnen, J., and Burle, B. (2004c). "Response inhibition in conflict tasks is revealed in delta plots," in *Cognitive Neuroscience of Attention*, ed. M. Posner (New York: Guilford Press), 369–377.

Roger, C., Bénar, C. G., Vidal, F., Hasbroucq, T., and Burle, B. (2010). Rostral cingulate zone and correct response monitoring: ICA and source localisation evidences for the unicity of correct- and error-negativities. *Neuroimage* 51, 391–403.

Scheffers, M. K., Coles, M. G. H., Bernstein, P., Gehring, W. J., and Donchin, E. (1996). Event-related brain potentials and error-related processing: an analysis of incorrect responses to go and no-go stimuli. *Psychophysiology* 33, 42–53.

Schumacher, E. H., Cole, M. W., and D'Esposito, M. (2007). Selection and maintenance of stimulus–response rules during preparation and performance of a spatial choice-reaction task. *Brain Res.* 1136, 77–87.

Schumacher, E. H., Elston, P. A., and D'Esposito, M. (2003). Neural evidence for representation-specific response selection. *J. Cogn. Neurosci.* 15, 1111–1121.

Simon, J. R. (1967). Ear preference in a simple reaction-time task. *J. Exp. Psychol.* 75, 49–55.

Simon, J. R. (1990). "The effect of an irrelevant directional cue on human information processing," in *Stimulus–Response Compatibility: An Integrated Perspective*, eds R. W. Proctor and T. G. Reeve (Amsterdam: North-Holland), 31–88.

Smid, H. G. O. M., Mulder, G., and Mulder, L. J. M. (1990). Selective response activation can begin before stimulus recognition is complete: a psychophysiological and error analysis of the continuous flow. *Acta Psychol.* 74, 169–201.

Speckman, P. L., Rouder, J. N., Morey, R. D., and Pratte, M. S. (2008). Delta plots and coherent distribution ordering. *Am. Stat.* 62, 262–266.

Stins, J. F., Polderman, J. C., Boomsma, D. I., and de Geus, E. J. C. (2007). Conditional accuracy in response interference tasks: evidence from the Eriksen flanker task and the spatial conflict task. *Adv. Cogn. Psychol.* 3, 389–396.

Stroop, J. (1935). Studies of interference in serial verbal reactions. *J. Exp. Psychol.* 18, 643–662.

Stürmer, B., Leuthold, H., Soetens, E., Schröter, H., and Sommer, W. (2002). Control over location-based response activation in the Simon task: behavioral and electrophysiological evidence. *J. Exp. Psychol. Hum. Percept. Perform.* 28, 1345–1363.

van den Wildenberg, W. P. M., Burle, B., Vidal, F., van der Molen, M. W., Ridderinkhof, K. R., and Hasbroucq, T. (2010). Mechanisms and dynamics of cortical motor enhibition in the stop-signal paradigm: a TMS study. *J. Cogn. Neurosci.* 22, 225–239.

van den Wildenberg, W. P. M., van Boxtel, G. J. M., van der Molen, M. W., Bosch, D. A., Speelman, J. D., and Brunia, C. H. M. (2006). Stimulation of the subthalamic region facilitates the selection and inhibition of motor responses in Parkinson's disease. *J. Cogn. Neurosci.* 18, 626–636.

Verbruggen, F., Aron, A. R., Stevens, M. A., and Chambers, C. D. (2010). Theta burst stimulation dissociates attention and action updating in human inferior frontal cortex. *Proc. Natl. Acad. Sci. U.S.A.* 107, 13966–13971.

Vidal, F., Hasbroucq, T., Grapperon, J., and Bonnet, M. (2000). Is the error negativity specific to errors? *Biol. Psychol.* 51, 109–128.

Wagenmakers, E.-J., Grasman, R. P. P. P., and Molenaar, P. C. M. (2005). On the relation between the mean and the variance of a diffusion model response time distribution. *J. Math. Psychol.* 49, 195–204.

Wiegand, K., and Wascher, E. (2005). Dynamic aspects of S–R correspondence: evidence for two mechanisms involved in the Simon effect. *J. Exp. Psychol. Hum. Percept. Perform.* 31, 453–464.

Wiegand, K., and Wascher, E. (2007). Response coding in the Simon task. *Psychol. Res.* 71, 219–233.

Wijnen, J. G., and Ridderinkhof, K. R. (2007). Response inhibition in motor and oculomotor conflict tasks: different mechanisms, different dynamics? *Brain Cogn.* 63, 260–270.

Wylie, S. A., Ridderinkhof, K. R., Bashore, T. R., and van den Wildenberg, W. P. M. (2010). The effect of Parkinson's disease on the dynamics of online and proactive cognitive control during action selection. *J. Cogn. Neurosci.* 22, 2058–2073.

Wylie, S. A., Ridderinkhof, K. R., Eckerle, M. K., and Manning, C. A. (2007). Inefficient response inhibition in individuals with mild cognitive impairment. *Neuropsychologia* 45, 1408–1419.

Wylie, S. A., Ridderinkhof, K. R., Elias, W. J., Frysinger, R. C., Bashore, T. R., Downs, K. E., van Wouwe, N. C., and van den Wildenberg, W. P. M. (2010). Subthalamic nucleus stimulation influences expression and suppression of impulsive behavior in Parkinson's disease. *Brain*. doi: 10.1093/brain/awq239

Wylie, S. A., van den Wildenberg, W. P. M, Ridderinkhof, K. R., Bashore, T. R., Powell, V. D., Manning, C. A., and Wooten, G. F. (2009a). The effect of Parkinson's disease on interference control during action selection. *Neuropsychologia* 47, 145–157.

Wylie, S. A., van den Wildenberg, W. P. M, Ridderinkhof, K. R., Bashore, T. R., Powell, V. D., Manning, C. A., and Wooten, G. F. (2009b). The effect of speed-accuracy strategy on response interference control in Parkinson's disease. *Neuropsychologia* 47, 1844–1853.

**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 September 2010; accepted: 16 November 2010; published online: 13 December 2010.*

*Citation: van den Wildenberg WPM, Wylie SA, Forstmann BU, Burle B, Hasbroucq T and Ridderinkhof KR (2010) To head or to heed? Beyond the surface of selective action inhibition: a review. Front. Hum. Neurosci. **4**:222. doi: 10.3389/fnhum.2010.00222*

frontiers in
HUMAN NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 29 December 2010
doi: 10.3389/fnhum.2010.00232

# Deconstructing the "resting" state: exploring the temporal dynamics of frontal alpha asymmetry as an endophenotype for depression

***John J. B. Allen*[1]* and *Michael X Cohen*[2,3]**

[1] *Department of Psychology, University of Arizona, Tucson, AZ, USA*
[2] *Department of Psychology, University of Amsterdam, Amsterdam, Netherlands*
[3] *Department of Physiology, University of Arizona, Tucson, AZ, USA*

***Edited by:***
*Judith M. Ford, Yale University School of Medicine, USA*

***Reviewed by:***
*Brian F. O'Donnell, Indiana University, USA*
*Ryan Thibodeau, St. John Fisher College, USA*

***Correspondence:**
*John J. B. Allen, Department of Psychology, University of Arizona, P.O. Box 210068, Tucson, AZ 85721-0068, USA.*
*e-mail: jallen@u.arizona.edu*

Asymmetry in frontal electrocortical alpha-band (8–13 Hz) activity recorded during resting situations (i.e., in absence of a specific task) has been investigated in relation to emotion and depression for over 30 years. This asymmetry reflects an aspect of endogenous cortical dynamics that is stable over repeated measurements and that may serve as an endophenotype for mood or other psychiatric disorders. In nearly all of this research, EEG activity is averaged across several minutes, obscuring transient dynamics that unfold on the scale of milliseconds to seconds. Such dynamic states may ultimately have greater value in linking brain activity to surface EEG asymmetry, thus improving its status as an endophenotype for depression. Here we introduce novel metrics for characterizing frontal alpha asymmetry that provide a more in-depth neurodynamical understanding of recurrent endogenous cortical processes during the resting-state. The metrics are based on transient "bursts" of asymmetry that occur frequently during the resting-state. In a sample of 306 young adults, 143 with a lifetime diagnosis of major depressive disorder (62 currently symptomatic), three questions were addressed: (1) How do novel peri-burst metrics of dynamic asymmetry compare to conventional fast-Fourier transform-based metrics? (2) Do peri-burst metrics adequately differentiate depressed from non-depressed participants? and, (3) what EEG dynamics surround the asymmetry bursts? Peri-burst metrics correlated with traditional measures of asymmetry, and were sensitive to both current and past episodes of major depression. Moreover, asymmetry bursts were characterized by a transient lateralized alpha suppression that is highly consistent in phase across bursts, and a concurrent contralateral transient alpha enhancement that is less tightly phase-locked across bursts. This approach opens new possibilities for investigating rapid cortical dynamics during resting-state EEG.

**Keywords: EEG asymmetry, resting-state, EEG dynamics, endophenotype, depression**

## INTRODUCTION

Prefrontal brain asymmetry has been proposed to serve as an indicator of risk for major depressive disorder (MDD; Allen et al., 2004b; Stewart et al., 2010) as well as internalizing disorders more generally (e.g., Allen et al., 1993; Bruder et al., 1997; Nitschke et al., 1999; Wiedemann et al., 1999; Davidson et al., 2000; Accortt and Allen, 2006; Mathersul et al., 2008; Stewart et al., 2008), and individuals low on this risk factor demonstrate significant signs of well-being (e.g., Kang et al., 1991; Schmidt and Fox, 1994; Fox et al., 1995; Davidson et al., 1999; Urry et al., 2004). Such findings support the proposition that prefrontal brain asymmetry indexes a dimension that spans risk for or resilience from depression and internalizing psychopathology, and that it provides an easily measured endophenotype (Iacono, 1998; Gottesman and Gould, 2003), a measurable endogenous characteristic of an individual that is related to underlying mechanisms conferring risk. Prefrontal brain asymmetry demonstrates several characteristics to support its promise as an endophenotype for depression(for overviews, see Allen et al., 2004b; Coan and Allen, 2004; Stewart et al., 2010) including that: (1) it differentiates depressed from non-depressed individuals not only during active episodes but when in remission (e.g., Henriques and Davidson, 1990, 1991; Allen et al., 1993, 2004b; Gotlib et al., 1998); (2) it is heritable (Anokhin et al., 2006; Smit et al., 2007); and (3) it is seen at elevated rates among non-depressed relatives of those with a history of depression (e.g., Dawson et al., 1997, 1999; Jones et al., 1997; Tomarken et al., 2004; Diego et al., 2010).

Prefrontal brain asymmetry in these contexts, however, has been measured almost exclusively with assessments of resting EEG alpha activity (Allen et al., 2004a; Coan and Allen, 2004), deriving summary metrics that may reflect anywhere from 2 to 12 min of resting brain electrical activity. Such resting assessments are often viewed from the somewhat endemic model that resting frontal EEG asymmetry taps trait-like dispositions for affective responses (Davidson, 1998b) of approach (indexed by relatively greater left frontal activity) or withdrawal (indexed by relatively greater right frontal activity; Coan et al., 2006). Although there is now a substantial literature of more than 80 studies to suggest that resting frontal EEG activity may serve as an indicator of a trait-like diathesis to respond to emotional situations with a characteristic pattern of emotional negativity and behavioral withdrawal (for review, see

Coan and Allen, 2004), this literature is not without studies that fail to support this notion (e.g., Bruder et al., 1997; Reid et al., 1998; Debener et al., 2000; Pizzagalli et al., 2002).

Such inconsistencies have inspired a healthy debate in this literature about the psychometric properties of the metrics of frontal asymmetry and the importance of various EEG reference montages for assessing prefrontal brain asymmetry (e.g., Davidson, 1998a; Hagemann et al., 2001, 2002; Allen et al., 2004a,b; Coan and Allen, 2004; Towers and Allen, 2009). Although resting frontal brain asymmetry shows high internal consistency (Towers and Allen, 2009) and reasonable stability over time (Hagemann et al., 2002; Allen et al., 2004b), there exists considerable variability within the resting assessment period, with non-depressed individuals demonstrating relatively greater left frontal activity approximately 70% of the time, and relatively greater right activity the remainder (Baehr et al., 1998; Allen et al., 2001). Such findings suggest that there may indeed be meaningful variability within the resting assessment period that deserves greater scrutiny, and that the assumption that resting EEG asymmetry is a trait may need to be reframed to reflect a recurring series of states within the resting period. Thus, investigations of this variability would be motivated by the assumption that the resting-state does not reflect a homogeneous psychological or physiological condition, but instead comprises multiple dynamic states, some of which may predominate sufficiently to produce meaningful, albeit suboptimal, summary metrics over the entire resting period. Further, some of these dynamic states may have greater relevance to questions of interest, such as identifying individuals at greatest risk for depression. Thus a finer-grained analysis of the resting-state may ultimately improve the status of prefrontal brain asymmetry as an endophenotype for depression. This assumption of dynamic networks active over the resting interval has a profitable precedent, as this assumption is the basis for analysis of resting fMRI data, which led to the characterization of the default mode network (Raichle et al., 2001; Smith et al., 2009).

Additionally, a finer-grained analysis of the resting-state may aid in linking the currently coarse and highly global metrics of prefrontal brain asymmetry to more specific aspects of neural function. Although resting frontal EEG asymmetry has been linked repeatedly to psychological traits and thus has reasonable psychological construct validity (Allen and Kline, 2004; Cacioppo, 2004), there is a conspicuous absence of links to functional or structural brain imaging data or brain neurochemistry, thus revealing that resting frontal EEG asymmetry has impoverished neurophysiological construct validity (Allen and Kline, 2004; Davidson, 2004). Because the measure of frontal EEG asymmetry suffers from both relatively crude temporal and also spatial resolution, a finer-grained analysis of temporal dynamics of the resting-state EEG, utilizing a spatially more specific approach, has promise in linking resting frontal EEG asymmetry to other aspects of neurophysiology. Addressing the spatial specificity, a recent investigation reduced the contribution of distal volume-conducted signals to frontal cortex via the use of the current-source-density (CSD) transformation (Stewart et al., 2010). In that study, the CSD derivation yielded a more robust relationship to a history of depression than did other traditional reference montages (i.e., average reference, averaged mastoids reference, Cz reference); CSD-based frontal EEG asymmetry differentiated individuals with any history of depression – current or previous episodes – from never-depressed individuals, independent of the current level of depressive symptoms. Still lacking from that study, however, was a fine-grained temporal analysis of frontal asymmetry, as the measures reflected the average activity across several minutes of resting sessions.

In the present investigation, therefore, we sought to improve both the spatial and temporal precision of resting frontal EEG asymmetry, characterizing this asymmetry in terms of a set of features that recurred dynamically during the resting period, and leveraging the improved anatomical specificity and sensitivity to depression history provided by the use of the CSD reference. By comparing metrics based on these recurring features to the conventional global metrics of resting EEG asymmetry, it was possible to assess whether the novel metrics retain the construct validity of the extant global metrics in terms of differentiating depressed and non-depressed individuals. Should these novel metrics that reflect the dynamics of the resting period show adequate correspondence to extant metrics and similarly differentiate depressed from non-depressed individuals, they may hold advantages for future investigations that can explore brain state dynamics and neural bases of activity that portends risk for depression.

## MATERIALS AND METHODS

### PARTICIPANTS

Participants were 306 young adults also reported in a recent report (Stewart et al., 2010). Prospective participants were identified using Beck depression inventory (BDI; Beck et al., 1961) scores, and were initially telephoned by a post-baccalaureate project manager to determine if they met exclusionary criteria: left handedness, history of head injury with loss of consciousness >10 min, concussion, epilepsy, electroshock therapy, use of current psychotropic medications, and active suicidal potential necessitating immediate treatment. Individuals receiving current psychotherapy were not excluded. Participants not excluded by this screen were invited for an intake interview. All participants accepted into the study were required to be strongly right handed (a score of greater than 35 on the 39 point scale of Chapman and Chapman, 1987). **Figure 1** provides a detailed flow chart summarizing study recruitment across a 4-year period. Participants were screened during the intake interview for Axis I psychopathology using the structured clinical interview for DSM-IV (SCID, First et al., 2002), and excluded if they met criteria for any current comorbid DSM-IV Axis I disorder other than lifetime MDD or current dysthymia. Inter-rater reliability was high for both current and past MDD diagnoses ($\kappa = 0.81$ and 0.91, respectively). The final sample of 306 participants (95 male) ranged in age from 17 to 34 years ($M = 19.1$, SE = 0.1), of whom 143 met criteria for the lifetime MDD. All participants provided informed consent, using procedures and a form approved by the Institutional Review Board of the University of Arizona.

**Figure 1** summarizes the flow of recruitment, **Table 1** provides participant demographics and symptom characterization, and **Table 2** provides diagnostic information about those with lifetime MDD. Depression severity was assessed during the intake interview with the BDI II (Beck et al., 1996) and the 17-item Hamilton rat-

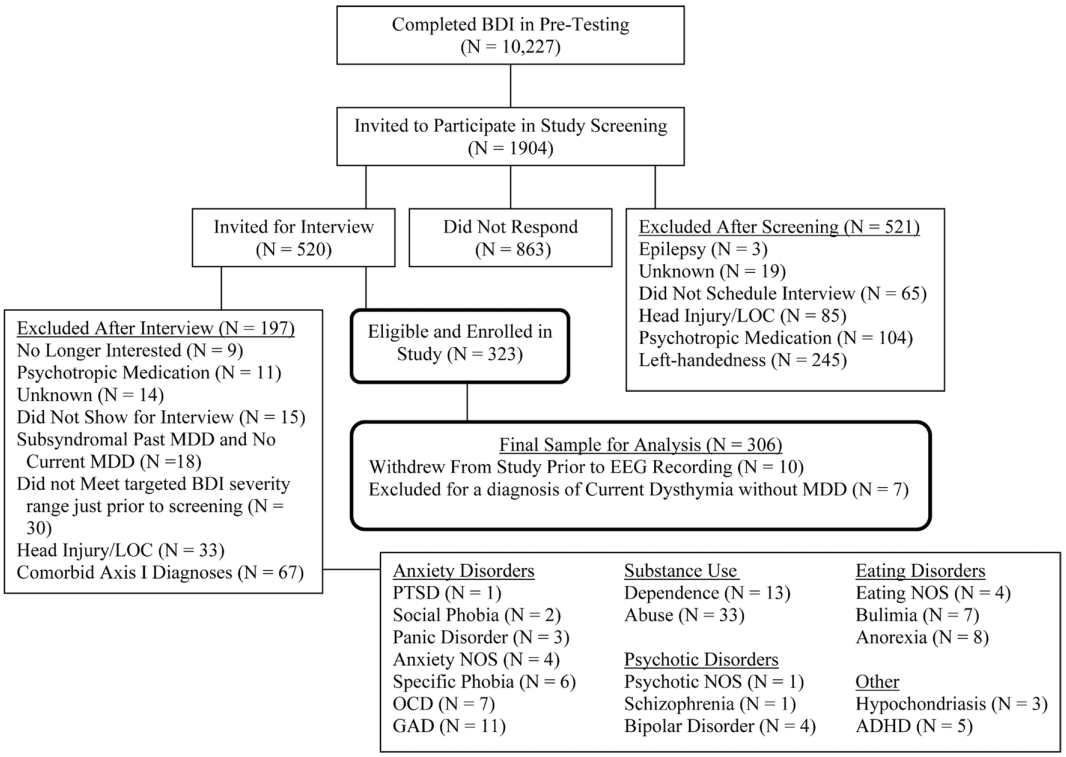



**FIGURE 1 | Flowchart of participant screening and enrollment.** Note: BDI, Beck depression inventory; LOC, loss of consciousness; MDD, major depressive disorder; PTSD, posttraumatic stress disorder; NOS, not otherwise specified; OCD, obsessive compulsive disorder; GAD, generalized anxiety disorder; ADHD, attention deficit hyperactivity disorder. After Stewart et al. (2010).

**Table 1 | Participant demographics.**

| MDD status | Biological sex | Caucasian, % | BDI-II, *M* (SE) | HRSD, *M* (SE) |
|---|---|---|---|---|
| Lifetime MDD+ | Men (*N* = 39) | 71.8 | 15.2 (1.3) | 9.6 (0.9) |
| | Women (*N* = 104) | 77.9 | 17.6 (0.8) | 11.7 (0.5) |
| Lifetime MDD– | Men (*N* = 56) | 66.1 | 5.7 (1.1) | 3.9 (0.7) |
| | Women (*N* = 107) | 71.0 | 6.2 (0.8) | 4.0 (0.5) |

*MDD, major depressive disorder; BDI-II, Beck depression inventory II; HRSD, Hamilton rating scale for depression.*

**Table 2 | DSM-IV diagnoses endorsed in lifetime MDD+ group (*N* = 143).**

| Diagnosis | Male | Female |
|---|---|---|
| Current MDD only | 5 | 9 |
| Past MDD only | 20 | 55 |
| Current MDD and past MDD | 10 | 29 |
| Current MDD and current dysthymia | 0 | 2 |
| Past MDD and current dysthymia | 1 | 5 |
| Current MDD, Past MDD, and current dysthymia | 3 | 4 |

ing scale for depression (HRSD; Hamilton, 1960), the latter that had an intra-class correlation of inter-rater agreement of 0.95 for a randomly selected sample of 10% of HRSD interviews.

### EEG DATA COLLECTION AND REDUCTION

Participants visited the laboratory on four separate days within a 2-week period. Two resting EEG sessions were completed each day. Nine participants attended fewer than all four EEG assessment days but all subjects were included in mixed linear model analyses that successfully handle missing data (Bagiella et al., 2000). During each resting EEG session, eight 1 min baselines were recorded, with eyes-open (O) or eyes-closed (C), in one of two counterbalanced orders (OCCOCOOC or COOCOCCO). Sessions within day were separated by approximately 20 min. All EEG data were acquired using a 64-channel NeuroScan Synamps2 amplifier, and imported into Matlab using the EEGLab Toolbox (Delorme and Makeig, 2004). A vertical electrooculogram (EOG) channel (superior and inferior orbit of the left eye) was recorded for later ocular artifact rejection. All impedances were kept under 10 kΩ. Data for each resting session were digitized continuously at 1000 Hz, amplified 2816 times, and filtered with 200 Hz low pass filter prior to digitization. EEG data were acquired with an online reference site immediately posterior to Cz and subsequently re-referenced offline to a CSD derivation using the CSD Toolbox of Kayser and Tenke (2006a,b) that is based on the spherical spline approach summarized by Perrin et al. (1989, 1990). CSD is a spatial filter that minimizes the contributions of deep and/or distant sources. It therefore provides a more accurate depiction of the topography of electrocortical dynamics (Srinivasan et al., 1996, 2007), and is appropriate for localizing time-frequency dynamics of cognitive processes at the level of the scalp (Cavanagh et al.,

2009; Cohen et al., 2009a). Modeling (Winter et al., 2007) and empirical (Srinivasan et al., 2007) papers suggest that volume conduction artifacts extend to less than 3 cm across the scalp (typical distance between two electrodes in a 128-channel cap, or 1/2 of an electrode in a 64-channel cap). It is believed that after CSD/Laplacian processing, each electrode measures activity from ~3 cm underneath the electrode. Thus, it is likely that the effects reported here are generated by tissue close to the surface of the brain. This is not meant to imply that deeper generators of alpha do not exist or are irrelevant; rather, our choice of methods helps topographically localize clinically relevant cortical dynamics.

After acquisition, epochs with movement and muscle artifacts were removed via visual inspection, following which segments were automatically rejected if vertical ocular activity exceeded ±75 μV. Finally, a custom artifact rejection algorithm rejected segments with large fast deviations in amplitude in any channel (e.g., DC shifts and spikes) that may have eluded human inspection. After these preprocessing steps, data were then processed in the standard fashion to derive the customary metrics of asymmetry (details below), and also processed using a novel approach that is based on identifying spontaneous asymmetrical bursts in the ongoing EEG (details below).

### *Conventional Data Reduction*

The standard data processing steps, and the rationale behind them, have been described in detail elsewhere (Allen et al., 2004a; Stewart et al., 2010). For this approach, summarized in **Figure 2**, data are segmented into short epochs, tapered with a Hamming window, transformed via fast-Fourier transform (FFT) to power spectra, and then the average across all spectra is derived and total alpha power is extracted for each site. In the present data, each resting session was segmented into 1 min EEG blocks and further into 117 epochs of 2.048 s per block, overlapping by 1.5 s. Overlapping epochs offsets the minimal weight applied to the end of the epoch due to the Hamming window function. Next, a power spectrum for each artifact-free epoch was derived via FFT, following which all power spectra across all 8 min were averaged to provide a summary spectrum for each resting session. At each site, from each resting session, total alpha power (8–13 Hz) was extracted and an asymmetry score for each resting session was computed with natural-log transformed scores [i.e., ln(Right) – ln(Left)] for homologous left and right leads (e.g., F7 and F8, F5 and F6, F3 and F4, F1 and F2), with higher scores interpreted to reflect relatively greater left activity (i.e., greater right than left alpha; cf. Allen et al., 2004a). Results from this approach on the present participant

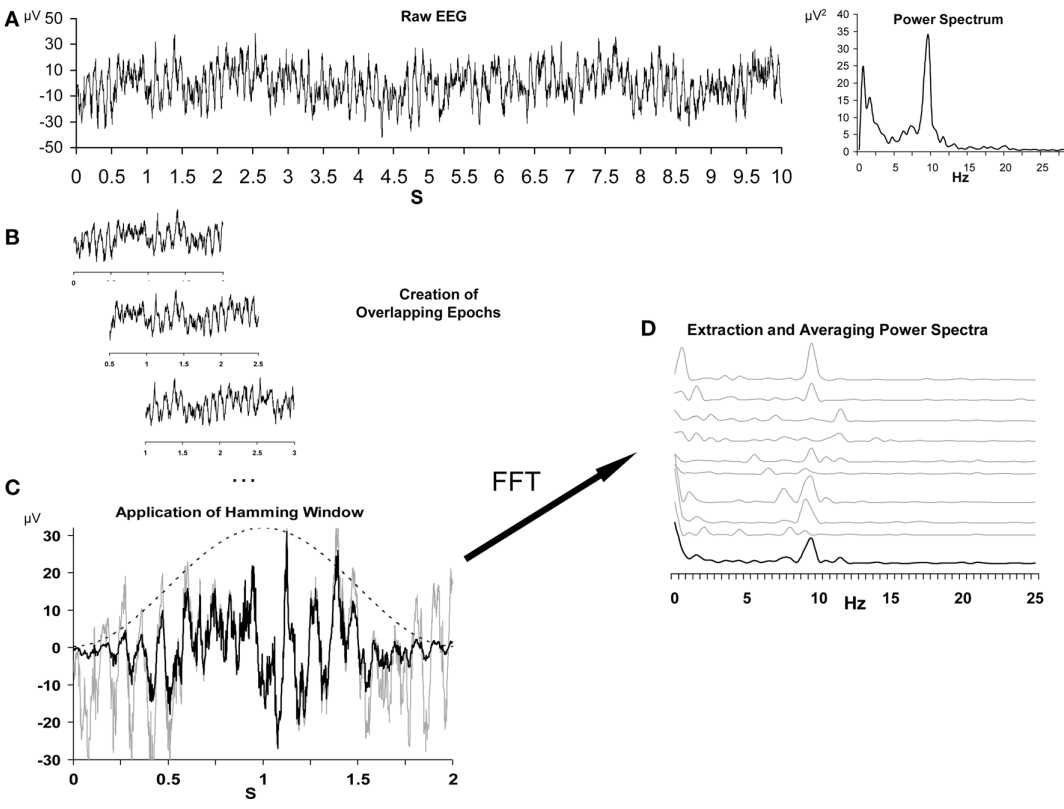



**FIGURE 2 | Schematic depiction of the data reduction steps for the conventional EEG asymmetry data reduction. (A)** depicts a 10-s segment of raw data from a single channel on the left, and the spectral representation of this epoch on the right. **(B)** illustrates the process of epoching the longer segment into shorter overlapping 2 s epochs. **(C)** depicts the impact of the Hamming window (dotted bell curve) on a single epoch, with the gray line representing the raw signal and the black line representing the signal after the application of the window. Panel D illustrates the extraction of individual power spectra via FFT, and finally the averaging of those spectra across epochs. The top nine gray lines are the spectral representation of nine 2 s epochs, and the lower black line is the average spectrum. Note that alpha power (8–13 Hz) is somewhat variable from epoch to epoch, but that the average spectrum reveals a distinct alpha peak. Vertical axis in **(D)** is power in microvolts-squared. Figure modeled after Allen et al. (2004a).

sample have been previously reported (Stewart et al., 2010), and a subset of those findings will be repeated here for comparison to the novel metrics.

### *Novel data reduction based on endogenous asymmetry bursts*

In cognitive/perceptual studies, transient as well as sustained alpha dynamics are thought to be related to "pulsed inhibition," a potential mechanism by which neural processing can be temporally organized and cognitive dynamics can be shaped (Klimesch et al., 2007; Palva and Palva, 2007; Handel et al., 2010). Many of these studies focus on transient (100 s of milliseconds) stimulus-induced alpha activity. On the other hand, the study of frontal alpha asymmetry as it relates to depression has focused on global measures of EEG recorded over several minutes (see previous section). Thus, we adopted methods from cognitive neuroscience to examine whether frontal alpha asymmetry fluctuates dynamically over time. Indeed, during exploratory data investigations, we noticed that alpha asymmetry is not stable over time, but rather is characterized by fluctuations at multiple time scales, and punctuated by "bursts" of increased or decreased power. An example of this can be seen in **Figure 3**. The reasoning behind the analyses reported here is that these bursts of frontal alpha asymmetry might reflect a phasic endogenous process related to the global alpha asymmetry typically reported. We consider each of these bursts to be an event of interest, and conducted time-frequency decomposition analyses surrounding each of these events to investigate these dynamic alpha asymmetries and link them to depression.

To identify endogenous lateralized alpha bursts (**Figure 3**), the ongoing EEG signal at sites F5 and F6 were filtered with an 8–13 Hz FIR bandpass filter. Next the Hilbert transform was applied, which provides the analytic representation of the data, from which estimates of instantaneous power (the natural log of the squared magnitude of the analytic signal) were taken for each site. Finally, the difference between sites F6 and F5 (F6–F5; right–left hemisphere) was taken as a continuous time-varying signal of alpha asymmetry. Sites F6 and F5 were selected as they were centrally positioned among sites examined with the conventional FFT-based metrics (Stewart et al., 2010).

To identify asymmetry bursts, we examined the distribution of the absolute value of alpha asymmetry values over time, and selected time points that were on the extreme upper 1% tail of the distribution for each recording session for each participant[1]. This gives both "positive" (greater relative right alpha power) and "negative" (greater relative left alpha power) bursts. Finally, if, after

[1]A subsequent re-analysis using a 5% cutoff yielded nearly identical results.

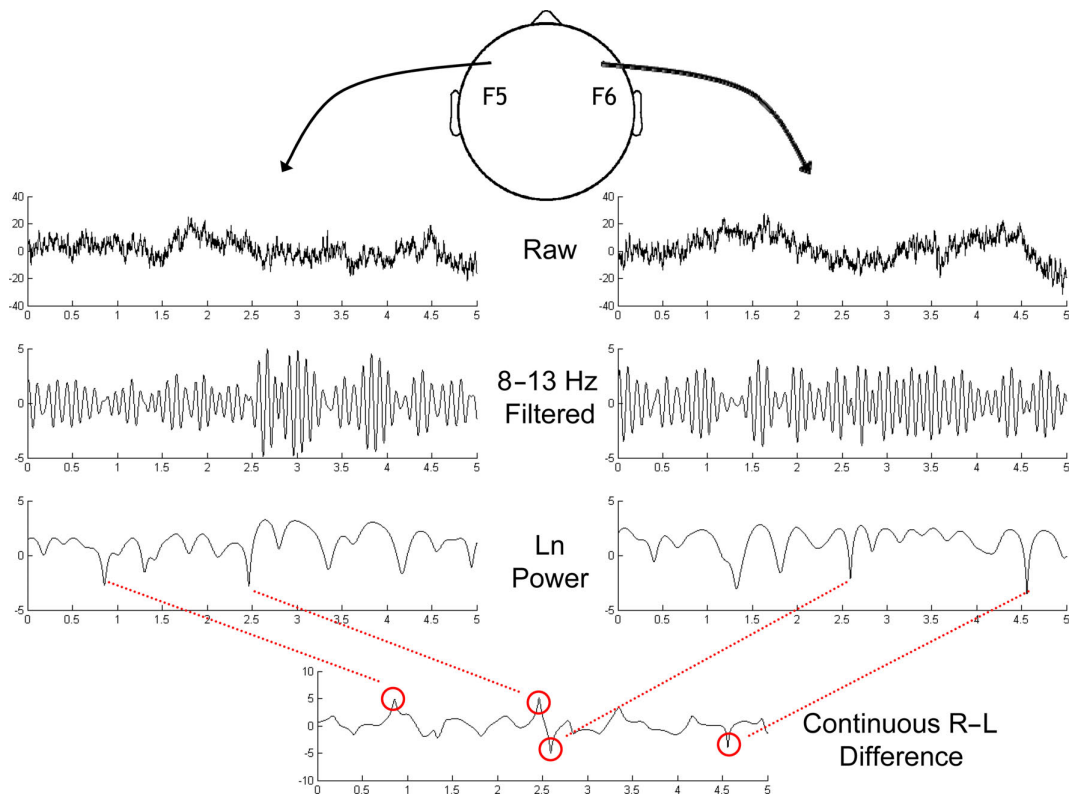



**FIGURE 3 | Schematic depiction of the identification of asymmetrical bursts in the ongoing resting EEG session.** Sites F5 and F6 were chosen based on previous findings as the sites of interest. Top row depicts 5 s of data after current-source density transformation. Next row is alpha-band (8–13 Hz) filtered version of the same data. The natural log of squared Hilbert-transformed alpha-band-pass filtered signals is depicted in the third row. Finally, the subtraction (Right–Left) of these two signals is depicted in the bottom panel, revealing the dynamic nature of frontal alpha asymmetry, with bursts identified by the red circles, and lines linking to the corresponding features in the right or left channel. The leftmost two bursts in this bottom panel are positive bursts, and the rightmost two bursts are negative bursts. Vertical axis for the upper two rows is $\mu V/cm^2$, and for the bottom two rows is log-$\mu V^2/cm^2$.

trial removal/rejection (using the procedure described above), there were fewer than 10 bursts in any one of the four conditions (eyes-open/closed × positive/negative bursts), then the entire session was removed from analyses. Although this might seem overly stringent because there might be valid data in the other conditions, because of the multitude of data from over 300 subjects, we decided to use a cautious approach that would favor selectivity. This procedure resulted in discarding 784 out of 2448 sessions. This does not imply that alpha asymmetry dynamics on different time scales or at less extreme values are irrelevant; rather, we chose to focus this initial investigation on endogenous events that can be unequivocally defined and differentiated.

To quantify the electrophysiological dynamics surrounding these endogenous asymmetry bursts, EEG data from each "trial" (i.e., burst, from 3 s before to 3 s after each burst) were convolved with a family of complex Morlet wavelets, defined as a complex sine wave windowed by a Gaussian: $e^{i2\pi ft} * e^{-t^2/(2*\sigma^2)}$ where $t$ is time, $f$ is frequency, which increased from 2.5 to 50 Hz in 40 logarithmically spaced steps, and $\sigma$ defines the width of each frequency band and is set to $4.5/(2\pi f)$. Estimates of frequency-specific instantaneous power ($\text{real}[z(t)]^2 + \text{imag}[z(t)]^2$) and phase ($\arctan(\text{imag}[z(t)]/\text{real}[z(t)])$) were extracted from the resulting analytic signal. Long epochs were extracted in order to discard data points contaminated by edge artifacts; visual inspection confirmed that no edge artifacts contaminated the data subjected to statistical analyses or shown in figures. Power was averaged across bursts and converted to decibel scale relative to the average power from −500 to +500 ms surrounding the burst; phase values were used for inter-burst phase coherence: $|1/n * \Sigma_{t=1}^{n} e^{i\phi_t}|$. For statistical analyses, data from −26 to +26 ms surrounding bursts were entered into repeated-measures ANOVA and regressions, using the procedures described below[2].

The analyses were designed to address three central questions: (1) How do the novel peri-burst metrics of dynamic asymmetry compare to the conventional FFT-based metrics? (2) Do the peri-burst metrics adequately differentiate depressed and non-depressed participants? and (3) What EEG dynamics surround the asymmetry bursts that are captured by the novel peri-burst metrics? To address these questions, a combination of approaches was utilized. Correlations between the novel peri-burst metrics and conventional asymmetry metrics were examined, and mixed linear models assessed whether the metrics differentiated those with lifetime depression from never-depressed individuals, and further whether these metrics would differentiate never-depressed individuals only from those with current depression (indicative of a state effect) or from those with any lifetime history of depression (indicative of a potential marker of risk).

## RESULTS

### A COMPARISON OF PERI-BURST DYNAMICS WITH CONVENTIONAL FFT-BASED METRICS

As depicted in **Figure 4**, peri-burst alpha power at sites F5 and F6 was most closely related to conventional FFT-derived power from the entire resting period at those same sites. And although the topographical pattern of the correlations of peri-burst alpha power at F5 and F6 with conventional FFT-based power estimates across the scalp was similar for both positive and negative bursts, the magnitude of the correlations were substantially larger for F5 during negative bursts and F6 during positive bursts than for F5 during positive or F6 during negative bursts. Finally, as evident in **Figure 4**, there is a general pattern of positive correlations across the scalp, likely due to the fact that both measures capture individual differences in overall power.

Of greater interest, however, is how the *asymmetry* in peri-burst alpha power relates to conventional and widely used *asymmetry* scores derived via FFT. Depicted in **Figure 5** are correlations between the peri-burst alpha power asymmetry [ln(F6)–ln(F5)] and conventional asymmetry scores [ln(Right)–ln(Left)] derived via FFT from the entire resting period. As seen in **Figure 5**, anatomical specificity is observed in the relationship of peri-burst alpha asymmetry to conventional asymmetry for positive bursts, negative bursts, and the pooled set of bursts, but the relationship is notably stronger for the combined dataset. Combining across both positive and negative bursts produces a peri-burst alpha power asymmetry metric that is most closely aligned with the conventional FFT-based asymmetry scores, accounting for 42% of the variance in conventional asymmetry scores at F6–F5.

Together, these findings demonstrate peri-burst alpha power from a small subset of the recording epoch is remarkably similar to, although not redundant with, conventional FFT-based power.

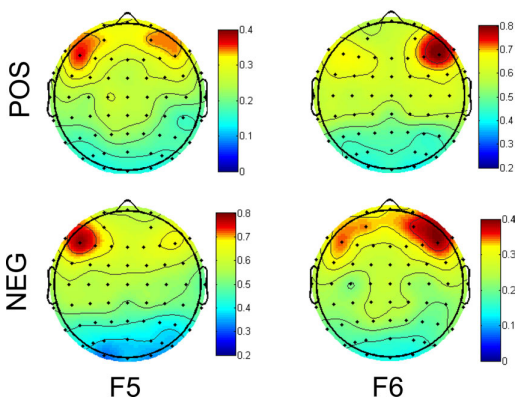



**FIGURE 4 | Correlation of peri-burst alpha power at sites F5 and F6 with the conventional FFT-derived metrics of power over the entire scalp.** Correlations are shown separately for positive bursts (top row) and negative bursts (bottom row). All individual site power values (both conventional and peri-burst power) were natural-log transformed prior to correlation, in keeping with the tradition of log-transformed power values in the EEG asymmetry literature. Against the backdrop of modest positive correlations, reflecting global power differences between subjects, there is anatomical specificity such that peri-burst power from F5 correlates most highly with conventional FFT power at F5, and similarly for these metrics at F6. This anatomical specificity is, in part, due to the effective high-pass spatial filter provided by CSD transformation. Note that the range of correlations differs such that positive bursts for F5 and negative bursts for F6 are on a common scale, and negative bursts for F5 and positive bursts for F6 are on a common scale. Maps were constructed by mapping Pearson correlations using the function topoplot from EEGLab (Delorme and Makeig, 2004).

[2]Matlab scripts to process data using these steps are available from www.psychofizz.org or from the authors. These scripts work with data imported to Matlab using the EEGLab data structure (Delorme and Makeig, 2004).

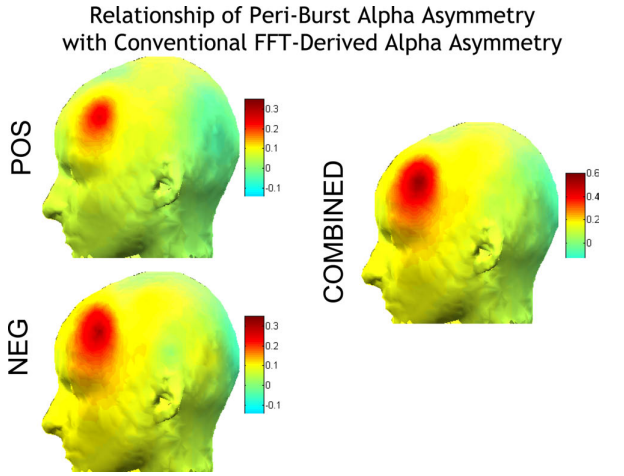



**FIGURE 5 | Correlations of peri-burst alpha asymmetry from sites F5 and F6 [ln(F6)–ln(F5)] with the conventional FFT-derived asymmetry score from all homologous EEG sites [ln(Right)–ln(Left)].** Correlations are shown separately for positive bursts (top left) and negative bursts (bottom left) and combined (right). Because asymmetry scores are difference scores, only one side of the head is depicted, as the opposite side would show identical topography. Anatomical specificity is observed for positive and negative bursts, and all bursts combined, but note that the scale for the combined is larger than either the positive or negative bursts alone, indicating that combining across both positive and negative bursts produces a metric that is most closely aligned with the conventional asymmetry scores, accounting for 42% of the variance in conventional asymmetry scores at F6–F5. Maps were constructed by mapping Pearson correlations using the function headplot from EEGLab (Delorme and Makeig, 2004).

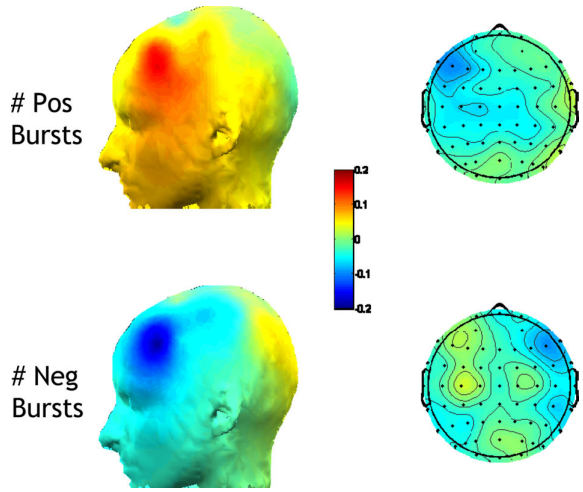



**FIGURE 6 | Correlations of number of positive and negative bursts with the conventional FFT-derived metrics.** Left panel: correlation with the asymmetry score from all homologous EEG sites [ln(Right)–ln(Left)]; Right panel: correlations with ln-transformed alpha power at individual sites. Correlations are shown separately for the number of positive bursts (top) and for the number of negative bursts (bottom). A greater number of positive bursts is associated with relatively greater ln(R)–ln(L) FFT-derived asymmetry scores, and less left frontal alpha. A greater number of negative bursts is associated with relatively lower ln(R)–ln(L) FFT-derived asymmetry scores, and less right frontal alpha.

Finally, the relationship between the number of positive and negative bursts and conventional FFT-derived metrics was examined. Because bursts were identified by selecting the extreme 1% of all activity, the number of positive or negative bursts reflects whether, among an individual's extreme amplitude bursts, the individual had more positive or negative bursts. As shown in **Figure 6**, anatomical specificity is observed in the relationship of number of positive and negative bursts to conventional asymmetry and power at individual sites, but the overall relationships are modest in magnitude.

## EXAMINING WHETHER PERI-BURST DYNAMICS ARE SENSITIVE TO LIFETIME DEPRESSION

To examine whether asymmetry metrics would differ as a function of depression status, mixed linear models were used, first with the conventional FFT-based EEG asymmetry and then with metrics of peri-burst dynamics.

### *Previous findings: conventional FFT-based alpha power asymmetry*

As previously reported (Stewart et al., 2010), conventional FFT-based frontal EEG alpha asymmetry differentiated individuals on the basis of lifetime MDD. A subset of the data (sites F5 and F6) from that report was used here to provide an analysis comparable to those planned for the peri-burst metrics. In the Stewart et al. report, CSD-referenced EEG asymmetry scores that were derived from the conventional FFT method were examined across a broad frontal region (F2–F1, F4–F3, F6–F5, F8–F7). For the model here, the specific electrode-pair F6–F5 was examined, as that is the site where peri-burst metrics were computed. In a model that included lifetime MDD status (past and/or current MDD = lifetime MDD+, never depressed = lifetime MDD–), recording day (4), and session within day (2), the FFT-derived EEG asymmetry score based on total 8–13 Hz alpha power was the dependent variable. Only effects involving MDD status were of interest. In the absence of any significant interactions involving Lifetime MDD status, there was a main effect of Lifetime MDD status [$F(1,2359.9) = 14.0, p < 0.001$], with Lifetime MDD participants having significantly lower frontal asymmetry scores than never-depressed individuals. To determine if these findings were due to currently depressed individuals among the Lifetime MDD group, the model was rerun, but instead of lifetime MDD status, current MDD status was used (current MDD+ = all participants with current MDD, regardless of past MDD status; past MDD+ = participants with past MDD but not current MDD or current dysthymia; MDD– = participants without current or past MDD or dysthymia; six participants with past MDD but current dysthymia were not included in these analyses). Again, there was a main effect of MDD status [$F(2,2307.2) = 5.9$, $p < 0.001$], which indicated that both current and past MDD+ participants had lower frontal EEG asymmetry scores than never-depressed controls (**Figure 7**), and did not differ significantly from one another.

### *Peri-burst alpha power asymmetry*

An identical approach was used to examine peri-burst alpha asymmetry. In a model that included lifetime MDD status (past and/or current MDD = lifetime MDD+, never depressed = lifetime MDD–), recording day (4), and session within day (2), the peri-burst EEG asymmetry score was the dependent variable. Only

effects involving MDD status were of interest. In the absence of any significant interactions involving Lifetime MDD status, there was a main effect of Lifetime MDD status [$F(1,1537.5) = 11.2$, $p < 0.001$], with Lifetime MDD participants having significantly lower frontal asymmetry scores than never-depressed individuals. Following-up the model using current MDD status instead of Lifetime MDD status revealed a main effect of current MDD status [$F(2,1492.8) = 4.9$, $p < 0.01$], which indicated that both current and past MDD+ participants had lower peri-burst frontal EEG asymmetry scores than never-depressed controls (**Figure 8**), and did not differ significantly from one another.

To directly compare the conventional frontal EEG alpha asymmetry to the peri-burst alpha power asymmetry, effect sizes were computed comparing lifetime and current MDD+ groups to never-depressed controls. **Table 3** shows that the effect sizes were roughly comparable for the two approaches; by computing the confidence interval for the effect size of the conventional EEG alpha asymmetry (Hedges and Olkin, 1985), the effect size for peri-burst alpha asymmetry was well within the confidence interval of the conventional alpha power asymmetry for each of the effects sizes in **Table 3**. Thus the novel metrics do not incur any disadvantage in terms of reduced effect size for discriminating depressed from non-depressed participants.

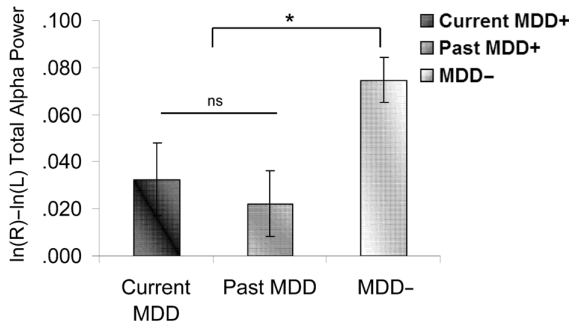



**FIGURE 7 | Conventional frontal alpha asymmetry scores as a function of MDD status.** Error bars reflect standard error. *Y*-axis is ln $\mu V^2/cm^2$. Both currently and previously depressed individuals had significantly lower asymmetry scores than never-depressed individuals.

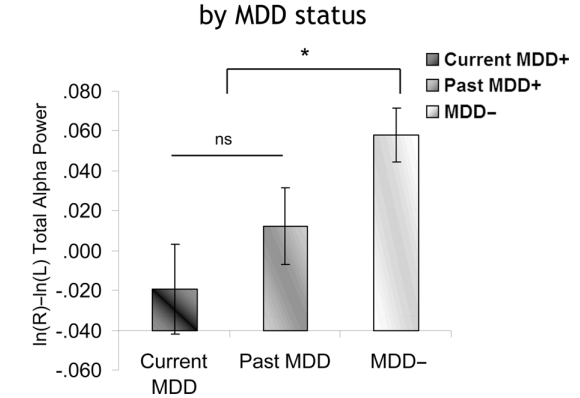



**FIGURE 8 | Peri-burst frontal alpha asymmetry scores as a function of MDD status.** Error bars reflect standard error. *Y*-axis is ln $\mu V^2/cm^2$. As with conventional metrics (**Figure 7**), both currently and previously depressed individuals had significantly lower peri-burst asymmetry scores than never-depressed individuals.

**Table 3 | Effect sizes (Cohen's *d*) comparing depressed groups to never-depressed controls.**

| Diagnosis | Conventional | Peri-burst |
|---|---|---|
| Lifetime MDD | 0.43 | 0.38 |
| Past MDD only | 0.43 | 0.27 |
| Current MDD (with or without Past MDD) | 0.35 | 0.45 |

### *Number of negative and positive bursts*

The number of positive and negative bursts was also examined, in two separate models, one for positive and one for negative bursts. Each model included lifetime MDD status (past and/or current MDD = lifetime MDD+, never depressed = lifetime MDD–), recording day (4), and session within day (2), and the number of bursts was the dependent variable. Only effects involving MDD status were of interest. For negative bursts, in the absence of any significant interactions involving Lifetime MDD status, there was a main effect of Lifetime MDD status [$F(1,1608.2) = 7.7$, $p < 0.01$], with Lifetime MDD participants having significantly more negative bursts than never-depressed individuals. Following-up the model using current MDD status instead of Lifetime MDD status revealed a main effect of current MDD status [$F(2,1558.7) = 4.1$, $p < 0.02$], which indicated that both current and past MDD+ participants had a greater number of negative bursts than never-depressed controls (**Figure 9**), and did not differ significantly from one another. For positive bursts, no significant effects involving either lifetime or current MDD status emerged.

### *Inter-burst phase coherence*

Inter-burst phase coherence was examined in a slightly different model, as an asymmetry score based on phase coherence had no analog in the conventional EEG asymmetry literature. Thus, peri-burst phase coherence values at each site were examined in a model that included lifetime MDD status (past and/or current MDD = lifetime MDD+, never depressed = lifetime MDD–), recording day (4), session within day (2), burst direction (positive/negative) and hemisphere (Left = F5, Right = F6) as factors. Interactions involving Day or Session were not of interest and were not specified in the model in the interest of parsimony. Only effects involving MDD status were of interest and thus reported. In the absence of any significant interactions involving Lifetime MDD status, there was a main effect of Lifetime MDD status [$F(1,11071.7) = 9.8$, $p < 0.001$], with Lifetime MDD participants having significantly higher peri-burst phase coherence than never-depressed individuals. This indicates that the fine timing of alpha activity surrounding alpha bursts is more consistent over repeated bursts in MDD+ individuals compared to that of MDD– individuals.

Following-up the model using current MDD status instead of Lifetime MDD status again revealed a main effect of MDD status [$F(2,10832.0) = 6.1, p < 0.01$], but in this model with current MDD status, this main effect was further qualified by both MDD by burst direction [$F(2,10827.9) = 5.9, p < 0.01$] and MDD by burst direction by hemisphere [$F(2,10827.9) = 3.2, p < 0.05$] interactions. This interaction was decomposed by examining positive and negative bursts separately. For negative bursts, no significant effects involving current MDD status emerged. By contrast, for positive bursts, and as depicted in **Figure 10**, there was a significant main effect of current MDD status [$F(2,5342.2) = 10.3, p < 0.001$] and also a significant Current MDD Status by hemisphere interaction [$F(2,5440.0) = 7.1, p = 0.001$]. Greater inter-burst phase coherence was seen among currently depressed participants compared to previously and never-depressed participants, but only at the left hemisphere site during positive bursts.

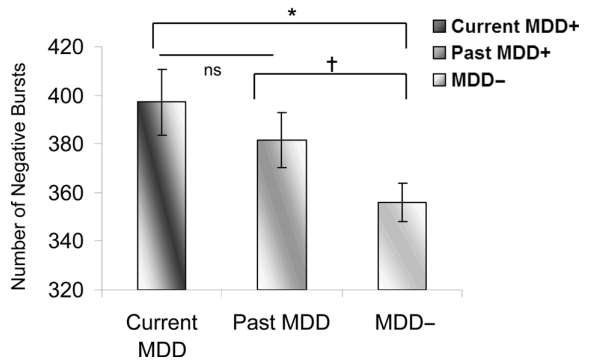



**FIGURE 9 | Number of negative bursts as a function of MDD status.** Error bars reflect standard error. Both currently and previously depressed individuals had significantly more negative bursts than never-depressed individuals. *$p < 0.05$; †$p < 0.07$.

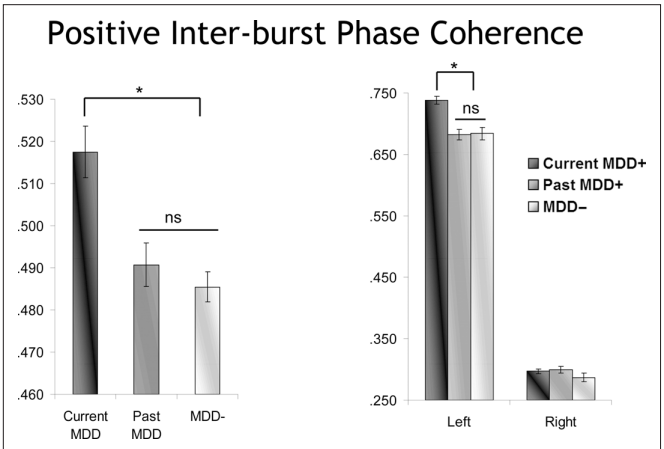



**FIGURE 10 | Inter-burst phase coherence as a function of MDD status for positive bursts.** The effect of MDD status is episode dependent, with only currently depressed individuals showing enhanced inter-burst phase coherence, and only for left frontal activity.

## EEG DYNAMICS SURROUNDING ASYMMETRY BURSTS

The previous sets of analyses demonstrate that transient asymmetry bursts are correlated with conventional measures of asymmetry and are related to current and past depression in similar ways as are conventional measures. Thus, transient asymmetry bursts may potentially open new doors for examining the neural dynamics that underlie frontal asymmetry, while maintaining a link to clinically relevant outcome measures. Our initial investigations focus on time-frequency dynamics in the power (amplitude of signal) and phase (timing of oscillations) of oscillations before and after bursts, and also the topographical distributions of these effects. Although exploratory in nature, the large *N* ensures that these patterns are robust across individuals.

### *Topography of peri-burst alpha power and inter-burst phase coherence*

We first examined the topographical distribution of alpha characteristics in the time frame surrounding bursts. **Figure 11** displays topographical distributions of the peri-burst dynamics for alpha oscillation power and alpha oscillation inter-burst phase coherence (the extent to which alpha oscillations have the same phase value over many bursts), separately for positive and negative bursts. Two features of these topographies are striking: First, each burst is characterized by a topographically focal inverse relationship between alpha power at F5 and F6, the sites that comprise the time-varying asymmetry waveform from which bursts were identified. Second, the bursts, which are defined solely on the basis of relative left-vs.-right hemisphere power, are accompanied by strong inter-burst phase coherence in the hemisphere that shows alpha suppression.

### *Temporal dynamics of peri-burst alpha power and inter-burst phase coherence*

The temporal dynamics of power and inter-burst phase coherence can also be seen in **Figure 12**. These dynamics begin approximately 200 ms (about two alpha cycles) prior to and

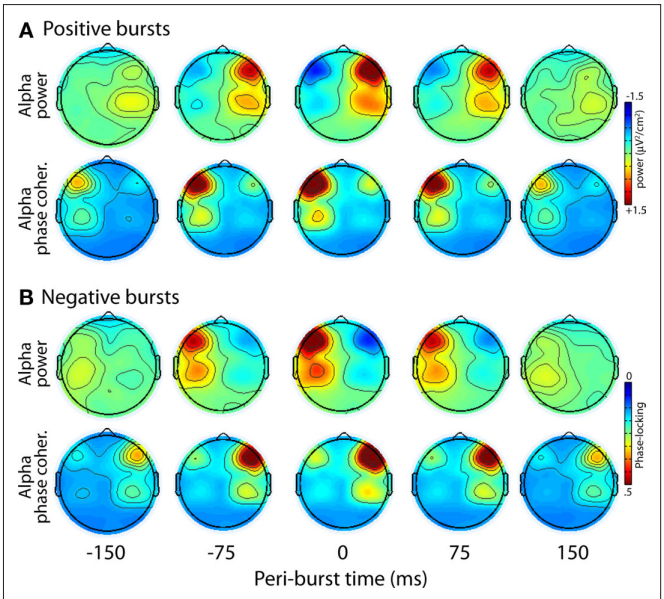



**FIGURE 11 | Topographical distribution of positive (A)** and negative **(B)** bursts for alpha oscillation power and alpha oscillation inter-burst phase coherence.

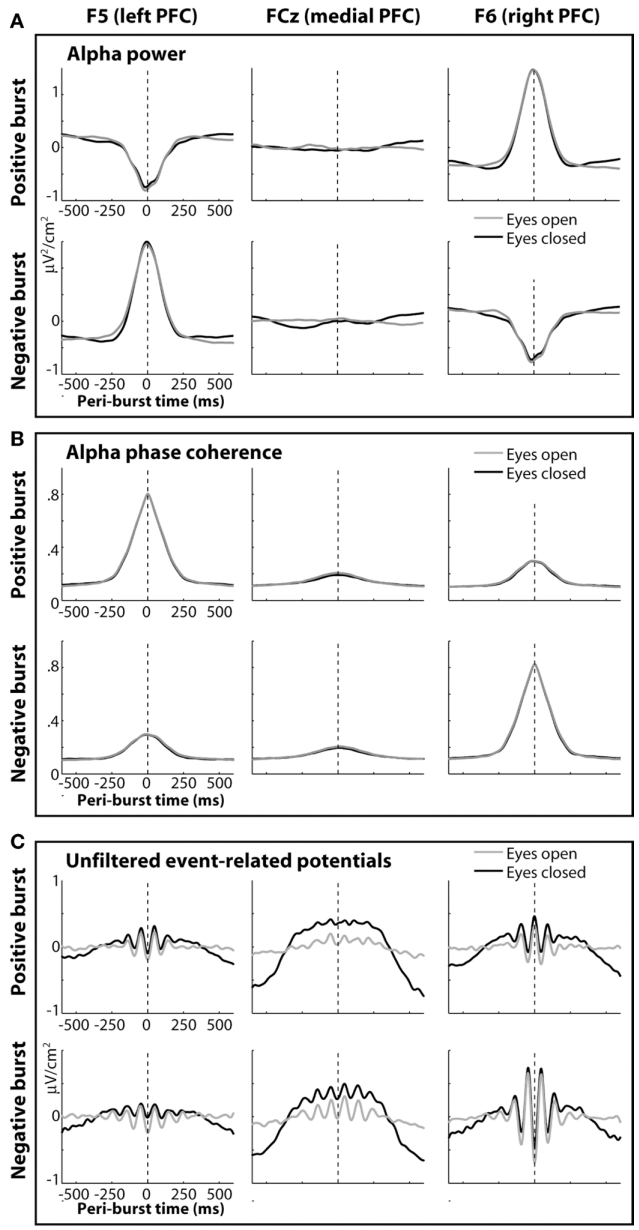



**FIGURE 12 | The temporal dynamics of power and inter-burst phase coherence for left (F5) medial (FCz) and right (F6) loci.** Bursts were defined based on the power difference ln(F6)–ln(F5), with positive bursts representing extreme positive values on this metric, and negative bursts representing extreme negative values on this metric. **(A)** displays peri-burst alpha power; **(B)** displays inter-burst phase coherence; **(C)** displays event-related potentials, created by time-locking unfiltered EEG to each burst.

decay approximately 200 ms following each burst. This enhancement of contralateral phase coherence may lead to the seemingly paradoxical observation that event-related potentials (**Figure 12C**) appear to contain more alpha activity in the contralateral hemisphere. That is, the burst-phase-locked alpha activity (giving rise to event-related potentials) is strongest when total peri-burst alpha power actually decreases. This effect was more pronounced in the right hemisphere compared to the left hemisphere (**Figures 12A,C**). The robust alpha in the event-related potentials in **Figure 12C** is striking considering no temporal filters were applied to the data.

### *Peri-burst oscillatory dynamics*

Finally, we examined the oscillatory dynamics in a broader frequency range through wavelet convolution techniques. Note that because the entire time period (−500 to +500 ms) was used for decibel conversion, in/decreases in peri-burst alpha power imply a relative de/increase in alpha power before and after the burst. As depicted in **Figure 13**, clear alpha modulation surrounds each burst, with F5 and F6 revealing effects of similar size but opposite direction. As also shown in the figure, there were lower theta-band dynamics that were similar at both F5 and F6, during both positive and negative bursts. Although blinks were rejected from the raw data prior to any signal processing, we suspected that given the low-frequency and ubiquitous nature of these lower theta-band dynamics, that they may be ocular in origin. Thus we examined oscillatory dynamics separately for bursts that occurred during eyes-open and eyes-closed resting periods. As shown in **Figure 14**, the site-specific alpha-band dynamics are highly consistent across eyes-open and eyes-closed resting periods, but the lower theta-band dynamics predominate only when eyes are open, with a relative theta-band suppression accompanying both positive and negative bursts.

To examine more specifically the topography of the lower theta-band peri-burst suppression, **Figure 15** shows the scalp topography of theta-band (4–7 Hz) power preceding and following the bursts. The topography is highly consistent for both positive and negative bursts, with relative theta-band suppression occurring surrounding the time of the alpha asymmetry bursts.

## DISCUSSION

Conventional FFT-based approaches to the analysis of frontal EEG asymmetry acquired during the resting-state produce metrics that lack temporal and spatial precision. As an initial endeavor to improve both the temporal and spatial precision by which frontal EEG asymmetry is assessed, we present metrics that capture a recurrent dynamic endogenous neural process – asymmetry bursts – that transpires during the lengthy resting-state. Moreover, these metrics were based on a CSD transformation, which had advantages for time-frequency dynamics at the level of the scalp (Cavanagh et al., 2009; Cohen et al., 2009a). These metrics, although reflecting a distinct minority of the resting assessment period, nonetheless account for a substantial proportion of variance in the conventional metrics that sum across the entire interval.

Peri-burst alpha power was found to be relatively closely related to conventional FFT-derived alpha power, at constituent target sites F5 and F6, as well as in terms of asymmetry in alpha power between F5 and F6. In the latter case of power asymmetry, peri-burst alpha power asymmetry, which represented but a small fraction (1%) of the total resting period, accounted for 42% of the variance in FFT-derived alpha power asymmetry across the entire resting period. Moreover, peri-burst alpha asymmetry, like FFT-derived alpha power asymmetry, was sensitive to a history of major depression, with both current and previously depressed individuals showing relatively lower scores, indicating greater relative left frontal alpha

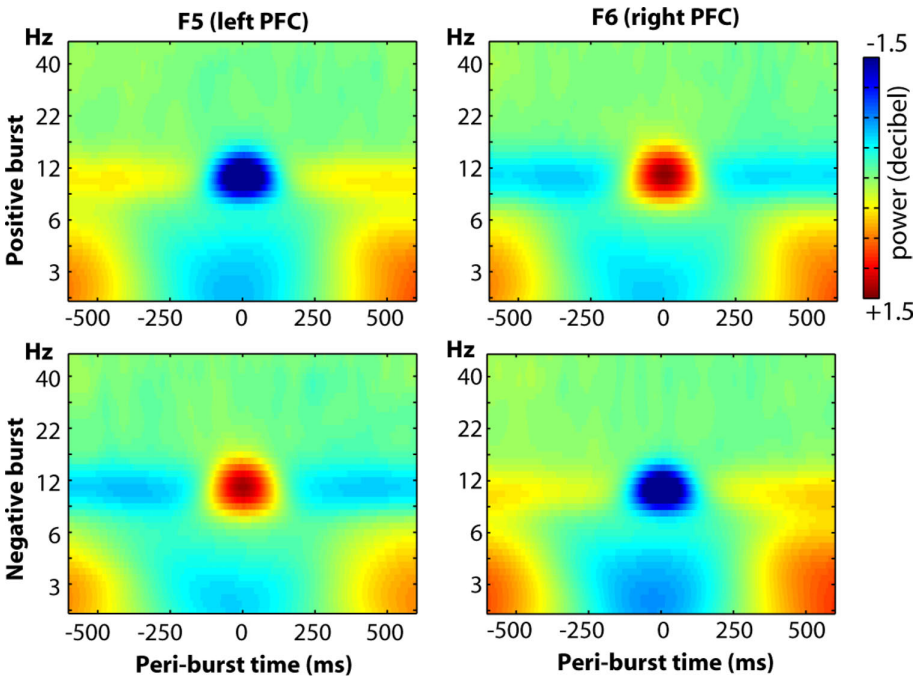



**FIGURE 13 | The oscillatory dynamics of a wider frequency range for sites F5 and F6, displayed as a function of positive and negative bursts.**

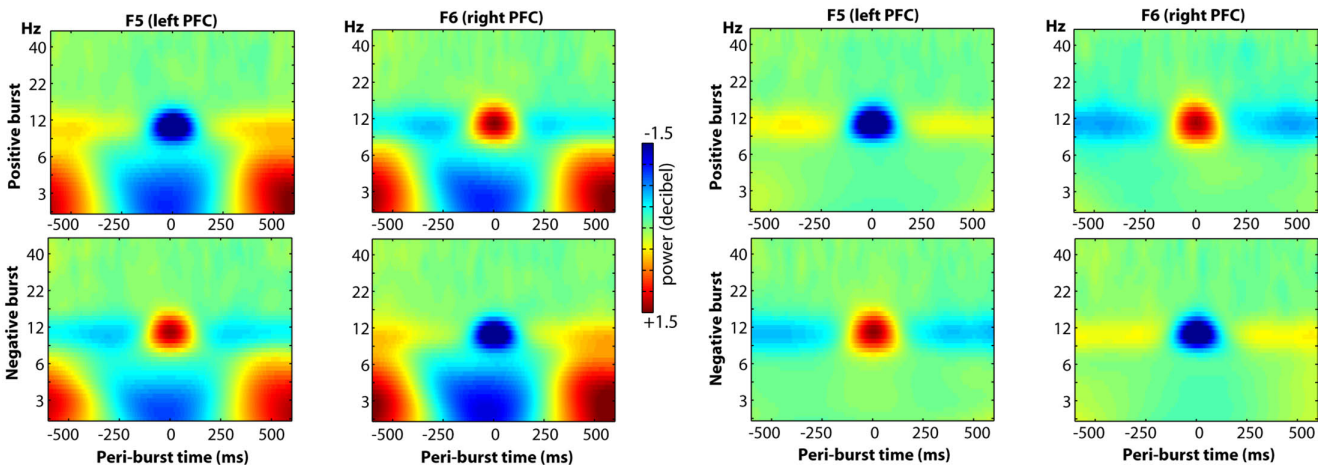



**FIGURE 14 | The oscillatory dynamics of a wider frequency range for sites F5 and F6, as a function of positive and negative bursts, depicted separately for eyes-open and eyes-closed epochs.** Clear alpha modulation surrounds each burst, with F5 and F6 revealing effects of similar size but opposite direction. These effects are highly consistent across both eyes-open **(A)** and eyes-closed **(B)** conditions, but a relative theta-band suppression accompanying both positive and negative bursts is seen only in the eyes-open **(A)** condition.

and by inference less left frontal activity, than never-depressed individuals. Thus peri-burst alpha power asymmetry is consistent with a rather sizable literature linking altered frontal EEG asymmetry to emotion and depression (Coan and Allen, 2004; Thibodeau et al., 2006), yet as a temporally more sensitive measure, peri-burst alpha power asymmetry holds the promise that future investigations may uncover more precisely what transpires in terms of neural dynamics that ultimately may create risk for depression.

The inter-burst phase metrics were also sensitive to MDD status, and in this case only to current, but not past, MDD status, and only over left frontal cortex during positive bursts. The findings with inter-burst phase coherence were exploratory but provocative, and further suggest altered neural dynamics in left frontal cortex as a feature of MDD, consistent with findings using a wide variety of methodologies including EEG (Coan and Allen, 2004; Thibodeau et al., 2006), positron emission tomography (e.g., Bench et al., 1993),

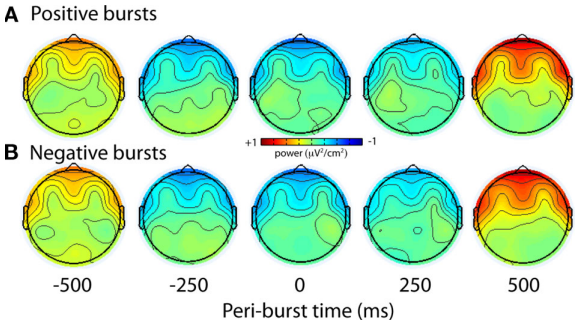



**FIGURE 15 | Topographical distribution of theta (4–7 Hz) activity during eyes-open condition.** The scalp topography of theta-band (4–7 Hz) power preceding and following the bursts is highly consistent for both positive **(A)** and negative **(B)** bursts, with relative theta-band suppression occurring at frontal and lateral frontal regions surrounding the time of the alpha asymmetry bursts.

and transcranial magnetic stimulation (e.g., George et al., 2010). The functional significance of the inter-burst phase metrics in MDD must await future research, but the specificity to current MDD suggests that such measures might be worth exploring within the context of treatment studies to observe is successful treatment is reflected in normalized inter-burst phase coherence, and whether baseline inter-burst phase coherence might predict treatment response. These findings also further highlight the rich temporal dynamics of the endogenous processes that underlie conventional FFT-based metrics to studying frontal asymmetry.

This initial investigation into characterizing the neural dynamics surrounding the bursts indicates that the asymmetry bursts are characterized by a transient lateralized alpha suppression that shows a highly consistent phase relationship across bursts, and a concurrent contralateral transient alpha enhancement that is less tightly phase-locked across bursts. These findings may be similar to those seen in event-related desynchronization investigations (e.g., Pfurtscheller, 1992), where regions involved in active processing show alpha suppression (desynchronization) while regions uninvolved in such processing show alpha enhancement (event-related synchronization; e.g., Pfurtscheller, 1992). Indeed, alpha has been implicated in the active suppression of irrelevant or distracting information during active task engagement (Klimesch et al., 2007; van Dijk et al., 2008). Speculatively, then, the recurrent burst-related alpha suppression may reflect recurrent bouts of active lateralized cortical processing across analogous frontal regions. In absence of an active task it is challenging to link these dynamics to specific cognitive/emotional processes; simultaneous autonomic recordings may help (e.g., to examine whether pupillary or cardiac responses coincide with asymmetry bursts). Examining such asymmetry bursts during specific cognitive tasks may also help elucidate candidate mental operations that underlie bursts. Nonetheless, the fact that the alpha suppression is particularly tightly phase-locked across bursts raises the possibility that the lateralized alpha suppression may drive or regulate cortical processing. It is possible that the reach of frontal alpha bursts extends beyond frontal cortical areas. For example, the human nucleus accumbens exhibits strong alpha-linked cross-frequency coupling, such that transient increases in gamma power occur preferentially at specific phases of nucleus accumbens alpha oscillations (Cohen et al., 2009b). Nucleus accumbens alpha was also phase-synchronized with frontal cortical alpha (unpublished observations in that same dataset), suggesting the possibility that alpha asymmetry dynamics regulate active processing in the ventral striatum, which has been implicated in depression and is one of the main targets for deep-brain-stimulation treatment of major depression (Schlaepfer et al., 2008).

We observed a decrease in theta activity at anterior frontal and lateral sites during alpha asymmetry bursts only during eyes-open condition. Although we cannot completely rule out that ocular activity contributed to these dynamics, they do not appear to be entirely driven by eye movements, for the following reasons: epochs with blinks were removed from the data, the extreme lateral and medial frontal spatial components are atypical of eye movements, and non-blink oculomotor artifacts in EEG have been associated with higher frequency (gamma) power (Yuval-Greenberg et al., 2008) and nose-reference (we used a CSD reference, which is a local referencing scheme and therefore minimizes contribution of ocular artifacts on EEG electrodes). Importantly, it is clear from both **Figures 13 and 14** that the lower theta-band effects do not extend into the alpha range, suggesting that the peri-burst alpha effects observed here are not influenced by the lower frequency modulations during eyes-open conditions, and, moreover, the alpha-band dynamics are robust during both eyes-open and eyes-closed conditions.

In the spirit of exploration, the present investigation may raise more questions than it answers, but it also lays a foundation for new directions in research on resting-state frontal brain asymmetry. The novel metrics presented here retain the advantage of the conventional metrics in terms of accounting for substantial variance in the conventional metrics and retaining sensitivity to a history of major depression, but these metrics provide the possibility to examine the spatio-temporal dynamics of brain activity that may underlie the resting-state, and thus give rise to risk for depression. Moreover, the ability to characterize specific relevant time points within the ongoing resting-state provides the possibility to utilize EEG-defined events, in the form of asymmetry bursts, to define time windows of interest with other concurrent imaging modalities that may reveal clues as to the localized origin of these bursts (e.g., concurrent fMRI, or ICA decomposition of EEG surrounding bursts), as well as autonomic and behavioral correlates of these recurrent endogenous events. The promise of this approach is bolstered by a recent investigation linking EEG-defined microstates to BOLD activation patterns, many within the widely studied resting-state networks (Musso et al., 2010). Alpha burst-driven transcranial magnetic stimulation may also be an important avenue of exploration not only for examining the neural dynamics surrounding bursts, but also as a possible improvement over TMS-guided depression treatment (e.g., Schutter, 2009; George et al., 2010).

The present approach is illustrative for future investigations of resting-state brain activity, including those utilizing frontal EEG asymmetry as well as other approaches such as fMRI investigations of resting-state activity (e.g., "default mode network," Raichle et al., 2001; Smith et al., 2009). Although much has been gleaned about neural systems under tightly controlled exogenous presentation of experimental stimuli, there remains much to be learned about how the brain functions under its own endogenous control. The

resting-state approaches have gained some traction in that pursuit, but have also suffered from the virtually inevitable decrease in precision that accompanies investigating lengthy uncontrolled time periods. By parsing relatively long resting time periods according to meaningful and recurrent endogenous signals such as asymmetry bursts or autonomic responses (e.g., Critchley et al., 2002; Siegle et al., 2003), or perhaps patterns of interconnectivity using the BOLD signal (cf., De Luca et al., 2006), greater progress in the elucidation of the neural basis of endogenous brain activity may be possible. It is with that hope that we offer the present exploration with endogenous EEG alpha asymmetry bursts.

## ACKNOWLEDGMENTS

This research was supported in part by grants from the National Institutes of Health (R01 MH066902) and the National Alliance for Research on Schizophrenia and Depression (NARSAD) to John Allen, and by a VIDI grant from the Dutch Organization for Scientific Research (NWO) to Michael Cohen. The authors wish to thank Jennifer Stewart and David Towers for their tireless assistance with data management and reduction, and to thank Andrew Bismark, Craig Santerre, Eynav Accortt, Amanda Brody, Jay Hegde, Eliza Fergerson, Jamie Velo, Dara Halpern, and myriad research assistants for their help on this project.

## REFERENCES

Accortt, E. E., and Allen, J. J. (2006). Frontal EEG asymmetry and premenstrual dysphoric symptomatology. *J. Abnorm. Psychol.* 115, 179–184.

Allen, J. J., Harmon-Jones, E., and Cavender, J. H. (2001). Manipulation of frontal EEG asymmetry through biofeedback alters self-reported emotional responses and facial EMG. *Psychophysiology* 38, 685–693.

Allen, J. J., Iacono, W. G., Depue, R. A., and Arbisi, P. (1993). Regional electroencephalographic asymmetries in bipolar seasonal affective disorder before and after exposure to bright light. *Biol. Psychiatry* 33, 642–646.

Allen, J. J., and Kline, J. P. (2004). Frontal EEG asymmetry, emotion, and psychopathology: the first, and the next 25 years. *Biol. Psychol.* 67, 1–5.

Allen, J. J. B., Coan, J. A., and Nazarian, M. (2004a). Issues and assumptions on the road from raw signals to metrics of frontal EEG asymmetry in emotion. *Biol. Psychol.* 67, 183–218.

Allen, J. J. B., Urry, H. L., Hitt, S. K., and Coan, J. A. (2004b). The stability of resting frontal electroencephalographic asymmetry in depression. *Psychophysiology* 41, 269–280.

Anokhin, A. P., Heath, A. C., and Myers, E. (2006). Genetic and environmental influences on frontal EEG asymmetry: a twin study. *Biol. Psychol.* 71, 289–295.

Baehr, E., Rosenfeld, J. P., Baehr, R., and Earnest, C. (1998). Comparison of two EEG asymmetry indices in depressed patients vs. normal controls. *Int. J. Psychophysiol.* 31, 89–92.

Bagiella, E., Sloan, R. P., and Heitjan, D. F. (2000). Mixed-effects models in psychophysiology. *Psychophysiology* 37, 13–20.

Beck, A. T., Steer, R. A., and Brown, G. K. (1996). *Manual for the Beck Depression Inventory-II*. San Antonio, TX: Psychological Corporation.

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., and Erbaugh, J. (1961). An inventory for measuring depression. *Arch. Gen. Psychiatry* 4, 561–571.

Bench, C. J., Friston, K. J., Brown, R. G., Frackowiak, R. S., and Dolan, R. J. (1993). Regional cerebral blood flow in depression measured by positron emission tomography: the relationship with clinical dimensions. *Psychol. Med.* 23, 579–590.

Bruder, G. E., Fong, R., Tenke, C. E., Leite, P., Towey, J. P., and Stewart, J. E. (1997). Regional brain asymmetries in major depression with or without an anxiety disorder: a quantitative electroencephalographic study. *Biol. Psychiatry* 41, 939–948.

Cacioppo, J. T. (2004). Feelings and emotions: roles for electrophysiological markers. *Biol. Psychol.* 67, 235–243.

Cavanagh, J. F., Cohen, M. X., and Allen, J. J. (2009). Prelude to and resolution of an error: EEG phase synchrony reveals cognitive control dynamics during action monitoring. *J. Neurosci.* 29, 98–105.

Chapman, L. J., and Chapman, J. P. (1987). The measurement of handedness. *Brain Cogn.* 6, 175–183.

Coan, J. A., and Allen, J. J. (2004). Frontal EEG asymmetry as a moderator and mediator of emotion. *Biol. Psychol.* 67, 7–49.

Coan, J. A., Allen, J. J., and McKnight, P. E. (2006). A capability model of individual differences in frontal EEG asymmetry. *Biol. Psychol.* 72, 198–207.

Cohen, M. X., van Gaal, S., Ridderinkhof, K. R., and Lamme, V. A. (2009a). Unconscious errors enhance prefrontal-occipital oscillatory synchrony. *Front. Hum. Neurosci.* 3:54. doi: 10.3389/neuro.09.054.2009.

Cohen, M. X., Axmacher, N., Lenartz, D., Elger, C. E., Sturm, V., and Schlaepfer, T. E. (2009b). Good vibrations: cross-frequency coupling in the human nucleus accumbens during reward processing. *J. Cogn. Neurosci.* 21, 875–889.

Critchley, H. D., Melmed, R. N., Featherstone, E., Mathias, C. J., and Dolan, R. J. (2002). Volitional control of autonomic arousal: a functional magnetic resonance study. *Neuroimage* 16, 909–919.

Davidson, R. J. (1998a). Anterior electrophysiological asymmetries, emotion, and depression: conceptual and methodological conundrums. *Psychophysiology* 35, 607–614.

Davidson, R. J. (1998b). Affective style and affective disorders: perspectives from affective neuroscience. *Cogn. Emot.* 12, 307–330.

Davidson, R. J. (2004). What does the prefrontal cortex "do" in affect: perspectives on frontal EEG asymmetry research. *Biol. Psychol.* 67, 219–233.

Davidson, R. J., Coe, C. C., Dolski, I., and Donzella, B. (1999). Individual differences in prefrontal activation asymmetry predict natural killer cell activity at rest and in response to challenge. *Brain Behav. Immun.* 13, 93–108.

Davidson, R. J., Marshall, J. R., Tomarken, A. J., and Henriques, J. B. (2000). While a phobic waits: regional brain electrical and autonomic activity in social phobics during anticipation of public speaking. *Biol. Psychiatry* 47, 85–95.

Dawson, G., Frey, K., Panagiotides, H., Osterling, J., and Hessl, D. (1997). Infants of depressed mothers exhibit atypical frontal brain activity: a replication and extension of previous findings. *J. Child. Psychol. Psychiatry* 38, 179–186.

Dawson, G., Frey, K., Panagiotides, H., Yamada, E., Hessl, D., and Osterling, J. (1999). Infants of depressed mothers exhibit atypical frontal electrical brain activity during interactions with mother and with a familiar, nondepressed adult. *Child Dev.* 70, 1058–1066.

De Luca, M., Beckmann, C. F., De Stefano, N., Matthews, P. M., and Smith, S. M. (2006). fMRI resting state networks define distinct modes of long-distance interactions in the human brain. *Neuroimage* 29, 1359–1367.

Debener, S., Beauducel, A., Nessler, D., Brocke, B., Heilemann, H., and Kayser, J. (2000). Is resting anterior EEG alpha asymmetry a trait marker for depression? Findings for healthy adults and clinically depressed patients. *Neuropsychobiology* 41, 31–37.

Delorme, A., and Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* 134, 9–21.

Diego, M. A., Jones, N. A., and Field, T. (2010). EEG in 1-week, 1-month and 3-month-old infants of depressed and non-depressed mothers. *Biol. Psychol.* 83, 7–14.

First, M. B., Spitzer, R. L., Gibbon, M., and Williams, J. B. W. (November, 2002). *Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Non-patient Edition (SCID-I/NP)* New York: Biometrics Research, New York State Psychiatric Institute.

Fox, N. A., Rubin, K. H., Calkins, S. D., Marshall, T. R., Coplan, R. J., Porges, S. W., Long, J. M., and Stewart, S. (1995). Frontal activation asymmetry and social competence at four years of age. *Child Dev.* 66, 1770–1784.

George, M. S., Lisanby, S. H., Avery, D., McDonald, W. M., Durkalski, V., Pavlicova, M., Anderson, B., Nahas, Z., Bulow, P., Zarkowski, P., Holtzheimer, P. E. III, Schwartz, T., and Sackeim, H. A. (2010). Daily left prefrontal transcranial magnetic stimulation therapy for major depressive disorder: a sham-controlled randomized trial. *Arch. Gen. Psychiatry* 67, 507–516.

Gotlib, I. H., Ranganath, C., and Rosenfeld, J. P. (1998). Frontal EEG alpha asymmetry, depression, and cognitive functioning. *Cogn. Emot.* 12, 449–478.

Gottesman, I. I., and Gould, T. D. (2003). The endophenotype concept in psychiatry: etymology and strategic intentions. *Am. J. Psychiatry* 160, 636–645.

Hagemann, D., Naumann, E., and Thayer, J. F. (2001). The quest for the EEG reference revisited: a glance from brain asymmetry research. *Psychophysiology* 38, 847–857.

Hagemann, D., Naumann, E., Thayer, J. F., and Bartussek, D. (2002). Does resting electroencephalograph asymmetry reflect a trait? An application of latent state-trait theory. *J. Pers. Soc. Psychol.* 82, 619–641.

Hamilton, M. (1960). A rating scale for depression. *J. Neurol. Neurosurg. Psychiatry* 23, 56–62.

Handel, B. F., Haarmeier, T., and Jensen, O. (2010). Alpha oscillations correlate with the successful inhibition of unattended stimuli. *J. Cogn. Neurosci*. doi: 10.1162/jocn.2010.21557. [Epub ahead of print].

Hedges, L., and Olkin, I. (1985). *Statistical Methods for Meta-Analysis*. New York: Academic Press.

Henriques, J. B., and Davidson, R. J. (1990). Regional brain electrical asymmetries discriminate between previously depressed and healthy control subjects. *J. Abnorm. Psychol.* 99, 22–31.

Henriques, J. B., and Davidson, R. J. (1991). Left frontal hypoactivation in depression. *J. Abnorm. Psychol.* 100, 535–545.

Iacono, W. G. (1998). Identifying psychophysiological risk for psychopathology: examples from substance abuse and schizophrenia research. *Psychophysiology* 35, 621–637.

Jones, N. A., Field, T., Fox, N. A., Lundy, B., and Davalos, M. (1997). EEG activation in 1-month-old infants of depressed mothers. *Dev. Psychopathol.* 9, 491–505.

Kang, D. H., Davidson, R. J., Coe, C. L., Wheeler, R. W., Tomarken, A. J., and Ershler, W. B. (1991). Frontal brain asymmetry and immune function. *Behav. Neurosci.* 105, 860–869.

Kayser, J., and Tenke, C. E. (2006a). Principal components analysis of Laplacian waveforms as a generic method for identifying ERP generator patterns: I. Evaluation with auditory oddball tasks. *Clin. Neurophysiol.* 117, 348–368.

Kayser, J., and Tenke, C. E. (2006b). Principal components analysis of Laplacian waveforms as a generic method for identifying ERP generator patterns: II. Adequacy of low-density estimates. *Clin. Neurophysiol.* 117, 369–380.

Klimesch, W., Sauseng, P., and Hanslmayr, S. (2007). EEG alpha oscillations: the inhibition-timing hypothesis. *Brain Res. Rev.* 53, 63–88.

Mathersul, D., Williams, L. M., Hopkinson, P. J., and Kemp, A. H. (2008). Investigating models of affect: relationships among EEG alpha asymmetry, depression, and anxiety. *Emotion* 8, 560–572.

Musso, F., Brinkmeyer, J., Mobascher, A., Warbrick, T., and Winterer, G. (2010). Spontaneous brain activity and EEG microstates. A novel EEG/fMRI analysis approach to explore resting-state networks. *Neuroimage* 52, 1149–1161.

Nitschke, J. B., Heller, W., Palmieri, P. A., and Miller, G. A. (1999). Contrasting patterns of brain activity in anxious apprehension and anxious arousal. *Psychophysiology* 36, 628–637.

Palva, S., and Palva, J. M. (2007). New vistas for alpha-frequency band oscillations. *Trends Neurosci.* 30, 150–158.

Perrin, F., Bertrand, O., Giard, M. H., and Pernier, J. (1990). Precautions in topographic mapping and in evoked potential map reading. *J. Clin. Neurophysiol.* 7, 498–506.

Perrin, F., Pernier, J., Bertrand, O., and Echallier, J. F. (1989). Spherical splines for scalp potential and current density mapping. *Electroencephalogr. Clin. Neurophysiol.* 72, 184–187.

Pfurtscheller, G. (1992). Event-related synchronization (ERS): an electrophysiological correlate of cortical areas at rest. *Electroencephalogr. Clin. Neurophysiol.* 83, 62–69.

Pizzagalli, D. A., Nitschke, J. B., Oakes, T. R., Hendrick, A. M., Horras, K. A., Larson, C. L., Abercrombie, H. C., Schaefer, S. M., Koger, J. V., Benca, R. M., Pascual-Marqui, R. D., and Davidson, R. J. (2002). Brain electrical tomography in depression: the importance of symptom severity, anxiety, and melancholic features. *Biol. Psychiatry* 52, 73–85.

Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., and Shulmani, G. L. (2001). A default mode of brain function. *Proc. Natl. Acad. Sci. U.S.A.* 98, 676–682.

Reid, S. A., Duke, L. M., and Allen, J. J. (1998). Resting frontal electroencephalographic asymmetry in depression: inconsistencies suggest the need to identify mediating factors. *Psychophysiology* 35, 389–404.

Schlaepfer, T. E., Cohen, M. X., Frick, C., Kosel, M., Brodesser, D., Axmacher, N., Joe, A. Y., Kreft, M., Lenartz, D., and Sturm, V. (2008). Deep brain stimulation to reward circuitry alleviates anhedonia in refractory major depression. *Neuropsychopharmacology* 33, 368–377.

Schmidt, L. A., and Fox, N. A. (1994). Patterns of cortical electrophysiology and autonomic activity in adults' shyness and sociability. *Biol. Psychol.* 38, 183–198.

Schutter, D. J. (2009). Antidepressant efficacy of high-frequency transcranial magnetic stimulation over the left dorsolateral prefrontal cortex in double-blind sham-controlled designs: a meta-analysis. *Psychol. Med.* 39, 65–75.

Siegle, G. J., Steinhauer, S. R., Stenger, V. A., Konecky, R., and Carter, C. S. (2003). Use of concurrent pupil dilation assessment to inform interpretation and analysis of fMRI data. *Neuroimage* 20, 114–124.

Smit, D. J., Posthuma, D., Boomsma, D. I., and De Geus, E. J. (2007). The relation between frontal EEG asymmetry and the risk for anxiety and depression. *Biol. Psychol.* 74, 26–33.

Smith, S. M., Fox, P. T., Miller, K. L., Glahn, D. C., Fox, P. M., Mackay, C. E., Filippini, N., Watkins, K. E., Toro, R., Laird, A. R., and Beckmann, C. F. (2009). Correspondence of the brain's functional architecture during activation and rest. *Proc. Natl. Acad. Sci. U.S.A.* 106, 13040–13045.

Srinivasan, R., Nunez, P. L., Tucker, D. M., Silberstein, R. B., and Cadusch, P. J. (1996). Spatial sampling and filtering of EEG with spline laplacians to estimate cortical potentials. *Brain Topogr.* 8, 355–366.

Srinivasan, R., Winter, W. R., Ding, J., and Nunez, P. L. (2007). EEG and MEG coherence: measures of functional connectivity at distinct spatial scales of neocortical dynamics. *J. Neurosci. Methods* 166, 41–52.

Stewart, J. L., Bismark, A. W., Towers, D. N., Coan, J. A., and Allen, J. J. B. (2010). Resting frontal EEG asymmetry as an endophenotype for depression risk: sex-specific patterns of frontal brain asymmetry. *J. Abnorm. Psychol.* 119, 502–512.

Stewart, J. L., Levin-Silton, R., Sass, S. M., Heller, W., and Miller, G. A. (2008). Anger style, psychopathology, and regional brain activity. *Emotion* 8, 701–713.

Thibodeau, R., Jorgensen, R. S., and Kim, S. (2006). Depression, anxiety, and resting frontal EEG asymmetry: a meta-analytic review. *J. Abnorm. Psychol.* 115, 715–729.

Tomarken, A. J., Dichter, G. S., Garber, J., and Simien, C. (2004). Resting frontal brain activity: linkages to maternal depression and socio-economic status among adolescents. *Biol. Psychol.* 67, 77–102.

Towers, D. N., and Allen, J. J. (2009). A better estimate of the internal consistency reliability of frontal EEG asymmetry scores. *Psychophysiology* 46, 132–142.

Urry, H. L., Nitschke, J. B., Dolski, I., Jackson, D. C., Dalton, K. M., Mueller, C. J., Rosenkranz, M. A., Ryff, C. D., Singer, B. H., and Davidson, R. J. (2004). Making a life worth living: neural correlates of well-being. *Psychol. Sci.* 15, 367–372.

van Dijk, H., Schoffelen, J. M., Oostenveld, R., and Jensen, O. (2008). Prestimulus oscillatory activity in the alpha band predicts visual discrimination ability. *J. Neurosci.* 28, 1816–1823.

Wiedemann, G., Pauli, P., Dengler, W., Lutzenberger, W., Birbaumer, N., and Buchkremer, G. (1999). Frontal brain asymmetry as a biological substrate of emotions in patients with panic disorders. *Arch. Gen. Psychiatry* 56, 78–84.

Winter, W. R., Nunez, P. L., Ding, J., and Srinivasan, R. (2007). Comparison of the effect of volume conduction on EEG coherence with the effect of field spread on MEG coherence. *Stat. Med.* 26, 3946–3957.

Yuval-Greenberg, S., Tomer, O., Keren, A. S., Nelken, I., and Deouell, L. Y. (2008). Transient induced gamma-band response in EEG as a manifestation of miniature saccades. *Neuron* 58, 429–441.

**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: 23 July 2010; accepted: 11 December 2010; published online: 29 December 2010.*

*Citation: Allen JJB and Cohen MX (2010) Deconstructing the "resting" state: exploring the temporal dynamics of frontal alpha asymmetry as an endophenotype for depression. Front. Hum. Neurosci.* ***4****:232. doi: 10.3389/fnhum.2010.00232*

frontiers in
HUMAN NEUROSCIENCE

HYPOTHESIS AND THEORY ARTICLE
published: 10 February 2011
doi: 10.3389/fnhum.2011.00017

# Mental training as a tool in the neuroscientific study of brain and cognitive plasticity

***Heleen A. Slagter[1]*, Richard J. Davidson[2,3,4] and Antoine Lutz[2]***

[1] *Brain and Cognition Unit, Department of Psychology, University of Amsterdam, Amsterdam, Netherlands*
[2] *Waisman Laboratory for Brain Imaging and Behavior, University of Wisconsin, Madison, WI, USA*
[3] *Department of Psychology, University of Wisconsin, Madison, WI, USA*
[4] *Center for Investigating Healthy Minds, University of Wisconsin, Madison, WI, USA*

***Edited by:***
*Michael X. Cohen, University of Amsterdam, Netherlands*

***Reviewed by:***
*Michael Posner, University of Oregon, USA*
*Tonya L. Jacobs, University of California, Davis, USA*

****Correspondence:***
*Heleen A. Slagter, Department of Psychology, University of Amsterdam, Roeterstraat 15, 1018 WB Amsterdam, Netherlands.*
*e-mail: h.a.slagter@uva.nl*

Although the adult brain was once seen as a rather static organ, it is now clear that the organization of brain circuitry is constantly changing as a function of experience or learning. Yet, research also shows that learning is often specific to the trained stimuli and task, and does not improve performance on novel tasks, even very similar ones. This perspective examines the idea that systematic mental training, as cultivated by meditation, can induce learning that is not stimulus or task specific, but process specific. Many meditation practices are explicitly designed to enhance specific, well-defined core cognitive processes. We will argue that this focus on enhancing core cognitive processes, as well as several general characteristics of meditation regimens, may specifically foster process-specific learning. To this end, we first define meditation and discuss key findings from recent neuroimaging studies of meditation. We then identify several characteristics of specific meditation training regimes that may determine process-specific learning. These characteristics include ongoing variability in stimulus input, the meta-cognitive nature of the processes trained, task difficulty, the focus on maintaining an optimal level of arousal, and the duration of training. Lastly, we discuss the methodological challenges that researchers face when attempting to control or characterize the multiple factors that may underlie meditation training effects.

**Keywords: plasticity, training, mental training, meditation, neuroimaging, cognition, brain**

## INTRODUCTION

The ability to learn is a function of brain plasticity and essential to the survival of all animals. Humans appear remarkable in this respect, as they can acquire a wide range of skills given appropriate training. Neuroscience research within the last few decades confirms that the human brain is plastic, and much more so than once thought possible (Buonomano and Merzenich, 1998). One of the most salient findings is perhaps the observation that the *adult* brain can still change significantly in both structure and function as a result of experience (e.g., Maguire et al., 2000; Draganski et al., 2004; Boyke et al., 2008). For example, taxi drivers, who obtain extensive navigational experience as adults, have larger posterior hippocampi, a brain structure important for spatial representation of the environment (Maguire et al., 2000). Not only long-term expertise, but also relatively short practice has been associated with neural changes in adults. For example, performing a five-finger piano exercise for 2 h on five consecutive days resulted in an enlargement of primary motor areas representing the finger muscles, which was accompanied by improved performance (Pascual-Leone et al., 1995, 2005). Remarkably, this study also showed that *mental* practice of the same finger exercise resulted in a similar reorganization of the motor cortex to the one observed in the group of participants that physically practiced the movements. Thus, plasticity appears to be an intrinsic property of the nervous system retained throughout a lifespan and the obligatory consequence of all neural activity, including mental practice. Indeed, it is now commonly held that the brain undergoes continuous changes in response to each sensory input, motor act, association, reward signal, action plan, and awareness itself (Pascual-Leone et al., 2005).

Yet, research also shows that learning is typically highly specific (Salomon and Perkins, 1989; Kramer and Willis, 2003; Green and Bavelier, 2008; Roelfsema et al., 2010). While individuals trained on a task will improve on that very task, other tasks, even very similar ones, often show little or no improvement. For example, in visual learning tasks, improvements are often not observed if the untrained test stimulus has a different orientation or contrast than the trained stimulus (Roelfsema et al., 2010). Similarly, memory training often enhances memory for the trained content, but not memory skills themselves. In the example of the taxi drivers, follow-up research showed that the taxi drivers did not exhibit improved performance on other memory tasks (Woollett and Maguire, 2009). Thus, training benefits are often stimulus or content specific rather than *process* specific, a fact that is well documented in the perceptual (Roelfsema et al., 2010), cognitive (Kramer and Willis, 2003), and motor (Bachman, 1961) domain. This is obviously a potentially severe impediment for training programs in educational or clinical settings, where the goal is to improve performance in everyday life and which thus necessarily require a general improvement in skills.

The aim of this paper is twofold: (i) to examine the hypothesis that systematic mental training can induce process-specific learning, that is, learning that generalizes to novel stimuli and task contexts, and (ii) to identify aspects of training contexts that may contribute to process-specific learning from the vantage point of mental training. If plasticity is indeed the obligatory consequence of not only each sensory input or motor act, but also of each thought process, changes in brain circuitry due to mental training

should lead to process-specific learning. Here, we use the term "process-specific learning" to denote learning effects that do not only improve performance on the trained task or tasks, but also transfer to new tasks and domains (cf. Green and Bavelier, 2008), i.e., learning that is not specific to the trained stimuli or tasks. The main focus of the paper will be on effects of mental training of cognitive control processes. "Cognitive control" refers to a collection of processes that allow us flexibly adapt our behavior in the pursuit of an internal goal, and includes processes such as selection of goal-relevant information, performance monitoring, interference resolution, and storage and manipulation of information in working memory. These processes are generally considered to be largely independent from one another, because they are differentially affected by independent variables, and can be behaviorally and neurally dissociated. There is an emerging consensus in the literature that this conglomerate of largely independent, but constantly interacting control processes is domain general (e.g., Baddeley, 1996; Fuster, 1997; Smith and Jonides, 1999; Miyake et al., 2000). Specifically, the general proposal is that the engagement of a particular cognitive control process (e.g., performance monitoring) within any one cognitive task is simply a matter of the degree to which load is exercised on that control mechanism and should extend to any cognitive challenge that incorporates sufficient control requirements of the same kind. The focus of this paper on the amenability of cognitive control processes to training is thus especially valuable because these abilities are fundamental to higher cognition and contribute to performance across cognitive domains (e.g., attention, working memory, long-term memory; Smith and Jonides, 1999; Duncan and Owen, 2000).

This paper will focus in particular on mental training of cognitive control processes as cultivated by meditation. Little is currently known about the amenability of cognitive control processes to mental training. Most studies to date have focused on mental training in the motor domain (for recent reviews, see, e.g., Guillot and Collet, 2005; Munzert et al., 2009). This is likely related to the fact that it is difficult to design controlled mental training contexts in which an individual can consistently practice a specific cognitive control skill. Yet, being able to enhance cognitive control function through training is especially valuable since, as mentioned above, cognitive control skills, such as the ability to focus attention in face of distraction, contribute to performance on virtually any task (Smith and Jonides, 1999; Duncan and Owen, 2000). Notably, many meditation practices are based on precisely descriptive and highly detailed theories (Gyatso and Jinpa, 1995; Gunaratana, 2007) and are explicitly designed to train such core cognitive skills, and typically in the absence of performing some external task (Lutz et al., 2007, 2008). Furthermore, many of these practices include a focus on maintaining optimal levels of arousal and motivation, and ascending stages through which the practitioner passes, ensuring continuously challenging training conditions. These are all factors that may foster more general learning (e.g., Green and Bavelier, 2008), as will be discussed in more detail below. The neuroscientific study of meditation may thus provide important insights into the potential of mental training to strengthen cognitive control skills, and the factors that contribute to, and the mechanisms that underlie process-specific learning.

The paper is organized as follows. First, to demonstrate the ability of purely mental training to influence brain function and structure, and thereby, behavior, we will briefly review findings from studies of mental practice in the motor domain. We will then focus on mental training in the cognitive domain. To this end, we will introduce standard cognitive-based meditation practices, and discuss recent key findings to illustrate the usefulness of mental training as cultivated by these meditation practices in the neuroscientific study of brain and cognitive plasticity. Next, we will review characteristics of mental training contexts that may promote process-related learning from the vantage point of meditation. In this context, we will also discuss recent findings from studies using other complex training protocols in non-laboratory settings, such as action video gaming, that have reported learning that appears more general than once thought possible. Then, we will discuss methodological challenges that the neuroscientific study of meditation faces in particular. The paper will end with a discussion in which important avenues for future meditation research are identified.

## MENTAL TRAINING IN THE MOTOR DOMAIN

Research has convincingly shown that mental training using motor imagery, like physical practice, can produce changes in brain structure and function that are associated with improved subsequent performance of a motor skill (Guillot and Collet, 2005; Munzert et al., 2009). In a seminal study, mentioned above, Pascual-Leone et al. (1995) showed that mental practice alone is sufficient to promote the modulation of neural circuits involved in the early stages of motor skill learning. Two hours of mental practice of a finger tapping sequence on five separate days was associated with both an enlargement of the representations of the fingers in primary motor cortex, as well as improved motor performance. In another study, Ranganathan et al. (2004) found that strength gains following mental training were related to increases in the amplitude of scalp-recorded brain potentials over midline and lateral motor and midline frontal regions. They proposed that mental training may have affected higher-order motor cortical regions, such as supplementary motor and prefrontal regions, which may in turn have influenced primary motor areas. A recent study using the high spatial resolution of functional magnetic resonance imaging (fMRI) confirmed that gains in motor behavior after mental training are associated with changes in higher-order motor cortical regions (Nyberg et al., 2006). Mental training-related changes in neural activity were observed in the supplementary motor area, as well as in visual cortex. Thus, mental practice of a motor act has been shown to produce changes in brain function and improve performance. Not surprisingly, therefore, mental practice has long found wide acceptance in the training of athletes (e.g., Suinn, 1984).

An important question is *how* mental training using motor imagery may improve motor performance (Jeannerod, 1996). Behavioral studies have shown that there is a strong correlation between imagined and executed actions along various behavioral dimensions, suggesting that mental and physical training rely in part on common mechanisms (Jeannerod, 1994; Grush, 2004). For example, the time it takes to imagine an action is closely correlated with the execution time of the action (Decety et al., 1989). In line with the idea that the contents of the representations that are formed and strengthened through the two types of practice

overlap to some extent, neuroimaging research has shown that motor imagery and motor execution activate similar, although not identical, neural networks (e.g., Jeannerod, 1995; Decety, 1996; Lotze et al., 1999; Kosslyn et al., 2007; Munzert et al., 2009). Yet, there are important differences between mental and physical practice. Mental training using motor imagery heavily depends on cognitive aspects of action control, such as motor planning, as well as on working memory to transform, maintain, and inspect information about a motor act. In addition, in contrast to physical practice, although a motor plan is generated, it is prevented from operating on the body (Jeannerod, 1994; Grush, 2004). These differences between mental and motor training are substantiated by neuroimaging research showing that mental and motor training can affect distinct neural networks (Lacourse et al., 2005; Nyberg et al., 2006; Olsson et al., 2008).

It has been proposed that mental training may be more likely to develop and strengthen effector-independent representations, rather than effector-dependent representations, whereas for physical practice, the reverse may be true (Olsson et al., 2008). It is notable in this respect that several studies have reported benefits of mental training using motor imagery on tasks that are largely cognitive (Minas, 1978; Driskell et al., 1994). Of further importance, in some studies, mental practice (Wohldmann et al., 2008), as well as combined mental and motor practice (Olsson et al., 2008) has been associated with greater transfer of learning than physical practice, that is, better performance on novel tasks that call upon the trained skill. These data support the idea that mental practice using motor imagery may strengthen more abstract representations that do not involve specific effectors, thereby fostering process-specific learning that more easily generalizes across stimuli and tasks. Initial evidence for this idea comes from an fMRI study by Olsson et al. (2008), who showed facilitated transfer of learning after combined motor and mental training that was related to changes in the connections from both motor and cognitive systems to the cerebellum.

To conclude, mental training of motor skills leads to changes in brain circuitry and behavior, just as physical training produces changes in brain circuitry and behavior. Yet, the mechanisms underlying mental and physical training of motor skills are not identical. Notably, mental training may improve motor performance by strengthening more abstract representations, and thereby promote transfer of learning to novel task contexts.

## MENTAL TRAINING IN THE COGNITIVE DOMAIN

We will now turn to mental training that is specifically focused at enhancing cognitive function rather than at improving motor performance. Just like processes related to motor control, processes related to the control of cognitive processes should be amendable to mental training. As mentioned in the introduction, we focus here on cognitive control processes, as their amenability to training is especially valuable because these abilities are fundamental to higher cognition and contribute to performance across a wide range of cognitively demanding tasks (Smith and Jonides, 1999; Duncan and Owen, 2000). Yet, unfortunately, relatively little is known about the amenability of such processes to mental training. As mentioned above, this is likely due to the fact that it is difficult to design controlled mental training contexts. Notably, many meditation practices are based on precisely descriptive and highly detailed theories and are explicitly designed to train specific, well-defined cognitive control skills. The neuroscientific study of meditation may thus provide important insights into the usefulness of mental training as a tool in the study of brain and cognitive plasticity (see also Tang and Posner, 2009; Xiong and Doraiswamy, 2009).

Despite a large number of scientific reports and theoretical proposals, the neurophysiological processes involved in meditation and the long-term impact of meditation on the brain and cognitive function are still largely unknown. The lack of statistical evidence, control populations and tasks, and the heterogeneity of the studied meditative states can, in part, account for the limited contributions made by neuroscience-oriented research on meditation. Therefore, rather than providing an exhaustive review of this literature, we will only discuss recent key findings that speak to the idea that mental training of cognitive skills can induce process-specific learning. Other articles have recently reviewed this literature more exhaustively (see, e.g., Cahn and Polich, 2006; Lutz et al., 2007). Specifically, in the below, we will first define meditation, and review recent key findings from neuroimaging studies of meditation. Modern neuroimaging techniques provide a new possibility for understanding process-specific learning by revealing its underlying mechanisms. We will then identify aspects of mental training as cultivated by meditation that may foster process-specific learning.

### MEDITATION

The term "meditation" refers to a broad variety of practices, ranging from techniques designed to promote relaxation or improve attentional function to exercises performed with more far-reaching goals such as a heightened sense of well-being or the cultivation of altruistic behaviors. Not all practices thus focus on explicit training of specific cognitive skills, and it is therefore essential to be specific about the type of meditation practice under investigation. Here we focus on two common styles of Buddhist contemplative techniques, because (1) Buddhist traditions offer extensive, precisely descriptive and highly detailed theories about these practices in a manner that lends itself readily to appropriation into a neuroscientific context (Gyatso and Jinpa, 1995; Gunaratana, 2007; Lutz et al., 2007), and (2) these techniques are now being applied in clinical settings (e.g., Kabat-Zinn et al., 2000; Teasdale et al., 2000). These two meditation styles are often combined, whether in a single session or over the course of practitioner's training, and are explicitly designed to train specific cognitive processes. The first style, focused attention (FA) meditation, entails voluntary focusing attention on a chosen object in a sustained fashion, such as a visual object, a visualized image, or breath sensations (Lutz et al., 2007, 2008). To sustain this focus, the practitioner must also constantly monitor and regulate the quality of attention. Thus, FA meditation is thought to not only train one's ability to sustain attention, but also to develop three regulatory skills; the first is the monitoring faculty that remains vigilant to distractions without destabilizing the intended focus. The next skill is the ability to disengage from a distracting object without further involvement. The last involves the ability to redirect focus promptly to the chosen object. One may note remarkable parallels between the processes involved in FA meditation, as described above, and recent cognitive (neuro)science conceptualizations of attention. Western scientists also recognize that the ability to focus and sustain attention on an intended object requires skills involved

in monitoring the focus of attention and detecting distraction, disengaging attention from the source of distraction, and (re)directing and engaging attention to the intended object (e.g., Posner and Petersen, 1990).

The second style of meditation, open monitoring (OM) meditation, involves non-reactively monitoring the content of experience from moment-to-moment, without focusing on any explicit object (Lutz et al., 2007, 2008). OM meditation typically starts by calming the mind and reducing distractions using FA meditation. The practitioner then gradually reduces the focus on an explicit object in FA, and the monitoring faculty is correspondingly emphasized. Usually, there is also an increasing emphasis on cultivating a "reflexive" awareness that grants one greater access to the rich features of each experience, such as the degree of phenomenal intensity, the emotional tone, and the active cognitive schema. OM meditation is thought to enhance non-reactive meta-cognitive monitoring, as well as increase awareness of automatic cognitive and emotional interpretations of sensory, perceptual, and endogenous stimuli, and thereby cognitive flexibility and reappraisal (Bishop et al., 2004; Lutz et al., 2008; Chambers et al., 2009).

Both FA and OM meditation are assumed to induce a predictable and distinctive state (or set of states) whose occurrence is clearly indicated by certain cognitive and emotional features. These states, which arise during practice and are relatively short-term, can allegedly result in enduring changes in mental function, i.e., in the development of certain traits (Lutz et al., 2007, 2008). For example, FA meditation is said to develop, as one advances, the three regulative skills to the point that, for example, advanced practitioners have an especially acute ability to notice when the mind has wandered. Eventually, this induces a trait change, whereby the attention rests more readily and stably on the chosen focus. At the most advanced levels, the regulative skills are invoked less and less frequently, and the ability to sustain focus thus becomes progressively "effortless." It should be noted that this explicit focus on inducing a particular state of mind (or set of mental states) comprises an important difference between meditation practices, like FA and OM, and computerized task training of cognitive function. While computerized task training might also lead to changes in mental state (e.g., being focused), this is typically not its primary aim. To summarize, FA and OM practices are designed to train specific mental processes and such changes are thought to gradually endure through time, i.e., result in process-specific learning and the development of traits.

## NEUROIMAGING STUDIES OF MEDITATION

Although the systematic study of meditation is still in its infancy, recent findings suggest that experience in FA and OM meditation is associated with changes in brain and cognitive function both during meditation and during performance of external tasks that do not require meditation (Cahn and Polich, 2006; Lutz et al., 2007, 2008; Chambers et al., 2009). Neuroimaging studies examining the neural correlates of meditation indicate that task demands of the two meditation styles are reflected in characteristic neural activity patterns. For example, a recent cross-sectional study found that FA meditation on an external visual point compared to a rest condition was associated with greater activation in multiple brain regions, implicated in attentional processing, including the superior frontal sulcus and the intraparietal sulcus, in long-term meditation experts compared to novices (Brefczynski-Lewis et al., 2007). More importantly, neuroimaging work suggests that FA and OM meditation may not only during meditation activate brain regions involved in the psychological processes that are allegedly evoked, but may also produce lasting changes in brain and mental function that translate to improved performance in novel task contexts. For instance, several recent studies of FA meditation have reported improvements in sustained voluntary attention during performance of external tasks that do not require meditation. In one longitudinal study, in which participants were randomly assigned to a meditation-retreat group or a waiting-list control group, 5 h/day of meditation practice that involved voluntary focusing attention on a chosen object in a sustained fashion for 3 months was found to improve sustained visual attention (MacLean et al., 2010). Specifically, compared to the waiting-list control group, meditation-retreat participants showed enhanced perceptual discrimination and vigilance on a sustained visual attention task requiring participants to discriminate a rare threshold-level short line from a standard long line. Notably, these enhancements in sustained attention ability were still observed 3 months after the retreat ended, demonstrating enduring changes in sustained attention. In another recent longitudinal study, Lutz et al. (2009) found that 3 months of intensive FA and OM meditation was associated with enhanced moment-to-moment stability of attention during FA meditation, as measured both by reduced response time variability to rare target events in an auditory sustained attention task (**Figure 1A**) and increased trial-to-trial consistency of oscillatory neural responses over frontal scalp regions to these events as recorded with electroencephalography (EEG; **Figure 1B**). Furthermore, those individuals who showed the greatest training-induced increase in neural response consistency showed the largest decrease in behavioral response variability (**Figure 1C**). In addition, this study reported that intense mental training affected cortical engagement, as reflected by a concomitant reduction in event-related desynchronization (ERD) to target tones (**Figure 1D**). ERD is usually viewed as a correlate of increased cellular excitability in thalamocortical systems during cortical information processing (Pfurtscheller and da Silva, 1999). Within this framework, previous studies have interpreted reductions in ERD after practice of external tasks as decreased cognitive effort (Pfurtscheller and da Silva, 1999; Romero et al., 2008). A mental training-related reduction in task effort would be consistent with traditional accounts of progress in this practice (Gyatso and Jinpa, 1995; Gunaratana, 2007). In these accounts, the regulative attentional skills allegedly are invoked less and less frequently with more advanced level of training, and the ability to sustain focus is said to become progressively "effortless." Since these effects were observed over time within the same individual and during external task performance, together the above findings suggest that purely mental training of FA generalizes to improvements in performance on novel tasks that call upon the trained skills. These longitudinal findings also indicate that the enhanced sustained attention ability most likely reflects plasticity in the adult brain, and more generally, that attentional skills are subject to training.

Long-term practice of OM meditation is also thought to result in enduring changes in cognitive and brain function. Specifically, because OM meditation fosters non-reactive awareness of the stream of experience without deliberate selection of a primary

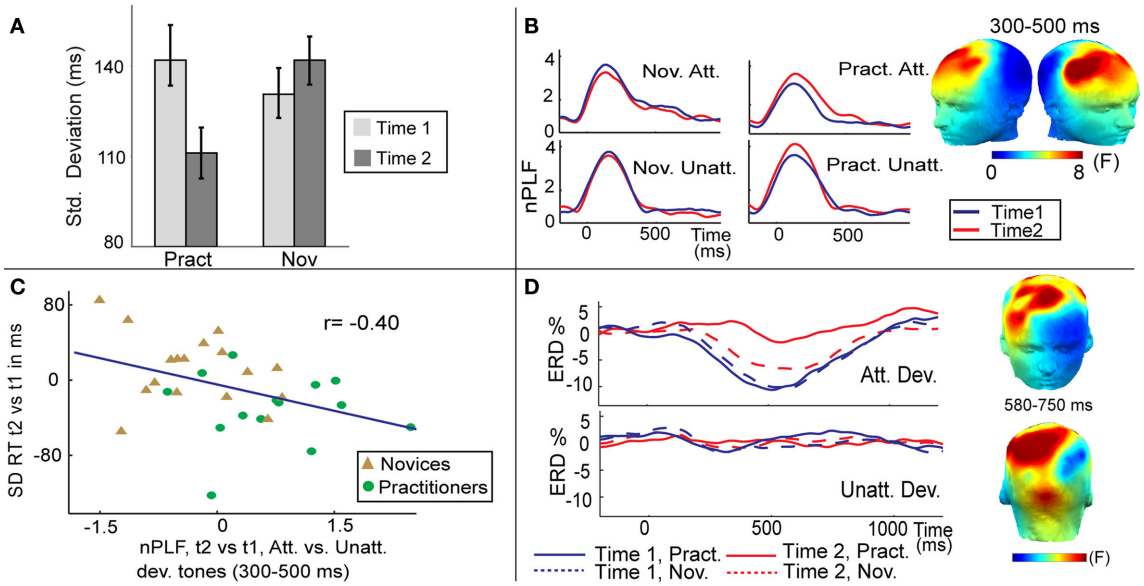



**FIGURE 1 | Effects of FA meditation. (A)** Intensive mental training reduced intra-individual variability of behavioral performance. Average SD of reaction time in response to target (attended deviant) tones, separately for each session (time 1, time 2) and group [practitioners (Pract), novices (Nov)]. **(B)** Intensive mental training increased trial-to-trial consistency of brain responses to attended deviant tones. The scalp maps show the spatial distribution of the mental training-related increase in theta-band (4–7 Hz) phase consistency (indexed by normalized PLF) to target tones, as indexed by a three-way interaction between group, session, and condition (attended, unattended deviant tones). *F* values are averaged between 300 and 500 ms. The time course of normalized PLF values averaged across the electrodes sites showing this significant three-way interaction is shown separately for each group, session, and attended (Att) and unattended (Unatt) deviant tones. Note that the observed increase in phase consistency to attended deviant tones over time was only observed for the practitioner group. **(C)** The correlation plot shows that the observed change in the trial-to-trial variability of brain responses predicted the observed behavioral change in the trial-to-trial variability of the RT. **(D)** Mental training selectively reduced cognitive effort as indexed by ERD. Time course of the ERD for attended and unattended deviant tones (Dev.) shown for the practitioners and novices separately. The scalp maps show the spatial distribution of the mental training-related decrease in beta-band (13–30 Hz) ERD to attended vs. unattended deviant tones [three-way interaction between group, session, and condition]. *F* values are averaged between 580 and 750 ms. Figure is adopted from Lutz et al. (2009).

object, intensive practice can be expected to reduce the elaborative thinking that would be stimulated by evaluating or interpreting a selected object (Cohen et al., 1996; Lutz et al., 2007, 2008). In line with this idea, Slagter et al. (2007, 2009) recently found that 3 months of intensive OM meditation reduced elaborative processing of the first of two target stimuli (T1 and T2) presented in a rapid stream of distracters, as indicated by a smaller T1-elicited P3b, a brain potential index of resource allocation (**Figure 2A**). Moreover, this reduction in resource allocation to T1 left neural resources more rapidly available for T2, as indexed by an increase in trial-to-trial phase consistency of neural oscillatory activity in the theta frequency range (**Figure 2B**), and was associated with improved detection of T2 when it followed T1 within half a second (**Figure 2C**). Because participants were not engaged in formal meditation during task performance in this study, these results provide support for the idea that one effect of an intensive training in OM meditation might be a general reduction in the propensity to "get stuck" on a target, as reflected in less elaborate stimulus processing, and the development of efficient mechanisms to engage and then disengage from target stimuli in response to task demands. OM meditation is also thought to cultivate a reflexive awareness of the subjective features of a given moment, such as its emotional tone, by engaging processes involved in interoception or awareness of bodily responses (Lutz et al., 2008). In line with this idea, a recent fMRI study found greater activity in network of brain areas involved in interoception, including the anterior insula and the somatosensory cortex, during a monitoring state relative to a narrative generation state in participants who were randomly assigned to an 8-week course incorporating OM meditation compared with a group of waitlist controls (Farb et al., 2007).

Next to these functional changes, several studies have also reported structural changes in the brain of long-term FA and OM meditators compared to novices in regions that are typically activated during these meditations (Lazar et al., 2005; Hölzel et al., 2008; Grant et al., 2010). For example, a study by Hölzel et al. (2008) found greater gray matter concentration in the anterior insula, a brain region, as mentioned above, important for awareness of internal experience, in experienced Vipassana meditators (mean practice: 8.6 years; 2 h daily). In another study, more than 1000 h of Zen meditation was associated with increased cortical thickness in the dorsal anterior cingulate, a region important in the adaptive control of behavior, and bilaterally, in the secondary somatosensory cortex. It is important that these cross-sectional findings are supplemented by findings from prospective studies to show that these structural changes result from meditation rather than, e.g., pre-existing individual differences. In addition, because both Vipassana and Zen meditation include forms of FA and OM meditation, future studies are necessary to determine how FA and OM meditation may each contribute to these anatomical changes.

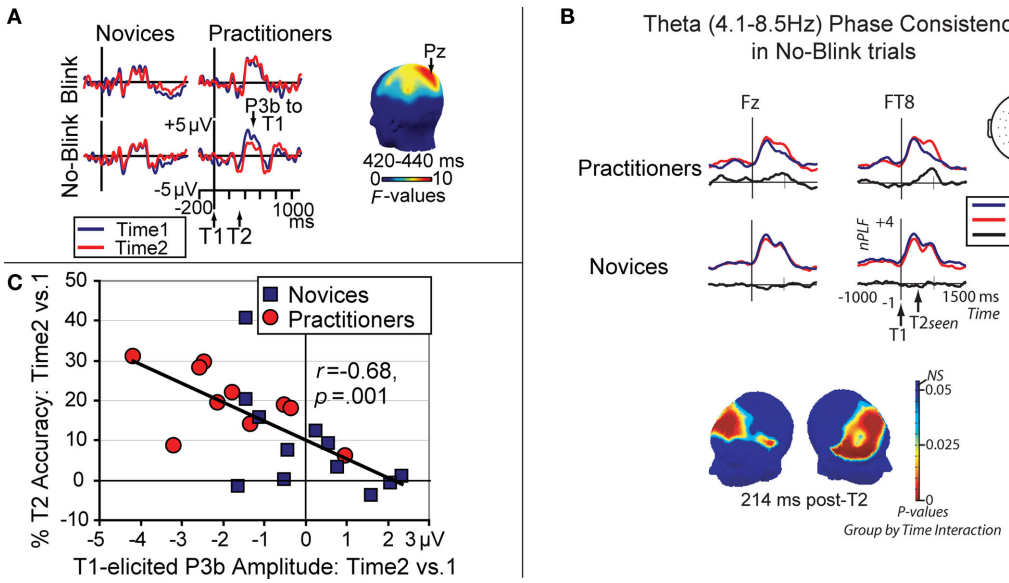



**FIGURE 2 | Effects of OM meditation. (A)** Brain potentials from electrode Pz, time-locked to T1 onset in short-interval trials as a function of session (time 2 vs. 1), T2 accuracy (no-blink vs. blink), and group (practitioners vs. novices). This figure panel illustrates that intensive OM meditation was associated with a selective reduction in T1-elicited P3b amplitude, a brain potential index of resource allocation, in no-blink trials in the practitioner group. The scalp map shows electrode sites where this three-way interaction between session, T2 accuracy, and group was significant between 420 and 440 ms. Adopted from Slagter et al. (2007); **(B)** Target locking of the theta frequency band phase at electrodes Fz and FT8, time-locked to T1 onset, shown for short-interval no-blink trials and separately for each session and group. Neural activity in the theta frequency band phase-locked robustly to consciously perceived target stimuli over frontal scalp regions. Importantly, a significant meditation-related increase in T2 phase consistency was observed over midline frontal and right lateral frontal and centro-parietal scalp regions (see scalp map). This increase in phase consistency over time was only observed for the practitioner group, indicating that intensive OM meditation may have reduced trial-to-trial variability in the recruitment of processes leading toward the conscious perception of T2. Adopted from Slagter et al. (2009); **(C)** Relationship between the observed change in brain resource allocation to T1, as indexed by T1-elicited P3b amplitude (for no-blink trials), and the corresponding the change in attentional blink magnitude over time. Note that those individuals that showed the largest decrease in T1-elicited P3b amplitude over time generally showed the largest increase in T2 accuracy over time. Adopted from Slagter et al. (2007); together, these data support the notion that the ability to accurately identify T2 depends upon the efficient processing of T1, and that OM meditation may reduce elaborate object processing.

To summarize, mental training of cognitive skills as cultivated by FA and OM meditation has been associated with changes in brain structure and function, as well as improved task performance. These findings provide initial support for the idea that systematic mental training of cognitive skills, as cultivated by meditation, can improve performance on external tasks that call upon the trained skills, and hence can strengthen specific cognitive processes. These findings also add to a growing body of literature demonstrating plasticity in the adult brain, and may provide initial insights into the basic mechanisms that underlie cognitive process-specific learning.

## FACTORS THAT DETERMINE PROCESS-SPECIFIC LEARNING

A long-standing issue in cognitive training is how to get training benefits to generalize to different stimuli and tasks (e.g., Schmidt and Bjork, 1992; Kramer and Willis, 2003; Green and Bavelier, 2008). As mentioned above, transfer of learning to novel task contexts has generally been the exception rather than the rule. Yet, purely mental training, as cultivated by meditation, as discussed above, appears to improve performance on novel, external tasks that call upon the trained skills but do not require meditation, and hence, induce process-specific learning. An important question is therefore what aspects of the meditation context may promote process-specific learning. In the below, we will discuss several factors that appear particularly important determinants of process-specific learning from the vantage point of mental training as cultivated by meditation. It should be noted that these factors are not specific to meditation training contexts, and likely represent common determinates of more generalized learning (e.g., Schmidt and Bjork, 1992; Ahissar and Hochstein, 2004; Green and Bavelier, 2008; Ahissar et al., 2009). Other non-laboratory training settings that appear to produce more general learning, such as action video gaming and musical training, also embody many of these characteristics (see Green and Bavelier, 2008 for an extensive review).

### COMPLEXITY OF THE TRAINING CONTEXT

From the above descriptions of FA and OM meditation, it is clear that each of these practices taps several cognitive skills in parallel. For example, FA meditation involves processes involved in directing and sustaining attention on a selected object (e.g., breath sensation), detecting mind wandering and distractors (e.g., thoughts), and disengagement of attention from distractors and shifting of attention back to the selected object. This is a key difference between mental training as cultivated by meditation and typical laboratory training settings that generally have been designed to train one specific cognitive skill, but not others. Like-wise, action video gaming, which has been shown to promote process-specific learning (e.g., Green and Bavelier, 2003, 2006; Donohue et al., 2010), naturally taps into many processes in parallel (Green and Bavelier, 2008).

Many action video games require players to simultaneously focus attention on task-relevant information, while ignoring distracting events, to monitor the environment and performance, and to track multiple moving objects. Thus, the complexity of training setting may be an important determinant of process-specific learning.

#### STIMULUS AND TASK VARIABILITY

Another important determinant of process-specific learning appears to be variability of input and task (Schmidt and Bjork, 1992). High variability at the level of the exemplars or movements to be learned, and the context in which they occur likely prevents learning at the level of specific stimuli/effectors, or stimulus-response relationships, thereby fostering learning at a higher or more abstract level of representation (Seidler, 2004; Green and Bavelier, 2008). It is notable in this respect that the observed structural brain changes in taxi drivers have not been observed in bus drivers, who operate along a constraint set of routes, unlike taxi drivers, who navigate widely around the city (Maguire et al., 2006). Of further importance, although many studies have reported lack of transfer of training benefits to new task contexts, recent studies of external task training in laboratory settings have reported improvements in cognitive task performance and transfer of learning to novel cognitive tasks when the training used a battery of tasks that all called upon the same (to-be-trained) cognitive skill, but involved different stimuli and task contexts (Olesen et al., 2004; Dahlin et al., 2008; Jaeggi et al., 2008; Persson and Reuter-Lorenz, 2008). For example, Dahlin et al. (2008) used six different training tasks, which all required working memory updating, a basic cognitive control skill. Five weeks of training resulted in enhanced performance not only on the trained task, but importantly also on a novel task, which required updating, but differed in terms of memorial content, set size, presentation rate, and response format. This transfer effect was mediated by the striatum, a subcortical brain structure important for updating. Thus, variability of input and task appears to be an important determinant of process-related learning and therefore, whether or not transfer of learning will occur. Mental training as cultivated by FA and OM meditation naturally includes many stimuli of various type and domain (e.g., auditory/somatosensory, cognitive/emotional, internal/external) that occur in different mental contexts. For example, OM meditation consists in being attentive moment by moment to *anything* that occurs in experience, whether it be a sensation, thought, or feeling. Such variability in stimulus and mental content may contribute to process-specific learning as a function of meditation.

#### TYPES OF PROCESSES TRAINED

It has been proposed that cognitive learning is most likely to occur when training programs focus on strengthening cognitive control functions that orchestrate thoughts and actions to make them consistent with internal goals (Persson and Reuter-Lorenz, 2008). As mentioned above, these include processes such as focusing attention, meta-cognitive monitoring, and switching between tasks. These processes are shared by many tasks and are widely viewed as amodal (Smith and Jonides, 1999; Duncan and Owen, 2000). Benefits of training programs focusing on such core cognitive skills may therefore more easily extend across materials and stimulus modalities, and hence, across tasks. Notably, FA and OM meditation are designed explicitly to train core cognitive skills (see above). This may provide another factor that fosters process-specific learning in meditation.

#### TASK DIFFICULTY

Learning is more likely to occur when the task remains challenging, albeit not too difficult, throughout the training (Ahissar and Hochstein, 2004). This is the case for example, in action video gaming, where the player advances from one level of difficulty to the next as expertise develops (Green and Bavelier, 2008). This is also true for meditation, where a practitioner is gradually introduced to more refined aspects of a technique while he or she gains more and more familiarity with the practice. Indeed, many contemplative traditions speak of ascending stages through which the practitioner passes, and a single practice may progress gradually through a number of meditative states, that might significantly differ from each other both phenomenally and in terms of the appropriate technique to be applied (Lutz et al., 2007). This progression ensures that the practice remains engaging and may also contribute to process-specific learning.

#### AROUSAL AND MOTIVATION

Arousal and motivation are important principles guiding learning. It is well known that training paradigms that lead to low levels of arousal will tend to lead to low amounts of learning, as will training paradigms that lead to excessively high levels of arousal (Frankenhaeuser and Gardell, 1976). Between these extremes is some level of arousal that leads to a maximum amount of learning (Green and Bavelier, 2008). Notably, central to both FA and OM meditation is maintenance of an optimal level of arousal or alertness via the regulation of attention or emotions (Lutz et al., 2007). For example, in most styles of FA meditation, after having placed the attention on the object, the practitioner seeks to avoid two overall "flaws" that represent under- and over-arousal. The first, dullness, is detected, in its most gross form, as a sensation of drowsiness. The second, excitement, manifests itself in distraction or attention wandering to other mental events. When suboptimal levels of arousal are detected, the practitioner can apply counteracting methods. For instance, to counteract drowsiness, the practitioner may add intensity to a visualized object or tension to the body. There are a variety of findings in the literature that already suggests that autonomic changes occur during FA and OM meditations. A frequent finding is a decrease in arousal during these practices (Cahn and Polich, 2006; Lutz et al., 2007). For instance, Takahashi et al. (2005) in a study of Zen meditation found changes in heart rate variability (reflecting parasympathetic nervous system activity) that were associated with changes in specific EEG oscillatory activity patterns. Yet, little is still known on the effect of FA and OM meditations on the maintenance of an optimal balance between excitement and dullness.

Like actively monitoring and regulating arousal levels, motivation is an inherent feature of many meditation practices. A formal meditation session will often begin and end with deliberately invocating some forms of soteriological or altruistic motivations (Gyatso and Jinpa, 1995; Gunaratana, 2007). Such deliberate practice, typically on a daily basis, is known to be a critical factor for the continued improvement of performance (Ericsson, 1996;

Ericsson and Lehmann, 1996). This feature, as well as the active monitoring and regulation of arousal levels, may be an additional aspect of meditation regimens important for inducing process-specific learning.

### DURATION OF TRAINING

In most laboratory manipulations cognitive training typically is no longer than several weeks, and total training duration in most previous neuroimaging studies of cognitive task training amounted to less than 20 h, even in studies demonstrating transfer of learning (e.g., Dahlin et al., 2008). This duration of training might be too short to induce long-lasting process-specific learning. Indeed, the attainment of an expert level of performance requires about 10,000 h of deliberate practice in the case of non-laboratory, domain-specific learning, such as performance of chess, surgery, music, and sports (Ericsson and Lehmann, 1996). Because the traditional duration of training in meditation also involves many hours of intensive practice over years (Lutz et al., 2007), the study of long-term meditators may provide a unique opportunity to investigate the development of process-specific learning. Reported differences in brain function between meditation experts (>10,000 h in life) vs. control subjects highlight the ability of meditation experts to generate new data that might not exist without such sustained mental training (Lutz et al., 2004; Carter et al., 2005; Brefczynski-Lewis et al., 2007). These data will need to be supplemented with data from studies that examine meditation-related changes over long time periods within the same individuals, from novice to expert. Such longitudinal data are necessary to exclude the possibility that observed training effects are due to pre-existing differences between groups (i.e., experts and novices) and will allow for a more precise delineation of the developmental trajectory of the trained abilities and the mechanisms that underlie process-specific learning. Thus, duration of training may represent a final determinant of process-specific learning.

Meditation has both short-term, intermediate, and long-term effects. Many studies have reported changes in cognitive task performance after relatively brief practice in the order of several days to weeks (e.g., Jha et al., 2007; Tang et al., 2007, 2010) or after a training period of several months (e.g., MacLean et al., 2010; Sahdra et al., in press). The extent to which these shorter-term changes are enduring and reflect lasting changes in brain function or structure is still unclear. Furthermore, the relationship between short- and long-term effects of meditation is at this point unknown, and longitudinal studies, that follow the same individual over time over longer time periods (i.e., years), are necessary to clarify this relationship.

## METHODOLOGICAL CHALLENGES

The problem of determining whether a treatment or intervention has an effect is ubiquitous in science, and the golden standard is to evaluate effects using randomized, double-blind experimental designs, in which participants are randomly assigned to either the experimental group or a control group, and neither the participants nor the experimenters know which group the participants are in. The neuroscientific study of meditation faces several methodological challenges in particular related to the fact that randomized assignment is not always possible, as in cross-sectional studies comparing meditation experts with non-meditators or less-experienced meditators, and the fact that it is impossible to blind participants to the nature of the study (in this case, meditation). Researchers should therefore be sensitive to potential confounding effects of motivation, demand characteristics, and outcome expectations on study outcomes. For example, participants with an interest or belief in meditation may be more likely to consent to participate and more motivated to follow the study procedure. They may also more strongly believe in positive study outcomes. It is well known that such non-specific factors can significantly influence study outcomes (e.g., Baskin et al., 2003).

One way to deal with these issues is to choose a control group with a clear conception of the research question and the hypothesized mechanism, and thorough consideration of how threats to validity may be best addressed for a given, clearly defined training procedure. For instance, a study may benefit from using an active control group that receives a structurally comparable, but different treatment than the meditation group, for example a physical exercise program (e.g., MacCoon et al., 2011). An active control group controls not only for factors related to the fact that participants cannot be blinded to the meditation intervention, such as expectancy effects, but also for other non-specific factors, such as social interactions, attention given by instructors, and time spend in the study. In studies of meditation to date, including most of the studies reviewed in this article, active control groups have been very rarely used. If we ultimately wish to attribute the changes observed in studies of meditation-based interventions to the active ingredient of meditation *per se* rather than the many non-specific factors, it is imperative to utilize active control groups that permit such a rigorous comparison. The state of research in this area is still in its infancy, but as the field moves forward, it will be increasingly important to use rigorous comparison groups to which participants are randomly assigned. Of course, in studies of long-term practitioners, this is not possible, but these studies need to be supplemented with longitudinal studies in less-experienced individuals where changes over time can be tracked (see below).

Next to choosing an appropriate control group, inclusion of a proper control task or task condition for which one does not expect to find meditation-related improvements, may also control for some of the above variables, in particular motivation and expectancy effects. For instance, based upon theoretical models of the attentional blink, we predicted that 3 months of Vipassana meditation, a form of OM meditation, (1) would selectively modulate *later* brain processes related to attentional resource allocation, but not early, sensory-driven brain processes (as indicated by early visual brain potential components), and (2) would only affect *T2 processing*, not T1 processing, in trials in which T2 followed T1 within the time-window of the attentional blink (in *short-interval*, but not long-interval trials; Slagter et al., 2007; see above). As predicted, we only found meditation-related improvements in T2 performance in short-interval trials that were associated with selective changes in higher-order T2 processing. Such a theoretically founded, selective effect is difficult to explain in terms of, for example, a placebo effect. Yet, it should be noted that alternate explanations, such as changes in working memory capacity, mood, or arousal, cannot be fully excluded in this study. For example, behavioral studies have reported a relationship between working memory capacity

and attentional blink size (Colzato et al., 2007), as well as between mood and arousal and the size of the attentional blink (Olivers and Nieuwenhuis, 2006; Jeffries et al., 2008). It is thus possible that OM meditation in this study affected the size of the attentional blink by modulating one of these factors next to, or rather than by training attentional processes itself. This example illustrates the need for future research to determine the precise mechanisms that underlie effects of FA and OM meditation on cognitive phenomena, such as the attentional blink, and their underlying brain circuitry. This research would benefit from including questionnaires and/or tests that measure changes in, e.g., mood and working memory capacity, as well as psychophysiological measures of mood and arousal, such as skin conductance and heart rate variability, to shed further light on the precise mechanisms that may underlie observed effects, and to better control for possible confounding variables.

To summarize, since individuals cannot be blinded with respect to the fact that they are participating in a meditation study, control groups and tasks are essential for validating the effects of meditation. Ideally, participants are randomly assigned to a practitioner and control group, and groups are matched in all aspects other than the factor of interest, most importantly in age, gender, race, socioeconomic status intelligence, motivation and study outcome expectancies. Questionnaires and psychophysiological measures may aid in interpreting observed effects of meditation.

#### CROSS-SECTIONAL VS. LONGITUDINAL STUDIES

There are two fundamental strategies for examining the effects of mental training, that each have distinct advantages and disadvantages: cross-sectional and longitudinal approaches. In the cross-sectional approach, individuals with varying levels of a given skill are compared and differences in some variables related to their skill level are identified. For example, neural function of long-term practitioners is compared with that of matched control subjects or less-experienced practitioners. This approach is useful for studying mental training-related changes over longer time periods (e.g., many years), where the longitudinal study of individuals may be difficult. Yet, as randomized group assignment is not possible in this approach, cross-sectional designs may suffer from cohort effects, in which different groups (i.e., expert vs. novice meditators) differ from each other by factors of no-interest, for instance in sleep or nutritional habits, or with respect to pre-existing differences in personality characteristics. For example, although lower neuroticism scores are reported by QiGong meditators, who have practiced for a greater number of years (Leung and Singhal, 2004), meditation training is more likely to be discontinued by people with higher trait neuroticism (Delmonte and Kenny, 1985; Delmonte, 1988).

The longitudinal approach studies the same individual at every point. This approach therefore does not necessarily suffer from cohort effects, and provides optimal power to identify training-related changes, because within-subject variability is typically smaller than across-subject variability. Yet, spending a number of months in a retreat environment, as is the case in many longitudinal meditation studies, including some of the studies discussed above (Slagter et al., 2007, 2009; Lutz et al., 2009; MacLean et al., 2010), also brings with it changes in for instance, sleep-wake cycle, mood, and arousal, and may require participants to practice in silence and with their eyes closed, depriving the senses of their usual stimulation. These confounding factors may well influence results. For example, it has been proposed that reduced sensory load does not only affect low-level sensory stimulus processing, for example, by increasing visual cortex excitability (e.g., Boroojerdi et al., 2000; Pitskel et al., 2007) or by modifying loudness perception (e.g., Formby et al., 2003; Munro and Blount, 2009), but can also lead to changes in higher-order cognitive information processing (e.g., Mahon and Caramazza, 2008). Most meditation practices are performed either with defocused open eyes, defocused half-open (hooded) eyes, or with closed eyes. It is possible that intensive meditation with defocused (half-) open eyes by reducing sensory load causes changes in the visual cortex similar to sensory deprivation. Notably, the longitudinal study by MacLean et al. (2010), discussed above, reported decreased visual perceptual threshold after 3 months of intensive meditation, and only found improved sustained attention task performance when retreat participants were allowed to use their pre-retreat visual stimuli. Improvements in sustained attention were not observed when perceptual thresholds were individually adjusted post-retreat to match each participant's pre-retreat perceptual threshold in difficulty. In this study, meditation-related reductions in sensory load may thus in part explain the observed changes in attentional function. The confound of reduced sensory load during some meditative practices is one that needs to be addressed in future studies. Neuroimaging measures may also be useful in this respect. For example, in the aforementioned Slagter et al. (2007) study, meditation was not associated with changes in early sensory processing (as indexed by the early visual brain potential components P1 and N1), suggesting that reduced sensory load likely did not contribute to the observed meditation-related changes in higher-order target processing in this study.

To summarize, meditation researchers should be circumspect of possible confounds related to the fact that individuals cannot be blinded to the meditation intervention, confounds that are associated with using cross-sectional and longitudinal study designs, and caveats, such as sensory deprivation or changes in sleep-wake cycle, that are more specific to the study of meditation. The above emphasizes the importance of conducting single-blind meditation studies. Those involved in data collection and analysis should be blinded with regard to group assignment and/or the specific study predictions to prevent experimenter biases (the favoring of certain outcomes over others) from influencing results. Of course, this is not possible in all cases, for example, when participants are studied in retreat settings.

It should be emphasized that many of the above challenges are not unique to the study of meditation. For example, one should always insure that effects of cognitive training cannot be explained by differences in motivation, arousal, mood, or expectancy bias between training and control groups, or by pre-existing differences between groups in age, intelligence, etc. Thus, in general, any cognitive training study should be designed with a clear conception of the research question and the hypothesized mechanism, and thorough consideration of how threats to validity may be best addressed for a given, clearly defined training procedure. Nevertheless, the development of adequate comparison groups against which to compare mental training as cultivated by meditation remains an important avenue for future research. Only when

done in a scientifically rigorous way can the study of meditation advance our understanding of the plasticity of cognitive processes and their underlying neural circuitry.

## DISCUSSION AND CONCLUSION

In this paper, we have reviewed the rationale for using mental training to study brain and cognitive plasticity. We have illustrated how the application of meditation in conjunction with neuroimaging methods has been used to shed light on the amenability of cognitive functions and their underlying brain circuitry to training. We have also identified several factors that may determine process-specific learning from the vantage point of mental training as cultivated by meditation, and discussed methodological challenges that the study of meditation needs to address. While the neuroscientific study of meditation is clearly still in its infancy, the initial findings reviewed here promise both to reveal the mechanisms by which such training may exert its effects and underscore the plasticity of the brain circuits that underlie complex mental functions. More generally, these findings support the idea that the nervous system is a continuously changing structure of which plasticity is an integral property and the obligatory consequence of not only sensory and motor processing, but also of more complex mental activities, such as focusing attention and meta-cognitive monitoring (Buonomano and Merzenich, 1998; Pascual-Leone et al., 2005). Just as a specific physical exercise will produce selected changes in brain circuitry and performance (Hillman et al., 2008), a specific mental exercise will lead to selected changes in brain circuitry that can significantly affect information processing and behavior.

Building upon previous work (Schmidt and Bjork, 1992; Ahissar and Hochstein, 2004; Green and Bavelier, 2008; Ahissar et al., 2009), we discussed several characteristics of meditation that may be responsible for process-specific learning, including the fact that typically multiple processes are trained in parallel (i.e., context complexity), variability in stimulus input, the manner in which task difficulty is progressed, the focus on maintaining an optimal level of arousal, the motivational state of the learner, and the duration of training. These features likely represent common principles of learning and transfer of learning, and many of them have previously been suggested to account for the more general learning observed after action video gaming, musical training, and athletic training (Green and Bavelier, 2008; Hillman et al., 2008). Like meditation, these more "natural" training regimens are exceedingly complex, tap many systems in parallel, and include variability in task and input (Green and Bavelier, 2008). Challenging and variable training thus appears important for inducing flexible learning that more easily generalizes to novel stimuli and task contexts. A key difference between meditation and these natural training regimens is that many meditation practices are explicitly designed to target specific, well-defined core cognitive processes. This too may contribute to the ability of mental training as cultivated by meditation to induce process-specific learning.

Future research should examine which specific aspects of a particular style of meditation actually produce changes in brain and cognitive function. As mentioned above, both FA and OM meditation tap several cognitive processes in parallel. It remains to be determined which mental training-related changes in cognitive processing actually produce, e.g., the observed differences in attention and brain function reported by some of the studies reviewed above. Future research should also be directed toward investigating the unique challenges that the studies on meditation practices present in designing appropriate controls. In addition, more research should be done on the "dose response" of meditation practices to determine what may be effective study durations and to help standardize training interventions.

While the neuroscientific study of meditation is still in its infancy, mainstream psychology and cognitive neuroscience will arguably be well served by engaging in a more open, but nonetheless critical and rigorous, examination of the findings from meditation studies. Such findings may help to determine the extent to which the adult brain is plastic or subject to change, identify the basic mechanisms that underlie process-specific learning, and could lead to further exploration of cognitive-neural systems that are resilient to damage, amenable to reorganization, and capable of improving efficiency of processing through training or pharmacological treatment.

## REFERENCES


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

Ahissar, M., Nahum, M., Nelken, I., and Hochstein, S. (2009). Reverse hierarchies and sensory learning. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 364, 285–299.

Bachman, J. C. (1961). Specificity vs. generality in learning and performing two large muscle motor tasks. *Res. Q.* 32, 3–11.

Baddeley, A. (1996). Exploring the central executive. *Q. J. Exp. Psychol.* 49, 5–28.

Baskin, T. W., Tierney, S. C., Minami, T., and Wampold, B. E. (2003). Establishing specificity in psychotherapy: a meta-analysis of structural equivalence of placebo controls. *J. Consult. Clin. Psychol.* 71, 973–979.

Bishop, S. R., Lau, M., Shapiro, S., Carlson, L., Anderson, N. D., Carmody, J., Segal, Z. V., Abbey, S., Speca, M., Velting, D., and Devins, G. (2004). Mindfulness: a proposed operational definition. *Clin. Psychol. Sci. Pract.* 11, 230–241.

Boroojerdi, B., Bushara, K. O., Corwell, B., Immisch, I., Battaglia, F., Muellbacher, W., and Cohen, L. G. (2000). Enhanced excitability of the human visual cortex induced by short-term light deprivation. *Cereb. Cortex* 10, 529–524.

Boyke, J., Driemeyer, J., Gaser, C., Buechel, C., and May, A. (2008). Training-induced brain structure changes in the elderly. *J. Neurosci.* 28, 7031–7035.

Brefczynski-Lewis, J. A., Lutz, A., Schaefer, H. S., Levinson, D. B., and Davidson, R. J. (2007). Neural correlates of attentional expertise in long-term meditation practitioners. *Proc. Natl. Acad. Sci. U.S.A.* 104, 11483–11488.

Buonomano, D. V., and Merzenich, M. M. (1998). Cortical plasticity: from synapses to maps. *Annu. Rev. Neurosci.* 21, 149–186.

Cahn, B. R., and Polich, J. (2006). Meditation states and traits: EEG, ERP, and neuroimaging studies. *Psychol. Bull.* 132, 180–211.

Carter, O. L., Presti, D. E., Callistemon, C., Ungerer, Y., Liu, G. B., and Pettigrew, J. D. (2005). Meditation alters perceptual rivalry in Tibetan Buddhist monks. *Curr. Biol.* 15, R412–R413.

Chambers, R., Gullone, E., and Allen, N. B. (2009). Mindful emotion regulation: an integrative review. *Clin. Psychol. Rev.* 29, 560–572.

Cohen, J. D., Braver, T. S., and O'Reilly, R. C. (1996). A computational approach to prefrontal cortex, cognitive control and schizophrenia: recent developments and current challenges. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 351, 1515–1527.

Colzato, L. S., Spape, M., Pannebakker, M. M., and Hommel, B. (2007). Working memory and the attentional blink: blink size is predicted by individual differences in operation span. *Psychon. Bull.* 14, 1051–1057.

Dahlin, E., Neely, A. S., Larsson, A., Backman, L., and Nyberg, L. (2008). Transfer of learning after updating training mediated by the striatum. *Science* 320, 1510–1512.

Decety, J. (1996). Neural representations for action. *Rev. Neurosci.* 7, 285–297.

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

Delmonte, M. (1988). Personality correlates of meditation practice and frequency in an outpatient population. *J. Behav. Med.* 11, 593–597.

Delmonte, M., and Kenny, V. (1985). Models of meditation. *Br. J. Psychother.* 1, 197–214.

Donohue, S. E., Woldorff, M. G., and Mitroff, S. R. (2010). Video game players show more precise multisensory temporal processing abilities. *Atten. Percept. Psychophys.* 72, 1120–1129.

Draganski, B., Gaser, C., Busch, V., Schuierer, G., Bogdahn, U., and May, A. (2004). Neuroplasticity: changes in grey matter induced by training – newly honed juggling skills show up as a transient feature on a brain-imaging scan. *Nature* 427, 311–312.

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

Duncan, J., and Owen, A. M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends Neurosci.* 23, 475–483.

Ericsson, K. A. (1996). The role of deliberate practice in the acquisition and maintenance of expert performance. *Int. J. Psychol.* 31, 4661.

Ericsson, K. A., and Lehmann, A. C. (1996). Expert and exceptional performance: evidence of maximal adaptation to task constraints. *Annu. Rev. Psychol.* 47, 273–305.

Farb, N. A., Segal, Z. V., Mayberg, H., Bean, J., McKeon, D., Fatima, Z., and Anderson, A. K. (2007). Attending to the present: mindfulness meditation reveals distinct neural modes of self-reference. *Soc. Cogn. Affect. Neurosci.* 2, 313–322.

Formby, C., Sherlock, L. P., and Gold, S. L. (2003). Adaptive plasticity of loudness induced by chronic attenuation and enhancement of the acoustic background. *J. Acoust. Soc. Am.* 114, 55–58.

Frankenhaeuser, M., and Gardell, B. (1976). Underload and overload in working life – outline of a multidisciplinary approach. *J. Human Stress* 2, 35–46.

Fuster, J. M. (1997). *The Prefrontal Cortex-Anatomy Physiology, Neuropsychology of the Frontal Lobe*, 3rd Edn. Philadelphia: Lippincott-Raven.

Grant, J. A., Courtemanche, J., Duerden, E. G., Duncan, G. H., and Rainville, P. (2010). Cortical thickness and pain sensitivity in zen meditators. *Emotion* 10, 43–53.

Green, C. S., and Bavelier, D. (2003). Action video game modifies visual selective attention. *Nature* 423, 534–537.

Green, C. S., and Bavelier, D. (2006). Effect of action video games on the spatial distribution of visuospatial attention. *J. Exp. Psychol. Hum. Percept. Perform.* 32, 1465–1478.

Green, C. S., and Bavelier, D. (2008). Exercising your brain: a review of human brain plasticity and training-induced learning. *Psychol. Aging* 23, 692–701.

Grush, R. (2004). The emulation theory of representation: motor control, imagery, and perception. *Behav. Brain Sci.* 27, 377–396.

Guillot, A., and Collet, C. (2005). Contribution from neurophysiological and psychological methods to the study of motor imagery. *Brain Res. Rev.* 50, 387–397.

Gunaratana, H. (2007). *Mindfulness in Plain English*. Somerville: Wisdom Publications.

Gyatso, T., and Jinpa, T. (1995). *The World of Tibetan Buddhism: an Overview of its Philosophy and Practice*. Somerville: Wisdom Publications.

Hillman, C. H., Erickson, K. I., and Kramer, A. F. (2008). Be smart, exercise your heart: exercise effects on brain and cognition. *Nat. Rev. Neurosci.* 9, 58–65.

Hölzel, B. K., Ott, U., Gard, T., Hempel, H., Weygandt, M., Morgen, K., and Vaitl, D. (2008). Investigation of mindfulness meditation practitioners with voxel-based morphometry. *Soc. Cogn. Affect. Neurosci.* 3, 55–61.

Jaeggi, S. M., Buschkuehl, M., Jonides, J., and Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. *Proc. Natl. Acad. Sci. U.S.A.* 105, 6829–6833.

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. (1996). Motor representations: one or many? *Behav. Brain Sci.* 19, 763–765.

Jeffries, L. N., Smilek, D., Eich, E., and Enns, J. T. (2008). Emotional valence and arousal interact in attentional control. *Psychol. Sci.* 19, 290–295.

Jha, A. P., Krompinger, J., and Baime, M. J. (2007). Mindfulness training modifies subsystems of attention. *Cogn. Affect. Behav. Neurosci.* 7, 109–119.

Kabat-Zinn, J., Lipworth, L., and Burney, R. (2000). The clinical use of mindfulness meditation for the self-regulation of chronic pain. *J. Behav. Med.* 8, 163–190.

Kosslyn, S. M., Ganis, G., and Thompson, W. L. (2007). "Mental imagery and the human brain," in *Progress in Psychological Science Around the World, Vol. 1, Neural, Cognitive and Developmental Issues*, eds. Q. Jing, M. R. Rosenzweig, G. d'Ydewalle, H. Zhang, H.-C. Chen, and K. Zhang (New York: NY Psychology Press), 195–209.

Kramer, R. F., and Willis, S. L. (2003). "Cognitive plasticity and aging," in *Psychology of Learning and Motivation: Advances in Research and Theory*, Vol. 43, ed. B. H. Ross (Amsterdam: Academic Press), 267–302.

Lacourse, M. G., Orr, E. L. R., Cramer, S. C., and Cohen, M. J. (2005). Brain activation during execution and motor imagery of novel and skilled sequential hand movements. *Neuroimage* 27, 505–519.

Lazar, S. W., Kerr, C. E., Wasserman, R. H., Gray, J. R., Greve, D. N., Treadway, M. T., McGarvey, M., Quinn, B. T., Dusek, J. A., Benson, H., Rauch, S. L., Moore, C. I., and Fischl, B. (2005). Meditation experience is associated with increased cortical thickness. *Neuroreport* 16, 1893–1897.

Leung, Y, and Singhal, A. (2004). An examination of the relationship between qigong meditation and personality. *Soc. Behav. Pers.* 32, 313–320.

Lotze, M., Montoya, P., Erb, M., Hülsmann, E., Flor, H., Klose, U., Birbaumer, N., and Grodd, W. (1999). Activation of cortical and cerebellar motor areas during executed and imagined hand movements: an fMRI study. *J. Cogn. Neurosci.* 11, 491–501.

Lutz, A., Dunne, J. D., and Davidson, R. J. (2007). "Meditation and the neuroscience of consciousness: an introduction," in *Cambridge Handbook of Consciousness*, eds P. Zelazo, M. Moscovitch, and E. Thompson (New York: Cambridge University Press), 81–109.

Lutz, A., Slagter, H. A., Rawlings, N. B., Francis, A. D., Greischar, L. L., and Davidson, R. J. (2009). Mental training enhances attentional stability: neural and behavioral evidence. *J. Neurosci.* 29, 13418–13427.

Lutz, A., Greischar, L. L., Rawlings, N. B., Ricard, M., and Davidson, R. J. (2004). Long-term meditators self-induce high-amplitude gamma synchrony during mental practice. *Proc. Natl. Acad. Sci. U.S.A.* 101, 16369–16373.

Lutz, A., Slagter, H. A., Dunne, J. D., and Davidson, R. J. (2008). Attention regulation and monitoring in meditation. *Trends Cogn. Sci.* 12, 163–169.

MacCoon, D. G., Sullivan, J. C., Davidson, R. J., Stoney, C. M., Johnson, L. L., Christmas, P. D., and Lutz, A. (2011). Health-Enhancement Program (HEP) Guidelines. Available at: http://digital.library.wisc.edu/1793/28198.

MacLean, K. A., Ferrer, E., Aichele, S. R., Bridwell, D. A., Zanesco, A. P., Jacobs, T. L., King, B. G., Rosenberg, E. L., Sahdra, B. K., Shaver, P. R., Wallace, B. A., Mangun, G. R., and Saron, C. D. (2010). Intensive meditation training improves perceptual discrimination and sustained attention. *Psychol. Sci.* 21, 829–839.

Maguire, E. A., Gadian, D. G., Johnsrude, I. S., Good, C. D., Ashburner, J., Frackowiak, R. S., and Frith, C. D. (2000). Navigation-related structural change in the hippocampi of taxi drivers. *Proc. Natl. Acad. Sci. U.S.A.* 97, 4398–4403.

Maguire, E. A., Woollett, K., and Spiers, H. J. (2006). London taxi drivers and bus drivers: a structural MRI and neuropsychological analysis. *Hippocampus* 16, 1091–1101.

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.* 102, 59–70.

Minas, S. C. (1978). Mental practice of a complex perceptual motor skill. *J. Hum. Mov. Stud.* 4, 102–109.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., and Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: a latent variable analysis. *Cogn. Psychol.* 41, 49–100.

Munro, K. J., and Blount, J. (2009). Adaptive plasticity in brainstem of adult listeners following earplug-induced deprivation. *J. Acoust. Soc. Am.* 126, 568–571.

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.

Nyberg, L., Eriksson, J., Larsson, A., and Marklund, P. (2006). Learning by doing versus learning by thinking: an fMRI study of motor, and mental training. *Neuropsychologia* 44, 711–717.

Olesen, P. J., Westerberg, H., and Klingberg, T. (2004). Increased prefrontal and parietal activity after training of working memory. *Nat. Neurosci.* 7, 75–79.

Olivers, C. N., and Nieuwenhuis, S. (2006). The beneficial effects of additional task load positive affect and instruction on the attentional blink. *J. Exp. Psychol. Hum. Percept. Perform.* 32, 364–379.

Olsson, C. J., Jonsson, B., and Nyberg, L. (2008). Learning by doing and learning by thinking: an fMRI study of combining motor and mental training. *Front. Hum. Neurosci.* 2:5. doi: 10.3389/neuro.09.005.2008

Pascual-Leone, A., Amedi, A., Fregni, F., and Merabet, L. B. (2005). The plastic human brain cortex. *Annu. Rev. Neurosci.* 28, 377–401.

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.

Persson, J., and Reuter-Lorenz, P. A. (2008). Gaining control training executive function and far transfer of the ability to resolve interference. *Psychol. Sci.* 19, 881–888.

Pfurtscheller, G., and da Silva, F. H. L. (1999). Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clin. Neurophysiol.* 110, 1842–1857.

Pitskel, N. B., Merabet, L. B., Ramos-Estabanez, C., Kauffman, T., and Pascual-Leone, A. (2007). Time-dependent changes in cortical excitability after prolonged visual deprivation. *Neuroreport* 18, 1703–1707.

Posner, M. I., and Petersen, S. E. (1990). The attention system of the human brain. *Annu. Rev. Neurosci.* 13, 25–42.

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.

Roelfsema, P. R., van Ooyen, A., and Watanabe, T. (2010). Perceptual learning rules based on reinforcers and attention. *Trends Cogn. Sci.* 14, 64–71.

Romero, S. G., McFarland, D. J., Faust, R., Farrell, L., and Cacace, A. T. (2008). Electrophysiological markers of skill-related neuroplasticity. *Biol. Psychol.* 78, 221–230.

Sahdra, B. K., MacLean, K. A., Ferrer, E., Shaver, P. R., Rosenberg, E. L., Jacobs, T. L., Zanesco, A. P., Aichele, S. R., King, B. G., Bridwell, D. A., Lavy, S., Mangun, G. R., Wallace, A. B., and Saron, C. D. (in press). Enhanced response inhibition during intensive meditation training predicts improvements in self-reported adaptive socio-emotional functioning. *Emotion*.

Salomon, G., and Perkins, D. N. (1989). Rocky roads to transfer – rethinking mechanisms of a neglected phenomenon. *Educ. Psychol.* 24, 113–142.

Schmidt, R. A., and Bjork, R. A. (1992). New conceptualizations of practice – common principles in 3 paradigms suggest new concepts for training. *Psychol. Sci.* 3, 207–217.

Seidler, R. D. (2004). Multiple motor learning experiences enhance motor adaptability. *J. Cogn. Neurosci.* 16, 65–73.

Slagter, H. A., Lutz, A., Greischar, L. L., Francis, A. D., Nieuwenhuis, S., Davis, J. M., and Davidson, R. J. (2007). Mental training affects distribution of limited brain resources. *PLoS Biol.* 5, 1228–1235. doi: 10.1371/journal.pbio.0050138

Slagter, H. A., Lutz, A., Greischar, L. L., Nieuwenhuis, S., and Davidson, R. J. (2009). Theta phase synchrony and conscious target perception: impact of intensive mental training. *J. Cogn. Neurosci.* 21, 1536–1549.

Smith, E. E., and Jonides, J. (1999). Storage and executive processes in the frontal lobes. *Science* 283, 1657–1661.

Suinn, R. M. (1984). "Imagery and sports," in *Cognitive Sport Psychology*, eds W. Straub and J. Williams (Lansing, NY: Sport Science Associates).

Takahashi, T., Murata, T., Hamada, T., Omori, M., Kosaka, H., Kikuchi, M., Yoshida, H., and Wada, Y. (2005). Changes in EEG and autonomic nervous activity during meditation and their association with personality traits. *Int. J. Psychophysiol.* 55, 199–207.

Tang, Y. Y., Lu, Q., Geng, X., Stein, E. A., Yang, Y., and Posner, M. I. (2010). Short-term meditation induces white matter changes in the anterior cingulate. *Proc. Natl. Acad. Sci. U.S.A.* 107, 15649–15652.

Tang, Y. Y., Ma, Y., Wang, J., Fan, Y., Feng, S., Lu, Q., Yu, Q., Sui, D., Rothbart, M. K., Fan, M., and Posner, M. I. (2007). Short-term meditation training improves attention and self-regulation. *Proc. Natl. Acad. Sci. U.S.A.* 104, 17152–17156.

Tang, Y. Y., and Posner, M. I. (2009). Attention training and attention state training. *Trends Cogn. Sci.* 13, 222–227.

Teasdale, J. D., Segal, Z. V., Williams, J. M., Ridgeway, V. A., Soulsby, J. M., and Lau, M. A. (2000). Prevention of relapse/recurrence in major depression by mindfulness-based cognitive therapy. *J. Consult. Clin. Psychol.* 68, 615–623.

Wohldmann, E. L., Healy, A. F., and Bourne, L. E. (2008). A mental practice superiority effect: less retroactive interference and more transfer than physical practice. *J. Exp. Psychol. Learn. Mem. Cogn.* 34, 823–833.

Woollett, K., and Maguire, E. A. (2009). Navigational expertise may compromise anterograde associative memory. *Neuropsychologia* 47, 1088–1095.

Xiong, G. L., and Doraiswamy, P. M. (2009). Does meditation enhance cognition and brain plasticity? *Ann. N. Y. Acad. Sci.* 1172, 63–69.

**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 July 2010; accepted: 26 January 2011; published online: 10 February 2011. Citation: Slagter HA, Davidson RJ and Lutz A (2011) Mental training as a tool in the neuroscientific study of brain and cognitive plasticity. Front. Hum. Neurosci.* ***5****:17. doi:10.3389/fnhum.2011.00017*

frontiers in
HUMAN NEUROSCIENCE

REVIEW ARTICLE
published: 18 March 2011
doi: 10.3389/fnhum.2011.00028

# Statistical analysis of fMRI time-series: a critical review of the GLM approach

***Martin M. Monti****

*Department of Psychology, University of California, Los Angeles, CA, USA*

***Edited by:***
*Michael X. Cohen, University of Amsterdam, Netherlands*

***Reviewed by:***
*Russell A. Poldrack, University of California, USA*
*Sean L. Simpson, Wake Forest University, USA*

****Correspondence:***
*Martin M. Monti, Department of Psychology, University of California, Los Angeles, CA 90095-1563, USA.*
*e-mail: monti@psych.ucla.edu*

Functional magnetic resonance imaging (fMRI) is one of the most widely used tools to study the neural underpinnings of human cognition. Standard analysis of fMRI data relies on a general linear model (GLM) approach to separate stimulus induced signals from noise. Crucially, this approach relies on a number of assumptions about the data which, for inferences to be valid, must be met. The current paper reviews the GLM approach to analysis of fMRI time-series, focusing in particular on the degree to which such data abides by the assumptions of the GLM framework, and on the methods that have been developed to correct for any violation of those assumptions. Rather than biasing estimates of effect size, the major consequence of non-conformity to the assumptions is to introduce bias into estimates of the variance, thus affecting test statistics, power, and false positive rates. Furthermore, this bias can have pervasive effects on both individual subject and group-level statistics, potentially yielding qualitatively different results across replications, especially after the thresholding procedures commonly used for inference-making.

**Keywords: functional magnetic resonance imaging, blood oxygenation level-dependent, general linear model, ordinary least squares, autocorrelation, multicollinearity, fixed effects, mixed effects**

## INTRODUCTION

Over the past 20 years the study of human cognition has benefited greatly from innovations in magnetic resonance imaging, in particular the development of techniques to detect physiological markers of neural activity. The most widely used of these techniques capitalizes on the changes in blood flow and oxygenation associated with neural activity (the hemodynamic response), and on the differing magnetic properties of oxygenated and deoxygenated blood. The paramagnetic properties of deoxyhemoglobin (dHb) create local field inhomogeneities, leading to reduced transverse (T2) relaxation times and therefore to a reduction in image intensity. Conversely, increased concentrations of oxyhemoglobin produce increased T2 relaxation times and a relative increase in image intensity. This blood oxygen level-dependent (BOLD) contrast mechanism forms the basis of functional magnetic resonance imaging (fMRI; Ogawa et al., 1990, 1992; Kwong et al., 1992).

While the idea of a hemodynamic response spatially localized to sites of neural activity dates back to Roy and Sherington (1890), the mechanisms by which neural activity triggers changes in cerebral blood volume, flow, and oxygenation are still not fully understood (c.f., Logothetis et al., 2001; Logothetis, 2002) posing a significant constraint on the interpretation of fMRI studies of cognition (Logothetis and Wandell, 2004; Logothetis, 2008). At the level of measurement, the degree to which changes in BOLD signal co-localize with neural activity depends on various imaging parameters, including the magnetic field strength and the imaging sequence used (Ugurbil et al., 2003; Logothetis, 2008). On a more fundamental physiological level, it is also unclear which aspects of neural activity are most closely linked to the hemodynamic response (see Logothetis, 2008, for an excellent overview). While activation of excitatory neurons has been shown to trigger changes in local blood flow (Lee et al., 2010), other studies have reported that input to and activity within local neuronal circuits are both better predictors of the hemodynamic response than output from pyramidal cells (e.g., Logothetis et al., 2001; Goense and Logothetis, 2008), suggesting that the hemodynamic response does not simply reflect the level of spiking activity. Indeed, in some cases the hemodynamic response has been observed in the absence of spiking output (Logothetis et al., 2001; Thomsen et al., 2004; Goense and Logothetis, 2008). A further complication is that both excitatory and inhibitory input create metabolic demands (Buzsáki et al., 2007), making it even harder to interpret the concomitant vascular response as a simple measure of neural firing rate (Logothetis, 2008; although in some cases, neuronal inhibition may produce metabolic and hemodynamic down-regulation; Stefanovic et al., 2004). Overall, the hemodynamic response is likely to reflect an average response to a range of metabolic demands imposed by neural activity, including both excitatory and inhibitory post-synaptic processing, neuronal spiking, as well as neuromodulation (Logothetis, 2008; Palmer, 2010).

Bearing these caveats in mind, the reminder of the paper will focus on statistical methods for separating noise from systematic fluctuations of the BOLD signal induced by experimental stimulation. Following a brief review of the general linear model (GLM) framework, the paper will focus on the degree by which single-subject fMRI time-series conform to the assumptions of the framework, and on the approaches used to mitigate infringements of these assumptions. Finally, the paper will also discuss methods to combine datasets from multiple subjects, with respect to their inferential scope and validity.

For clarity, the rest of the manuscript will use the term "volume" to refer to an individual data acquisition point, typically a three dimensional image of the MRI signal at multiple points throughout

the brain acquired over the course of several seconds. Multiple volumes acquired as one continuous stream of data are referred to as a "run" or "scan." Each subject usually undergoes multiple runs, sometimes with brief interruptions in between, but usually whilst remaining inside the bore of the magnet. A set of multiple runs is referred to as a "session." A standard fMRI dataset for a complete experiment usually consists of one (or more) session from each of a number of subjects.

## SINGLE-SUBJECT ANALYSIS (I): THE GENERAL LINEAR MODEL APPROACH

An fMRI dataset, can be seen as a set of cuboid elements (i.e., voxels) of variable dimension, each of which has an associated time-series of as many time-points as volumes acquired per session. The aim of a (conventional) statistical analysis is to determine which voxels have a time-course that correlates with some known pattern of stimulation or experimental manipulation. The first step in fMRI data analysis is to apply a series of "pre-processing" transformations with the aim of correcting for several potential artifacts introduced at data acquisition. Each transformation can be applied as required depending on the specific experimental design or acquisition protocol. The most typical steps include adjusting for differences in the acquisition time of individual image slices, correction for subject motion, warping individual subjects data into a common space ("normalization"), and temporal and spatial smoothing (see Jezzard et al., 2002). Following pre-processing, data analysis is often carried in two steps: a separate first-level analysis of data from each individual subject, followed by a second-level analysis in which results from multiple subjects are combined.

In the GLM approach, the time-course associated with each voxel is modeled as a weighted sum of one or more known predictor variables (e.g., the onset and offset of an experimental condition) plus an error term. The aim of the analysis is to estimate if, and to what extent, each predictor contributes to the variability observed in the voxel's time-course. Consider, for example, an experiment in which the BOLD response, $y$, is sampled $n$ times (i.e., volumes). The intensity of the BOLD signal at each observation ($y_i$) can be modeled as the sum of a number of known predictor variables ($x_1 \ldots x_p$) each scaled by a parameter ($\beta$):

$$y_1 = x_{1,1}\beta_1 + x_{1,2}\beta_2 + \quad + x_{1,p}\beta_p + \varepsilon_1$$
$$y_2 = x_{2,1}\beta_1 + x_{2,2}\beta_2 + \quad + x_{2,p}\beta_p + \varepsilon_2$$
$$\ldots$$
$$y_n = x_{n,1}\beta_1 + x_{n,2}\beta_2 + \quad + x_{n,p}\beta_p + \varepsilon_n$$

The aim of the first-level statistical analysis is to determine how large the contribution of each predictor variable $x_i$ is to the observed values of $y$. That is to say, how large each scaling parameter $\beta_i$ is, and whether it is significantly different from zero. Using the more compact matrix notation, the GLM can be re-expressed, in its simplest formulation, as:

$$Y = X\beta + \varepsilon \quad (1)$$

Where $Y$ is an $n \times 1$ column vector (i.e., $n$ rows, 1 column) representing the BOLD signal time-series associated with a single voxel. $X$ is the $n \times p$ design matrix, with each column representing a different predictor variable. Of interest are the columns representing manipulations or experimental conditions, although the matrix typically also includes regressors of non-interest, modeling nuisance variables such as low-frequency drifts and motion. $\beta$ is the $p \times 1$ vector of unknown weights setting the magnitude and direction of the (unique) association between each given predictor variable and the data $Y$. Finally, $\varepsilon$ is an $n \times 1$ vector containing the error values associated with each observation (i.e., the value of each observation that is not explained by the weighted sum of predictor variables).

**Figure 1** depicts the GLM model for an imaginary voxel with associated time-series $Y$, as the linear combination of three regressors of interest (e.g., tasks A,B,C) and a number of nuisance variables (here six motion regressors and a linear drift), each scaled by a vector of unknown amplitudes ($\beta$), plus an error term $\varepsilon$.

The standard approach to fMRI analysis is to fit the same model independently to the time-course of each voxel. Spatial covariance between neighboring voxels is thus typically ignored at the model fitting stage. The presence of more response variables (i.e., voxels) than observations (i.e., volumes), together with the aim of making topographically specific claims about BOLD activity, has traditionally motivated this "massive-univariate" approach. Recently, however, a lot of effort has gone into developing multivariate techniques to address the question of what information specific brain areas represent (as opposed to the "localizationalist" approach typical of univariate analysis; c.f., Kriegeskorte et al., 2006, 2008; Bowman, 2007; Bowman et al., 2007).

Several methods are available to estimate the value of the unknown parameter $\beta$, and therefore assess whether a given predictor variable significantly explains some portion of the variance observed in a voxel's time-course, including the ordinary least squares (OLS), (feasible) generalized least squares (GLS), and the so-called Smoothing and "Sandwich" approaches (see Waldorp, 2009, for a clear overview of these methods). In its simplest OLS form, the optimal parameters are defined as those

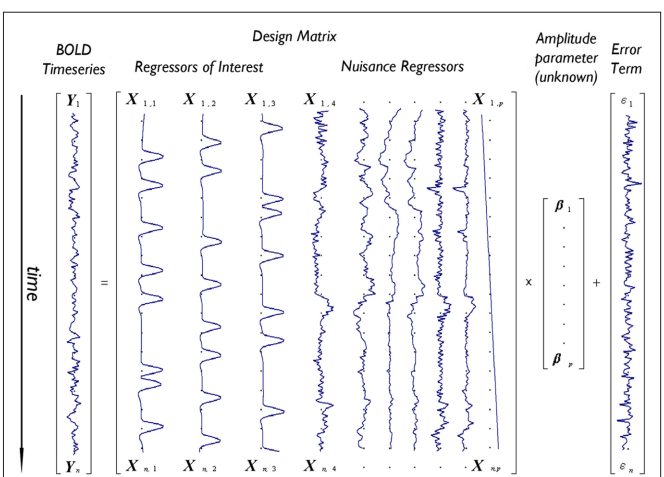



**FIGURE 1 | Depiction of the GLM model for an imaginary voxel with time-series *Y* predicted by a design matrix *X* including 10 effects (three regressors of interest – e.g., tasks A,B,C – and seven nuisance regressors – e.g., six motion parameters and one linear drift) of unknown amplitude $\beta_p$ and an error term.**

that minimize the sum of squared residuals[1]: $\sum_{i=1}^{n}\left(Y_i - X_i \times \hat{\beta}\right)^2$ (i.e., the squared difference between the observed signal $Y$ and the expected signal as specified by the $X$ matrix scaled by the $\beta$ parameters). The unknown parameter and its variance are thus estimated as follows[2]:

$$\hat{\beta} = \left(X^T X\right)^{-1} X^T Y \tag{2}$$

$$var\left(\hat{\beta}\right) = \sigma^2 \left(X^T X\right)^{-1} \tag{3}$$

According to the Gauss–Markov ($\mathcal{GM}$) theorem, the OLS will correspond to the best linear unbiased estimator (BLUE) of the population parameters, in the class of unbiased estimators, provided the following assumptions relating to the properties of the error term and the parameters hold true[3]:

(A1) Errors are independently and identically distributed (i.i.d.) $\sim N(0, \sigma^2 I)$
(A2) he regressors in the $X$ matrix are independent of error [i.e., $E(\varepsilon, X) = 0$], non-stochastic (i.e., deterministic), and known.
(A3) No regressor is a linear transformation of one (or more) other regressors.

It should be noted that (A1) is in fact a three-part assumption requiring that (A1a) the errors for different observations (time-points) are not correlated [i.e., $Cov(\varepsilon_i, \varepsilon_j) = 0$]; (A1b) the expected value of the error term is zero [i.e., $E(\varepsilon) = 0$]; and (A1c) the variance of the error is $\sigma^2$ at all observations.

Crucially, a statistical model is only valid inasmuch as its assumptions are met. When this is not the case, inferences drawn from it will be biased and can even be rendered invalid. The reminder of this paper focuses on the degree to which fMRI data abide by the above assumptions, the consequences of their infringement, and describes the main methods currently available to adjust for such situations.

## SINGLE-SUBJECT ANALYSIS (II): THE $\mathcal{GM}$ ASSUMPTIONS AND fMRI TIME-SERIES

### AUTOCORRELATION

One major potential violation of the model's assumptions arises from the fact that fMRI data represent a time-series. In particular, correlations between residuals at successive time-points can violate the i.i.d. assumption (A1a). Common sources of noise that can introduce serial correlations include hardware related low-frequency drifts, oscillatory noise related to respiration and cardiac pulsation, and residual movement artifacts not accounted for by image registration (c.f., Weisskoff et al., 1993; Friston et al., 1994; Boynton et al., 1996). The presence of serial correlation does not directly affect unbiasedness of the $\hat{\beta}$, but can produce biased estimates of the error variance. This can in turn lead to biased test statistics (e.g., $t$ or $F$ values), and affect statistical inferences based on those statistics.

In fMRI data, the problem is typically one of systematically over-estimating the error degrees of freedom, since the presence of serial correlations means that the true number of independent observations (the effective degrees of freedom) will be lower than the apparent number of observations. In turn, this produces an underestimate of the error variance and thus an inflated test statistics (since the error variance is included in its denominator). The artificially liberal nature of inferences drawn in the absence of correction for autocorrelation was demonstrated by Purdon and Weisskoff (1998) who found false positive rates as high as 0.16 could occur for a nominal $\alpha$-level of 0.05. Several different approaches have been suggested to minimize this problem.

#### *Temporal smoothing ("pre-coloring")*

Friston et al. (1995) and Worsley and Friston (1995) suggest an extension of the GLM to accommodate serial correlation via "temporal smoothing." Their proposal is to re-frame (1) as:

$$Y = X\beta + \Sigma\varepsilon \tag{4}$$

where $\Sigma$ represents some process hidden in the residual characterizing the serial correlation, and $\varepsilon$ represents a "well-behaved" error term $\sim N(0, \sigma^2 I)$. By then imposing a linear transformation $S$ to (4) the idea is to "swamp" and thereby minimize the endogenous – unknown – correlation structure with some exogenously imposed, therefore known, correlation structure $S$, obtaining:

$$SY = SX\beta + S\Sigma\varepsilon \tag{5}$$

The assumption underlying this method is that the $S$ transformation is robust enough so that $S\Sigma S^T \sim SS^T$, thereby effectively "swamping" the unknown endogenous serial correlation. If this assumption holds, the "colored" noise is identically distributed $\sim N(0, \sigma^2 S^T)$ (Friston et al., 1995). The derived $\beta$-estimates remain unbiased but do not retain maximal efficiency, thus degrading power, as a function of how (in)effectively the endogenous correlation is swamped. In addition, the pre-coloring smoothing function acts as a low-pass filter, and may risk attenuating experimentally induced signals of interest (Marchini and Ripley, 2000; Woolrich et al., 2001). Partly as a response to these problems, the pre-coloring approach has now been largely superseded by the pre-whitening approach.

#### *Pre-whitening*

Rather than attempting to mask an unknown covariance structure with a known one, the pre-whitening strategy attempts to estimate and remove the autocorrelation prior to estimating the model parameters (Bullmore et al., 1996). This technique makes use of a two-pass procedure. In the first pass, a GLM is fit to the data under the infringed i.i.d. error assumption. The residuals derived from this model are then used to estimate the autocorrelation structure actually present. The autocorrelation is then modeled with a simple Auto-Regressive model of order 1 [AR(1)], in which the error at each time-point is assumed to be a combination of the error at the previous time-point and some "fresh" error. Once the parameters for this model have been estimated, the raw data is "pre-whitened"

[1]As we will discuss in the remainder of the paper, several sophistications on top of this simple approach are necessary to compensate for specific features of BOLD time-series.

[2]Where the superscript "$T$" denotes the transposed matrix.

[3]In addition to the unbiasedness and maximal efficiency, the BLUE is asymptotically normal, a desirable property for subsequent parametric testing.

by removing the estimated covariance structure. Finally, a second pass of the GLM estimation is carried out on the whitened data. The intuition underlying such an approach is that if a good estimate of the autocorrelation structure can be obtained and removed from the data, the i.i.d. assumption (A1a) will hold. As a parallel to (5), this approach can be expressed as:

$$K^{-1}Y = K^{-1}X\beta + K^{-1}\Sigma\varepsilon \quad (6)$$

where $K \approx \Sigma$. Thus, instead of the convolution with a temporal smoothing matrix S as in Friston et al. (1995), the pre-whitening approach uses a de-convolution matrix obtained by data driven estimation of the $\Sigma$ structure. Unlike the pre-coloring approach, pre-whitening also has the advantage that the resulting parameter estimates will be the BLUEs. Thus, equations (2) and (3) can be re-written as:

$$\hat{\beta} = (X^T K^{-1} X)^{-1} X^T K^{-1} Y$$

$$var(\hat{\beta}) = \sigma^2 \left(X^T K^{-1} X\right)^{-1}$$

If the process $\Sigma$ is exactly characterized by $K$, then in the transformed data the error variance is equal to $\sigma^2 I$ again. Compared with pre-coloring, pre-whitening has the advantage of being more efficient (Woolrich et al., 2001) across different experimental designs, and particularly for rapid event-related designs where much of the experimentally induced signal is concentrated in high temporal frequencies. The approach, however, relies heavily on accurately characterizing the endogenous correlation structure and non-optimal modeling can reduce efficiency and induce significant bias, affecting the magnitude of test statistics (Friston et al., 2000a). In response to these problems, several authors have suggested the use of more complex models allowing for serial correlations over longer periods of time (Worsley et al., 2002), greater flexibility in the local noise modeling (e.g., Locascio et al., 1997; Purdon et al., 2001), and non-parametric approaches (Woolrich et al., 2001).

Lenoski et al. (2008) compared the performance of several different strategies for whitening residuals, including the global linearized AR(1) algorithm implemented in SPM2/5 (Frackowiak et al., 2004), a regularized non-parametric correction (Woolrich et al., 2001) similar to the one implemented in FSL (Smith et al., 2004), and a voxel-wise, regularized AR($m$) autocorrelation algorithm (Worsley et al., 2002) as implemented in fMRIStat. Across six algorithms, the global linearized AR(1) model proved to be the least effective, decreasing the count of non-white residuals to about 47%. Notably, the algorithm was effective at eliminating positive autocorrelation structures. Due to the inflexibility of the global approach, in which the same, positive correlation structure is assumed to exist at every voxel, white and negative autocorrelations were poorly modeled. As a consequence, this approach induced negative correlations that were not originally in the data (the practical implications of this remain debatable however, as Lenoski et al. (2008) also report that positive autocorrelations are present in the overwhelming majority of voxels; see also Woolrich et al., 2001; Worsley et al., 2002). Regularized algorithms, including AR(1) and non-parametric methods, successfully reduced non-whitened residuals to about 41 and 37% of the total voxels. As with the global linearized AR(1) approach, positive correlations were greatly reduced, although there was again a slight increase in negative correlations. The best performance was obtained with non-regularized non-parametric and AR(2) algorithms, which decreased the number of non-white residuals to an average of 1.6 and 0.5%, respectively. Overall, this supports the idea that the major source of autocorrelation in BOLD time-series is successfully captured by an AR(2) model.

#### *Explicit noise modeling*

Finally, an alternative approach to the data driven estimation of serial correlation has been proposed by Lund et al. (2006). In their interpretation, autocorrelation amongst residuals can often be taken as evidence of unmodeled, but potentially known, sources of variance. In their "nuisance variable regression" (NVR) approach they attempt to explicitly model within the design matrix several factors believed to induce autocorrelation, including hardware related low-frequency drift, residual movement effects, and aliased physiological noise (i.e., cardiac pulsation and respiration; see also Lund et al., 2006). Empirical tests suggested that the approach is effective in capturing the modeled sources of noise such as cardiac and respiratory effects (see also Razavi et al., 2003). Indeed, the NVR approach resulted superior to both the AR(1) and the simple high-pass filtering approaches for dealing with serial correlation. However, despite the potential appeal of this approach, it is highly dependent on the sources of noise being well characterized, and will still require some form of data cleaning to account for additional unmodeled sources of temporal correlation.

### HETEROSCEDASTICITY

Assumption (A1c) requires the variance of residuals to be constant across observations (i.e., time-points), and the covariances (i.e., the off-diagonal elements of the variance–covariance matrix) to be all equal to zero [i.e., $var(\varepsilon) = \sigma^2 I$]. Violation of such assumption is referred to as heteroscedasticity. When this assumption is not upheld the estimator ($\hat{\beta}$) is still unbiased, but no longer efficient, usually resulting in confidence intervals that are either too wide or too narrow. As for the case of autocorrelation, if the variance is biased subsequent parametric testing will yield incorrect statistics. In the fMRI literature heteroscedasticity has received relatively little attention. As a notable exception, Luo and Nichols (2002, 2003) discuss the possibility of heteroscedasticity in fMRI data, for example due to a dependency of the variances on the response, or because of other factors such as time or physical ordering (in Luo and Nichols, 2002, violations of homoschedasticity are indeed found, mostly due to artifacts).[4]

### MULTICOLLINEARITY AND *X* MATRIX MIS-SPECIFICATION

#### *Multicollinearity*

Assumption (A3) requires that none of the explanatory variables (i.e., columns of the $X$ matrix) is perfectly correlated with any other explanatory variable, or any linear combination of. In the presence of perfect multicollinearity the $X$ matrix is rank-deficient, the inverse of $(X^T X)$ does not exist, and infinite solutions

[4]Consequently, Luo and Nichols (2003) include a specific diagnostic test to assess the homoschedasticity assumption in their diagnostic package (see Luo and Nichols, 2003, p. 1016).

equally satisfy the GLM system of equations. In reality, however, the problem with multicollinearity is one of degree. The impact of multicollinearity on the $\beta$-estimates of the correlated columns is to reduce their efficiency (i.e., increase $var(\hat{\beta})$) as a positive function of the degree of collinearity present. The fundamental problem is that as two columns become more and more correlated it becomes impossible to disentangle the unique impact of each on the dependent measure. The confidence intervals for the coefficients thus become wide, possibly including zero, making it difficult to assess whether an increase in a regressors is associated with a positive or negative change in the dependent measure (i.e., $Y$). The consequence of this issue is then a strong bias in the inferential statistics (e.g., $t$-test – which can be either positive or negative). It should be noted, however, that multicollinearity only affects the repartition of variance among the individual regressors. Overall model statistics such as $R^2$ and the significance of the model remain unaffected, possibly leading to the seemingly paradoxical situation where individual $\beta$s have low significance but the overall model fit is high. A clear example of the impact of multicollinearity in functional neuroimaging experiments is provided in Andrade et al. (1999), where, in a PET experiment, the activations associated with a given regressor in the presence, or absence, of a strongly covarying second effect are qualitatively different. As the authors remark, the results obtained from the correlated and uncorrelated models could lead to qualitatively different conclusions (Andrade et al., 1999). Overall, this example shows that in the presence of multicollinearity functional neuroimaging results can be misleading and misinterpreted.

As for the homoscedasticity assumption, the multicollinearity issue did not find, so far, much space in the fMRI methodology literature. There may be at least two reasons for this: first, the standard use of the pseudo-inverse method; second, this problem is typically dealt with at the creation of the experimental design (i.e., the $X$ matrix). Indeed, much more energy has been spent on the issue of design efficiency (including multicollinearity minimization; c.f., Dale, 1999; Wager and Smith, 2003; Henson, 2007; Smith et al., 2007).

#### *X matrix mis-specification (I): model building*

Mis-characterizing the expected BOLD signal, by mis-modeling the $X$ matrix, can be an important source of error. According to Petersson et al. (1999), the specification of the $X$ matrix faces two connected trade-offs related to the cases of over- and under-specification. On the one hand, inclusion of a maximum number of effects in the model would be desirable to increase fit, though at the cost of reducing power (by consuming one $df$ for each additional effect), while the marginal increase of explained variance decreases with each additional factor. Further, over-modeling of the signal may degrade the generalizability of the results. On the other hand, exclusion of relevant factors from the model may have the effect of inflating the error variance, reducing power, and possibly introducing serial dependencies in the error term, thus infringing assumption (A1a). At the same time, however, exclusion of irrelevant effects from the model has the (positive) consequence of increasing power via increase of the $df_{model}$ (one per each excluded variable; see Petersson et al., 1999, pp. 1246–1247 for a complete discussion on the point).

Overall, the consequences of even minor cases of mis-modeling, can result in severe loss of statistical power and inflation of the false positive rate far beyond the nominal $\alpha$-level (Loh et al., 2008). Along similar lines, Razavi et al. (2003) analyze the importance of model building with respect to the impact of model mis-specification (by either excluding appropriate effects or including inappropriate effects) on its goodness of fit and validity. Using a forward selection approach, inclusion of appropriate regressors (e.g., task and several noise sources) increased both statistics. Inclusion of an inappropriate term, on the other hand, increased the model goodness of fit, but strongly decreased the model validity. Finally, the authors also point out that the sheer number of activated voxels under each model, a heuristic often used for model selection, was often in conflict with model validity and goodness of fit statistics. In particular, inclusion of an inappropriate effect in the model reduced its validity, but increased the number of active voxels.

#### *X matrix mis-specification (II): the hemodynamic response function*

Once the $X$ matrix is properly specified, the regressors are typically convolved with a hemodynamic response function (HRF), in order to transform neural responses to on–off stimulation into an expected vascular signal (see Boynton et al., 1996). The HRF thus characterizes the input–output behavior of the system (Stephan et al., 2004), imposing an expectation on how the BOLD signal in a voxel should vary in response to a stimulus. Even for a well specified $X$ matrix, incorrect modeling of the HRF might cause significant discrepancy between the expected and the observed BOLD signal, increasing the variance of the GLM coefficients, degrading power, and decreasing the model validity (c.f., Aguirre et al., 1998; Loh et al., 2008; Waldorp, 2009). Indeed, even minor model mis-specification can result in substantial power loss and bias, possibly inflating Type I errors (Lindquist et al., 2009). It is then all the more problematic that the HRF is known to be highly variable across individuals, and, within individuals, across tasks, regions of the brain, and different days (Aguirre et al., 1998; Handwerker et al., 2004). The input–output relationship between stimulation and BOLD response can be modeled in one of several ways. The most typical approach is to assume a linear time-invariant system, where the HRF is modeled by a set of smooth functions that, when overlapping, add up linearly (Friston et al., 1994; Boynton et al., 1996). In this approach, many models have been proposed, with various degrees of flexibility. At the low end of the spectrum, the HRF is considered to have a fixed shape (e.g., the difference between two gamma functions) except for its amplitude (Worsley and Friston, 1995). A popular alternative is to employ a canonical HRF together with two derivatives, to allow for (small) variations in latency and dispersion (Friston et al., 1998). Even greater flexibility can be afforded by using a larger set of basis functions, typically constrained to the space of plausible HRF shapes (Woolrich et al., 2004; Penny and Holmes, 2007) and its relevant parameters (Liao et al., 2002). At the top end of the flexibility spectrum, finite impulse response (FIR) basis sets allow estimation of the height of the BOLD response at each time-point (Glover, 1999; Ollinger et al., 2001).

In general, inclusion of multiple parameters has the advantage of acknowledging and allowing for known HRF variability, thus increasing sensitivity (Woolrich et al., 2004). At the same time, however, the increased flexibility comes at the cost of potentially

fitting physiologically ambiguous or implausible shapes (Calhoun et al., 2004; Woolrich et al., 2004), fewer degrees of freedom, and decreased power (Lindquist et al., 2009). Furthermore, when multiple HRF shapes are tested, it is not clear how to aggregate the results for group-level analysis (c.f., Calhoun et al., 2004; Steffener et al., 2010), nor it is clear how to interpret differences between tasks when spread over a multitude of parameters (Lindquist et al., 2009). A different approach is to adopt physiologically informed models of BOLD response, such as the Balloon model (Buxton et al., 1998; Friston et al., 2000b). In this model, a set of differential equations specifies a dynamic link between neuronal activity and transient increases in the rate of cerebral blood flow, in terms of volume and dHb content. The elicited BOLD signal is then considered to be proportional to the ratio of these two quantities (Friston et al., 2000b). While more biologically plausible, this approach face several difficulties relating to estimation of a large number of parameters, unreliability of estimates in the presence of noisy data, and the lack of a direct framework for inference-making (Lindquist, 2008).

Overall, most cognitive fMRI research to date appears to be exclusively focused on estimating the magnitude of evoked activations (Lindquist and Wager, 2007; Lindquist et al., 2009), and does not pay much attention to HRF variability. Indeed, as revealed by a recent survey of 170 fMRI studies, 96% of experiments adopting an event-related design used a canonical HRF model, thus ignoring differences in shape between individuals or areas of the brain (Grinband et al., 2008). As noted by Lindquist (2008), building of more sophisticated HRF models is likely to be, in the coming years, one of the areas of great multidisciplinary focus.

## LINEARITY

The GLM approach also assumes effects to add linearly to compose the response measurements. Boynton et al. (1996) tested this assumption by parametrically varying a visual stimulus' duration and contrast. Investigating the additivity of the noise in V1, they concluded that although deviations from linearity were measurable, these were not strong enough to reject the GLM. Support for the use of a linear approach was also offered in Cohen (1997), where response amplitudes to parametric variations of the stimuli were well modeled by a piecewise linear approximation. Despite this initial evidence, it has now been extensively shown that there are at least two sources of non-linearities in the BOLD signal: the vascular response, especially the vasoelastic properties of the blood vessels (see Buxton et al., 1998), and adaptive behavior in neuronal-response (e.g., Logothetis, 2003).

Vazquez and Noll (1998) tested the linearity of the BOLD response to (visual) stimulation of varying length. Under the linearity assumption, it should be possible to predict the amplitude of the response at a given duration by multiplying the amplitude of the response at a shorter duration an appropriate number of times. When stimuli of at least 5 s were used to predict the BOLD response amplitude at longer intervals, this expectation was met. However, when stimuli of 4 s or less were used to predict the response at longer durations, these were found to greatly overestimate the observed amplitudes. A similar result was reported, using auditory stimuli, by Robson et al. (1998). Consistently with the results in Vazquez and Noll (1998), it was possible to predict the BOLD response amplitude at long durations with stimuli of at least 6 s. Stimuli of shorter duration yield dramatic overestimates of the response amplitude at longer durations, as a positive function of the time difference between the predictor and the predicted stimulus length (see Robson et al., 1998, Figure 4, p. 191). The authors thus suggest including in the model an adaptive component that may discount the response amplitude for short stimulations specified as:

$$E(t) = (1 - A) + Ae^{-t\alpha} \tag{7}$$

Equation (7) essentially represents a scaling factor to be applied to the amplitudes of short latency stimuli in order to correct for "transient" non-linearity. In Robson et al. (1998), this approach reduced the discrepancy between the observed response at the longest latency (25.5 s) and the predicted one (from the shortest latency – 100 ms) from 11.09% signal change to only 0.88%.[5]

Friston et al. (1998) used parametric variations in the rate of word presentation to assess the presence of non-linear BOLD effects. The observed departure from linearity was interpreted in terms of a hemodynamic "refractoriness," according to which a prior stimulus interacts with a following – temporally contiguous – stimulus by modulating its response amplitude. As a solution, the authors proposed "linearizing the problem," by employing Volterra series to overtly characterize the non-linear component of the response. The observed BOLD signal $Y(t)$ can then be modeled as:

$$Y(t) = g^0 + \sum_{i=1}^{P} g_i^1 X_i(t) + \sum_{i=1}^{P}\sum_{j=1}^{P} g_{ij}^2 X_i(t) \cdot X_j(t) + e(t) \tag{8}$$

The second term of the equation represents the change in output (i.e., $Y$) for a given change in input. The third term is the part of the model that describes the effect of the response at one time-point on a temporally contiguous time-point, with the parameters $g^0$, $g^1$, $g^2$ representing the scaling factor of a series of $P$ basis functions approximating the zeroth, first, and second Volterra smoothing kernels (see Friston et al., 1998, p. 42). One criticism to this approach raised in Calvisi et al. (2004) and Friston et al. (2000b) is that while data driven computation of Volterra series parameters may allow for a better input–output mapping, it does so in a black-box fashion without being informative on what are the processes generating the non-linearities. In response to these criticisms Friston et al. (2000b) present evidence for the non-linearities expressed in the Balloon model of hemodynamic signal transduction (see Buxton et al., 1998) being compatible with a second order Volterra characterization, thus adding biological plausibility to the model.

A different approach has been proposed by Wager et al. (2005) who report substantial non-linearities in the magnitude, peak delay, and dispersion of the response for a stimulus presentation rate of 1 s. Noting the consistency of such non-linearities across the brain they suggest empirically deriving the functional form of each of these characteristics of the response as a function of stimulus history. The authors approximate the non-linearities with a biexponential model:

$$y = Ae^{-\alpha X} + Be^{-\beta X}$$

[5]The values of parameters $A$ and $\alpha$ were computed empirically by minimizing the discrepancy between predicted and actual signal.

By fitting the parameters $A$ and $\alpha$, $B$ and $\beta$, the scaling and exponent of two exponential curves, the authors empirically characterize the non-linear changes in BOLD magnitude, onset time, and peak delay. The idea is to first run an experiment from which to derive the fixed parameters estimates and then use the non-linear characterizations as scaling factors for individual responses – according to the history of stimulation up to each response – in following experiments.

The issue of non-linearity is particularly relevant to fast event-related designs. When short intervals separate periods of stimulation, the response to the individual stimuli will superpose, and will do so sub-additively, reducing the observed signal, presumably as a consequence of neuronal and vascular factors (e.g., Birn and Bandettini, 2005; Heckman et al., 2007; Zhang et al., 2008). Several studies have now documented the decrease in the estimated response amplitude at short inter-stimulus intervals (ISI). In Miezin et al. (2000), for example, average ISIs of about 5 s (with a minimum of 2.5 s) resulted in a decrease of 17–25% of the signal obtained in widely spaced trials (e.g., 20 s). Similarly, Zhang et al. (2008) reported significant decreases in response amplitude for ISIs of 1 and 2 s, as compared to longer spacing (i.e., 4, 6, and 8 s). At the lower end of stimulus spacing, Heckman et al. (2007) compared BOLD response in visual cortex to stimuli of varying contrast under a "spaced" (3 s ISI) and a rapid (1 s ISI) presentation. Qualitatively, the patterns of response across designs were similar. Quantitatively, however, the rapid presentation had the effect of scaling the strength of the response. In particular, the response reduction observed when switching from a spaced to a rapid presentation was found to be similar to the reduction observed when decreasing the stimulus contrast (under the same presentation condition). In primary visual cortex, for example, switching to a rapid design induced a response reduction equivalent to that observed when decreasing the stimulus contrast by 84%. Finally, while the response non-linearity appears stable within a given region (Miezin et al., 2000), it may vary greatly across different primary cortices (e.g., Miezin et al., 2000; Soltysik et al., 2004) and may increase in associative areas (Huettel and McCarthy, 2001; Boynton and Finney, 2003; Heckman et al., 2007).

Overall, as noted by Wager et al. (2005) non-linearities are largely ignored in the neuroscientific and psychological BOLD-fMRI literature. Several reasons may underlie this observation. First, there exists an "envelope" within which the linearity assumption holds (e.g., ISIs ≥ ~4 s). An informal PubMed survey of 20 papers published in 2010 mentioning the words "rapid/fast event-related fMRI" in the abstract revealed that half the designs used an average ISI of at least 4 s. Of the remaining, half the designs made use of an ISI between 3.5 and 4 s, and the remaining employed even shorter intervals. The large majority of studies thus falls at the boundary of the envelop, or well within it. No single study mentioned non-linearities in BOLD response at short ISIs. All studies except for one, however, made use of stimuli (pseudo-)randomization and/or ISI random jittering, with the aim of maximizing power and mitigating non-linear effects by making each stimulus category equiprobable at each trial (c.f., Dale, 1999; Henson, 2007). Second, the increased statistical power conferred by a greater number of trials, under fast designs, may well outweigh the amplitude reduction induced by overlapping BOLD responses (calculated to be about 20% in Miezin et al., 2000). Third, the bulk of the work has been devoted to determining canonical responses to a single stimulus rather than to exploring interactions among multiple stimuli. Finally, most proposed solutions (e.g., Volterra series; see Friston et al., 1998, 2000a) require fitting of a large number of parameters which may cause severe degradation of power.

## MULTIPLE SUBJECTS ANALYSIS

Once single-subject data has been analyzed for a set of participants, individual results are aggregated to assess commonality and stability of effects within or across groups of interest (Holmes and Friston, 1998; Worsley et al., 2002). Prior to group analysis, however, for datasets to be comparable across subject, individual results are warped into a common reference space (typically either the Talairach, Talairach and Tournoux, 1988, or the MNI152, Evans et al., 1997), in order to "align" corresponding cerebral structures across subjects with differing brain anatomy. This normalization procedure is all but uncontroversial, especially in relation to its effectiveness (see Brett et al., 2002), however, a discussion of the issue extends beyond the scope of this review.

One of the central issues in group analysis concerns the scope of the inferences one may validly draw from aggregate statistics. As will be discussed below, some group analysis strategies only afford making valid inferences about the specific sample (i.e., participants) one has tested. Other strategies, however, afford making valid, and typically more interesting, inferences about the population from which the sample was drawn.

### FIXED EFFECTS

As Holmes and Friston (1998) nicely put, classical statistical hypothesis testing proceeds by comparing the difference between the observed and hypothesized effect against the "yardstick" of variance (p. S745). The scope of an inference is then bound by the yardstick employed. In a fixed effects (FFX) analysis, the variance considered is that derived from *scan-to-scan* measurement error, and represents the within-subject variability ($\sigma_w^2$). This variability may include physiological task-related (e.g., adaptation, learning, and strategic changes in cognitive or sensory-motor processing) and task non-related effects (e.g., changes in global perfusion secondary to vasopressin secretion in the supine position), as well as non-physiological noise, such as gradient instabilities (c.f., Friston et al., 1999a). In this approach statistical testing assesses whether a response is significant with respect to the precision with which it can be measured (Friston et al., 2005). Paraphrasing the very intuitive example offered in Mumford and Nichols (2009), if one were to measure the length of hair from the same set of participants twice, it is reasonable to assume that the only difference across the two measurements should relate to small variation around the average hair length of each participant. In this sense, for each subject the magnitude of the effect is considered to be fixed. FFX analyses thus represent the population variance as being a sole function of within-subject variability divided by the product of subjects ($N$) and number of observations per subject ($n$) (c.f., Penny and Holmes, 2007). For a given subject $i$, the observed response in trial $j$ (i.e., $y_{ij}$) is modeled as varying around the subject's mean effect $d_i$ plus a within-error component $e_{ij}$ (with mean zero and variance $\sigma_w^2$):

$$y_{ij} = d_i + e_{ij} \quad (9)$$

For a single subject *i*, then, the (maximum likelihood) parameter estimate and its variance are:

$$\hat{d}_i = \frac{1}{n}\sum_{j=1}^{n} y_{ij} \tag{10}$$

$$var\left(\hat{d}_i\right) = \frac{\sigma_w^2}{n} \tag{11}$$

As shown in equation (11), the variance of the effect for each subject *i* is the within-error divided by the number of observations *n*. To compute the group-level analysis of the population effect ($d_{pop}$), then, it suffices to aggregate the individuals' effects ($\hat{d}_i$) for all *N* subjects, yielding (see Penny and Holmes, 2007, Section 3 for the full derivation):

$$\hat{d}_{pop} = \frac{1}{N}\sum_{i=1}^{N} \hat{d}_i \tag{12}$$

$$Var\left(\hat{d}_{pop}\right) = \frac{\sigma_w^2}{Nn} \tag{13}$$

The crucial point shown in (13) is that the group estimate's variance in an FFX approach is a function of the scan-to-scan (i.e., within-subject) variability $\sigma_w^2$ only. This group analysis is thus conceptually equivalent to concatenating all data and performing a single GLM on a "super-subject" with $N \times n$ observations.

Importantly, the inferences drawn from a FFX analysis are not invalid, rather they are only valid with respect to the yardstick employed (i.e., $\sigma_w^2$). Inferences are thus supported at the level of the sample analyzed, but not at the level of the population from which the sample is drawn (given that there is no consideration of "sampling variability"). As noted by Friston et al. (1999a), a FFX approach makes the assumptions that each subject makes the same contribution to the observed activation thus discounting random variation from subject to subject (see the data presented in Penny and Holmes, 2007, Figure 2 for a dramatic example of *subject-to-subject* variability). This type of analysis can thus be seen as relevant in "single case" studies (Penny and Holmes, 2007), but seems unacceptable for "standard" fMRI experiments of healthy volunteers, and their (desired) inferential scope.

## RANDOM EFFECTS, MIXED EFFECTS, AND SUMMARY STATISTICS

### *Random and mixed effects*

For inferences to apply at the population level it is necessary to account for the fact that individual subjects themselves are sampled from the population and thus random quantities with associated variances (c.f., Mumford and Nichols, 2006; Mumford and Poldrack, 2007, for a very clear explanation and examples). The yardstick of variance must thus account for the *subject-to-subject* variation ($\sigma_b^2$). In the random effects (RFX) approach, the magnitude of the effect in each subject is no longer considered fixed, as in FFX analyses, but rather is a random variable itself. In this approach, statistical testing assesses whether the magnitude of an effect is significant with respect to the variability across subjects. There are several reasons for assuming that across-subject variation is present in fMRI data. In particular, this variation can be due to any (and any interaction) of several factors such as general subject differences in neural or hemodynamic response to stimulation, and/or differing underlying anatomy (c.f., Friston et al., 1999a). Further, any of the above-mentioned within-subject variations may be of different magnitude across subjects and, finally, many non-physiological noise sources could affect the way in which a BOLD effect (even assuming this was actually the same across several subjects) could give rise to different data (e.g., radio-frequency and gradient instabilities, re-calibration of the scanner, repositioning effects or differential shimming effects). It should be noted, however, that unless the true vector of (single subject) $\beta$s is known, it is not possible to draw pure random effects inferences (Bianciardi et al., 2004). The standard approach (often incorrectly referred to as a RFX nonetheless; see Smith et al., 2005) is then to include a mixture of within-session fixed effects and across-session random effect, thus generating a so-called mixed effect model (MFX; c.f., Beckmann et al., 2003; Smith et al., 2005). In this approach, the single observation for each individual (i.e., $y_{ij}$) is still centered around the subject's true mean $d_i$ plus a within-subject error component $e_{ij}$, as in equation (9). The subject's mean $d_i$ itself, however, is now characterized as a random variable that is centered around the real population mean $d_{pop}$ plus a between-subjects error component $z_i$ (with zero mean and variance $\sigma_b^2$). If we restate $d_i$ in equation (9) as $d_{pop} + z_i$ we obtain the "all-in-one" model:

$$y_{ij} = d_{pop} + z_i + e_{ij} \tag{14}$$

The group effect estimate, and its associated variance, are now equal to (see Penny and Holmes, 2007, for the full derivation):

$$\hat{d}_{pop} = \frac{1}{Nn}\sum_{i=1}^{N}\sum_{j=1}^{n} d_{ij} \tag{15}$$

$$Var\left(\hat{d}_{pop}\right) = \frac{\sigma_w^2}{Nn} + \frac{\sigma_b^2}{N} \tag{16}$$

It is immediately clear from (16) that in MFX analyses the yardstick used for statistical testing results from a mixture of the "within" ($\sigma_w^2$) and "across" ($\sigma_b^2$) sources of variability. It is also important to notice that, in equation (16), both sources of variance are scaled by the total number of subjects (*N*), while only the within-subject variance is scaled by the number of observations per subject (*n*). Thus, as a general rule, more subjects may be better than more observations per subject (Penny and Holmes, 2007).

### *Summary statistics: the hierarchical approach*

A straight-forward strategy to perform group analysis is to formulate a "single-level" GLM in which various parameters of interest at the group level are estimated directly from all of the original single sessions' time-series data (Beckmann et al., 2003). This all-in-one model can be specified as:

$$Y_G = XX_G\beta_G + \gamma \tag{17}$$

where $Y_G$ is now the full data vector [composed of all the individual subjects' time-series $Y_i$ from equation (1)], *X* is the single-subject design matrix, $X_G$ is a group-level matrix specifying how the individual subjects' data are to be related (e.g., all averaged in a single group, divided into two groups of interest), and $\gamma$ is the error

term (comprised of within-subject and across-subject variance; see below). This approach, though simple, is computationally very inefficient because of the size of matrices for datasets of more than 100 time-points (i.e., volumes) for more than 100,000 voxels, for one or more sessions, for each of 15 (or more) subjects (Mumford and Poldrack, 2007). For this reason, Holmes and Friston (1998) first proposed a computationally simpler hierarchical model of group analysis. This approach, typically referred to as the *Summary Statistics* approach, is based on a two-level strategy. First, a single GLM is carried out for each subject individually (i.e., first-level analysis). Following, the single-subject estimates (e.g., the $\hat{\beta}$s or a contrast of interest $c\hat{\beta}$), and not the time-series themselves, are carried forward to the second step where a group-level test is performed (these ideas are further developed in Penny and Holmes, 2007, and in Beckmann et al., 2003, though with some important differences, as discussed below).

A hierarchical two-level linear model of (fMRI) data analysis can be written as follows (c.f., Bianciardi et al., 2004):

$$\begin{cases} Y = X\beta + \epsilon & (1^{\text{st}}\,\text{level; fixed effects}) \\ \beta = X_G\beta_G + \epsilon_G & (2^{\text{nd}}\,\text{level; random effects}) \end{cases} \quad (18)$$

As mentioned above, however, the true vector of effect size $\beta$ is not known, hence, in the summary statistics approach it is the *estimated* parameters (i.e., $\hat{\beta}$, or $c\hat{\beta}$) that are brought forward from the first-level analysis to the second. The hierarchical model in (18) can thus be restated as:

$$\begin{cases} Y = X + \epsilon & (1^{\text{st}}\,\text{level; fixed effects}) \\ \hat{\beta} = X_G\beta_G + \eta & (2^{\text{nd}}\,\text{level; mixed effects}) \end{cases} \quad (19)$$

where the error term $\eta$ is not equal to the random effect component $\epsilon_G$, but contains a mixture of both the within (i.e., fixed) and between (i.e., random) variability (hence the characterization as a MFX).

It is important to notice that the summary statistics strategy is equivalent to the all-in-one model described by equation (17) only under the assumption that the first-level variances are homoschedastic, and can thus be assumed to be equal across subjects (Beckmann et al., 2003; Penny and Holmes, 2007). More in general, the concern is whether sphericity can be assumed (i.e., error terms are identically and independently distributed). When this assumption is not met, the error term will no longer be a scalar multiple of the identity matrix (i.e., $\sigma^2 I$), which will reduce efficiency of the estimators. As discussed in Friston et al. (2005), three main factors determine whether the sphericity assumption is tenable in group analysis. First, the within-session error (co) variance must be the same for all subjects (i.e., subjects must exhibit the same amount of measurement error). Second, the first-level $X$ matrix must be the same for all subjects (i.e., the design must be "balanced"). Third, a one-dimensional contrast is brought forward from the first-level to the group-level analysis. Under these circumstances, non-sphericity at the second level induced by differences in first-level variances can be ignored (and the group-level effect can be computed using the efficient hierarchical summary statistic approach). These conditions, however, cannot always be met, as in the case where the specification of the $X$ matrix is dependent on subject-specific performance (e.g., *post hoc* classification of trials in "remembered" or "forgotten"), or when one or more subjects exhibit particularly high variance, compared to the rest of the sample. Friston et al. (2005) thus propose a summary statistics model formally identical to that described by Penny and Holmes (2007), except for the use of a restricted maximum likelihood (ReML) approach to estimate, from the first level, the amount of non-spehricity induced by the individual subjects' variance components. The ReML estimates of non-sphericity (over responsive voxels only) can then be entered in the group-level parameter estimation as a known quantity (see Friston et al., 2005, p. 247, for a schematic representation of this approach) removing the dependence on the sphericity assumption. The authors then address in real data the performance of the Holmes and Friston (1998) summary statistic approach under the violation of group-level sphericity, and the performance of the ReML approach under the same conditions. Interestingly, the Holmes and Friston (1998) approach appeared to be robust to heterogeneity in first-level design matrices and unequal first-level variance. However, the ReML approach did perform marginally better, in terms of group-level statistics and associated $p$-values.

In contrast to the Friston et al. (2005) results, numerical simulations conducted by Beckmann et al. (2003) show that the conventional Holmes and Friston (1998) approach can indeed yield suboptimal group-level statistics across a wide variety of designs (e.g., mean group activation, paired $t$-tests, $F$-tests) when the second-level assumption of sphericity is not met. Beckmann et al. (2003) thus propose a generalization of the summary statistic approach that retains mathematical equivalence with the all-in-one analysis also when group-level sphericity is violated. In particular, they show that the summary statistic approach described in (19) can be made equivalent to the all-in-one approach [equation (17)] if the group-level variance is set equal to the sum of the estimated between-subjects variance and the first-level parameter variance structures (c.f., Beckmann et al., 2003, Section II.C). According to this approach, it suffices to carry forward to the group-level analysis both the first-level estimates (i.e., $\hat{\beta}$ or $c\hat{\beta}$) and their (co)variance structures to correctly implement the hierarchical equivalent of the all-in-one model. The mathematical argument is empirically supported by a substantial increase in group-level $Z$-scores in the generalized model, as compared to the Holmes and Friston (1998) approach, under different violations of the group-level sphericity. More importantly, as Beckmann et al. (2003) point out, the increase in $Z$-scores is about typical threshold values (i.e., from values of 2.0 to ~3.0), thus likely to affect inferences made on thresholded images.

In a more recent study, Mumford and Nichols (2009) compared, with respect to power and specificity (i.e., Type I error rates), the performance of the Holmes and Friston (1998) approach with models that include first-level variances. Over a range of sample sizes and non-sphericity (induced by outlier variance), their simulated data shows that while the weighted approaches are more optimal in ensuring outlier down-weighting, in the case of a one-sample $t$-test, the conventional summary statistics model is robust to group-level sphericity violations. In particular, while this latter strategy does suffer some power loss (up to about 9% at the lower end of the simulated sample sizes), it still correctly controls (if

slightly conservatively) for Type I errors. It is important to stress, however, that these results are unlikely to replicate in other cases (e.g., simple linear regression).

Finally, a last issue relates to the sensitivity of R/MFX analyses. Indeed, while this approach has the desirable property of allowing valid inference at the population level, comparing the magnitude of an effect of interest to both the within- and the across-subject variability may result in significantly less sensitivity, as compared to FFX approaches (Friston et al., 1999b). To achieve sufficient power and acceptable reliability, it might thus be necessary to obtain a sample of 25–27 participants (Desmond and Glover, 2002; Thirion et al., 2007), which is about 30% more than the current typical sample size of 15–20.

## DISCUSSION

Throughout the past 20 years, the GLM has arguably become the most widely employed approach to analyzing fMRI data. One of the main advantages of this framework is its great flexibility, which allows for a multitude of testing strategies (e.g., *t*/*F*-test, ANOVA, ANCOVA). However, for the statistical model to be valid the assumptions on which it relies must be met.

### HOW WELL DO FMRI TIME-SERIES CONFORM TO THE MODEL'S ASSUMPTIONS?

Overall, some of the $\mathcal{GM}$ assumptions appear to be particularly problematic for fMRI datasets; however, the increased (but variously implemented) sophistication of fMRI analysis strategies mitigates this issue. At the first-level analysis, the presence of autocorrelation in the residuals, its biasing effects on the precision of the estimates, and its (possibly severe) inflation of Type I errors, has been long discussed and addressed in a variety of ways. Currently, the pre-whitening approach seems to be the standard choice. However, even within the domain of pre-whitening strategies, different algorithms can yield substantially different results in terms of power and false positives rate, possibly leading to very different inferences (Lenoski et al., 2008). Furthermore, it should also be stressed that results produced assuming white residuals (as done by some fMRI analysis software) should be interpreted with great caution due to the inflated effective $\alpha$-level. This is particularly true for studies conducted at the single-subject level, as in single-patient reports. The specification of the $X$ matrix also appears to be problematic. Correlation among the columns, for example, can lead to entirely erroneous qualitative interpretation of the data (see Andrade et al., 1999), stressing the importance of employing tools to assess and build experimental designs as efficiently as possible (see Dale, 1999; Wager et al., 2005; Smith et al., 2007). In addition, even for well specified and efficient designs, mis-specification of the HRF, something that – simply stated – occurs regularly in the cognitive fMRI literature (c.f., Grinband et al., 2008), can also result in substantial power loss and bias (Lindquist et al., 2009). The linearity assumption appears to be only partially problematic since it is only really violated in a specific subspace of experimental designs (e.g., stimuli spacing < ~4 s), and even when violated it does not result in excessive amplitude reduction (c.f., Miezin et al., 2000). Furthermore, the increase in power obtained by including a greater number of trials (at shorter ISIs), in conjunction with condition randomization and ISI jittering, may well outweigh the response amplitude reduction due to non-linearities. Other assumptions, such as homoschedasticity, have received little attention, also because they have not been found to be overly problematic (e.g., Razavi et al., 2003; Mumford and Nichols, 2009). With respect to second-level analysis, the most discussed issue is whether subjects can be assumed to have similar variance or not, and the possibility of unbalanced designs (Holmes and Friston, 1998; Beckmann et al., 2003). While there still is debate concerning the extent and severity of sphericity violations, some analyses show that it is generally not a problem in the context of 1-sample *t*-tests (Mumford and Nichols, 2009), although it may well be in other designs. Finally, it should be noted that there are many other important issues that could not be reviewed here (e.g., gaussianity of the BOLD signal, Hanson and Bly, 2001, and correction for multiple comparison, e.g., Thirion et al., 2007).

### WHAT IS THE ALTERNATIVE?

One last question relates to what alternatives to the GLM are available. Exploratory approaches (e.g., ICA) notwithstanding, there are at least three alternatives that, while making use of the GLM for the purposes of *estimation*, do not rely on it for *inference-making*. Non-parametric approaches, for example, can be employed to this end under the main constraint of exchangeability of observations (Holmes et al., 1996; Nichols and Holmes, 2001). This strategy has been recently argued to be generally preferable to parametric testing (Thirion et al., 2007). Bayesian methods have also been proposed, where "posterior probability maps" (i.e., images of the probability that an activation exceeds some specific threshold, given the data) can be used for inference-making (Friston and Penny, 2003). Posterior inference also has the advantage (among others) of not suffering from the multiple comparison problem since, as false positives cannot occur, the probability that activation has occurred, given the data, at any particular voxel is the same, irrespective of the number of analyzed voxels (see Friston et al., 2002, for an overview of advantages of Bayesian inference over classical one). Finally, a different approach is to derive $\beta$-estimates from a GLM but then assess significance of spatial distribution of activations, rather than individual voxels (Kriegeskorte et al., 2006). This strategy, by switching from a massive-univariate to a (local) multivariate approach has the promising advantage of assessing patterns of information representation, rather than localization of information, something that may be of great interest from a cognitive neuroscience point of view.

## CONCLUSION

Overall, the GLM approach to fMRI time-series remains a relatively intuitive and highly flexible tool, especially in light of the many sophistications that have been introduced to resolve assumption infringements. Nonetheless, it is also clear that in the current literature some problems are almost entirely ignored (e.g., HRF mis-specification; see Grinband et al., 2008), while others, because of different approaches to correction, can still lead to substantially different results (e.g., autocorrelation; see Lenoski et al., 2008). The main problem, however, is typically not one of bias, but rather one of variance of the estimators, power, and false positive rate. Furthermore, even though the first-level assumptions are typically the most problematic, assumption infringement in single-subject

analysis affects *both* first-level and group-level estimates, *Z*-scores, power, and false positives (particularly for the widely employed event-related design; see Bianciardi et al., 2004, for an experimental demonstration). This "percolation" effect from first- to higher-level analysis (Friston, 2007) is present regardless of whether the within-session variance is overtly carried forward to group-level analysis.

These statistical concerns are made worse by the standard "quantitative-to-qualitative" transformation of SPMs by which continuous measures (e.g., *Z*/*t*-scores) are converted into binary information (i.e., "active/not-active") for the purposes of inference. Indeed, the values around which statistics can vary due to assumption infringements and different correction strategies is just about standard thresholding levels (e.g., in the range of *Z*-values of 2.0 and 3.0; see Beckmann et al., 2003). As a consequence, even assuming a similar distribution of $\hat{\beta}$ values in a given brain region, fairly small differences in the amount of assumption infringement induced by any of the issues discussed above, along with differences in the effectiveness of the corrections used, will affect activation statistics just enough that the resulting post-thresholding map may be qualitatively very different, affecting overall convergence of results – something not infrequent in the fMRI literature. This issue is all the more problematic in light of the many different thresholding strategies available (c.f., Thirion et al., 2007), and the standard use of subjectively set parameters (e.g., *Z*-score cut-off values).

Considering all these issues, it is perhaps surprising that, as opposed to other experimental sciences, checking of the model's assumptions is virtually absent, or at least almost never reported, in the cognitive fMRI literature. A major factor in this respect is certainly the complexity of such procedures due to the massive amount of data (see Loh et al., 2008). However, the increased availability of diagnostic tools to assess hypotheses (e.g., Luo and Nichols, 2003; Loh et al., 2008) and other important statistics such as effect size, and power calculation (e.g., Mumford and Nichols, 2008), will hopefully lead to a more widespread checking of the model's assumption, and perhaps, as a consequence, greater replicability across studies.

## ACKNOWLEDGMENTS

The author would like to thank Andrew Engell, Rik Henson, Ian Nimmo-Smith, Daniel Osheron, Russell Thompson, and Daniel Wakeman for helpful discussion and feedback on previous versions of the manuscript. Any error or misunderstanding is sole responsibility of the author.

## REFERENCES

Aguirre, G. K., Zarahn, E., and D'Esposito, M. (1998). A critique of the use of the Kolmogorov–Smirnov (KS) statistic for the analysis of BOLD fMRI data. *Magn. Reson. Med.* 39, 500–505.

Andrade, A., Paradis, A.-L., Rouquette, S., and Poline, J.-B. (1999). Ambiguous results in functional neuroimaging data analysis due to covariate correlation. *Neuroimage* 10, 483–486.

Beckmann, C. F., Jenkinson, M., and Smith, S. M. (2003). General multilevel linear modeling for group analysis in fMRI. *Neuroimage* 20, 1052–1063.

Bianciardi, M., Cerasa, A., Patria, F., and Hagberg, G. E. (2004). Evaluation of mixed effects in event-related fMRI studies: impact of first-level design and filtering. *Neuroimage* 22, 1351–1370.

Birn, R. M., and Bandettini, P. A. (2005). The effect of stimulus duty cycle and 'off' duration on BOLD response linearity. *Neuroimage* 27, 70–82.

Bowman, F. D. (2007). Spatiotemporal models for region of interest analyses of functional neuroimaging data. *J. Am. Stat. Assoc.* 102, 442–453.

Bowman, F. D., Guo, Y., and Derado, G. (2007). Statistical approaches to functional neuroimaging data. *Neuroimaging Clin. N. Am.* 17, 441–458.

Boynton, G. M., Engel, S. A., Glover, G. H., and Heeger, D. J. (1996). Linear systems analysis of functional magnetic resonance imaging in human v1. *J. Neurosci.* 16, 4207–4221.

Boynton, G. M., and Finney, E. M. (2003). Orientation-specific adaptation in human visual cortex. *J. Neurosci.* 23, 8781–8787.

Brett, M., Johnsrude, I. S., and Owen, A. M. (2002). The problem of functional localization in the human brain. *Nat. Rev. Neurosci.* 3, 243–249.

Bullmore, E., Brammer, M., Williams, S., Rabe-Hesketh, S., Janot, N., David, A., Mellers, J., Howard, R., and Sham, P. (1996). Statistical methods of estimation and inference for functional MR image analysis. *Magn. Reson. Med.* 35, 261–277.

Buxton, R. B., Wong, E. C., and Frank, L. R. (1998). Dynamics of blood flow and oxygenation changes during brain activation: the balloon model. *Magn. Reson. Med.* 39, 855–864.

Buzsáki, G., Kaila, K., and Raichle, M. (2007). Inhibition and brain work. *Neuron* 56, 771–783.

Calhoun, V., Stevens, M., Pearlson, G., and Kiehl, K. (2004). fMRI analysis with the general linear model: removal of latency-induced amplitude bias by incorporation of hemodynamic derivative terms. *Neuroimage* 22, 252–257.

Calvisi, M. L., Szeri, A. J., Liley, D. T. J., and Ferree, T. C. (2004). Theoretical study of BOLD response to sinusoidal input. *IEEE* 1, 659–662.

Cohen, M. S. (1997). Parametric analysis of fMRI data using linear systems methods. *Neuroimage* 6, 93–103.

Dale, A. (1999). Optimal experimental design for event-related fMRI. *Hum. Brain Mapp.* 8, 109–114.

Desmond, J., and Glover, G. (2002). Estimating sample size in functional MRI (fMRI) neuroimaging studies: statistical power analyses. *J. Neurosci. Methods* 118, 115–128.

Evans, A., Collins, L., Paus, C. H. T., MacDonald, D., Zijdenbos, A., Toga, A., Fox, P., Lancaster, J., and Mazziota, J. (1997). A 3d probabilistic atlas of normal human neuroanatomy. Third International Conference on Functional Mapping of the Human Brain. *Neuroimage* 5, S349.

Frackowiak, R. S. J., Friston, K. J., Frith, C., Dolan, R., Price, C. J., Zeki, S., Ashburner, J., and Penny, W. D. (eds). (2004). *Human Brain Function*, 2nd Edn. New York: Academic Press.

Friston, J. K., and Penny, W. (2003). Posterior probability maps and SPMs. *Neuroimage* 19, 1240–1249.

Friston, K. (2007). "Statistical parametric mapping," in *Statistical Parametric Mapping*, eds K. Friston, J. Ashburner, S. Kiebel, T. Nichols, and W. Penny (London: Academic Press), 10–31.

Friston, K. J., Holmes, A., Poline, J.-B., Grasby, P., Williams, S., Frackowiak, R., and Turner, R. (1995). Analysis of time-series revisited. *Neuroimage* 2, 45–53.

Friston, K. J., Holmes, A. P., Price, C. J., Bruchel, C., and Worsley, K. J. (1999a). Multisubject fMRI studies and conjunction analyses. *Neuroimage* 10, 385–396.

Friston, K. J., Holmes, A. P., and Worsley, K. J. (1999b). How many subjects constitute a study? *Neuroimage* 10, 1–5.

Friston, K. J., Jezzard, P., and Turner, R. (1994). Analysis of functional MRI time-series. *Hum. Brain Mapp.* 1, 153–171.

Friston, K. J., Josephs, O., Rees, G., and Turner, R. (1998). Nonlinear event-related responses in fMRI. *Magn. Reson. Med.* 39, 41–52.

Friston, K. J., Josephs, O., Zarahn, E., Holmes, A. P., Roquette, S., and Poline, J. B. (2000a). To smooth or not to smooth? – Bias and efficiency in fMRI time-series analysis. *Neuroimage* 12, 196–203.

Friston, K. J., Mechelli, A., Turner, R., and Price, C. J. (2000b). Nonlinear responses in fMRI: the balloon model, volterra kernels, and other hemodynamics. *Neuroimage* 14, 466–477.

Friston, K. J., Penny, W., Phillips, C., Kiebel, S., Hinton, G., and Ashburner, J. (2002). Classical and Bayesian inference in neuroimaging: theory. *Neuroimage* 16, 465–483.

Friston, K. J., Stephan, K. E., Lund, T. E., Morcom, A., and Kiebel, S. (2005). Mixed-effects and fMRI studies. *Neuroimage* 24, 244–252.

Glover, G. (1999). Deconvolution of impulse response in event-related BOLD fMRI. *Neuroimage* 9, 416–429.

Goense, J., and Logothetis, N. (2008). Neurophysiology of the BOLD fMRI signal in awake monkeys. *Curr. Biol.* 18, 631–640.

Grinband, J., Wager, T., Lindquist, M., Ferrera, V., and Hirsch, J. (2008). Detection of time-varying signals in event-related fMRI designs. *Neuroimage* 43, 509–520.

Handwerker, D., Ollinger, J., and D'Esposito, M. (2004). Variation of BOLD hemodynamic responses across subjects and brain regions and their effects on statistical analyses. *Neuroimage* 21, 1639–1651.

Hanson, S., and Bly, B. (2001). The distribution of BOLD susceptibility effects in the brain is non-Gaussian. *Neuroreport* 12, 1971.

Heckman, G. M., Bouvier, S. E., Carr, V. A., Harley, E. M., Cardinal, K. S., and Engel, S. A. (2007). Nonlinearities in rapid event-related fMRI explained by stimulus scaling. *Neuroimage* 34, 651–660.

Henson, R. N. A. (2007). "Efficient experimental design for fMRI," in *Statistical Parametric Mapping: The Analysis of Functional Brain Images*, eds K. J. Friston, J. Ashburner, S. J. Kiebel, T. E. Nichols, and W. D. Penny (London: Academic Press), 193–210.

Holmes, A. P., Blair, R. C., Watson, G., and Ford, I. (1996). Nonparametric analysis of statistic images from functional mapping experiments. *J. Cereb. Blood Flow Metab.* 16, 7–22.

Holmes, A. P., and Friston, K. J. (1998). Generalisability, random effects and population inference. *Neuroimage* 7, S754.

Huettel, S. A., and McCarthy, G. (2001). Regional differences in the refractory period of the hemodynamic response: an event-related fMRI study. *Neuroimage* 14, 967–976.

Jezzard, P., Matthews, P. M., and Smith, S. M. (eds). (2002). *Functional MRI: An Introduction to Methods*. Oxford: Oxford University Press.

Kriegeskorte, N., Goebel, R., and Bandettini, P. (2006). Information-based functional brain mapping. *Proc. Natl. Acad. Sci. U.S.A.* 103, 3863–3868.

Kriegeskorte, N., Mur, M., and Bandettini, P. (2008). Representational similarity analysis – connecting the branches of systems neuroscience. *Front. Syst. Neurosci.* 2:4. doi: 10.3389/neuro.06.004.2008

Kwong, K. K., Belliveau, J. W., Chesler, D. A., Goldberg, I. E., Weisskoff, R. M., Poncelet, B. P., Kennedy, D. N., Hoppel, B. E., Cohen, M. S., Turner, R., Cheng, H.-M., Brady, T. J., and Rosen, B. R. (1992). Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *Proc. Natl. Acad. Sci. U.S.A.* 89, 5675–5679.

Lee, J., Durand, R., Gradinaru, V., Zhang, F., Goshen, I., Kim, D., Fenno, L., Ramakrishnan, C., and Deisseroth, K. (2010). Global and local fMRI signals driven by neurons defined optogenetically by type and wiring. *Nature* 465, 788–792.

Lenoski, B., Baxter, L. C., Karam, L. J., Maisog, J., and Debbins, J. (2008). On the performance of autocorrelation estimation algorithms for fMRI analysis. *IEEE* 2, 828–838.

Liao, C., Worsley, J., Poline, J., Aston, J., Duncan, G., and Evans, A. (2002). Estimating the delay of the fMRI response. *Neuroimage* 16, 593–606.

Lindquist, M. (2008). The statistical analysis of fMRI data. *Stat. Sci.* 23, 439–464.

Lindquist, M., Meng Loh, J., Atlas, L., and Wager, T. (2009). Modeling the hemodynamic response function in fMRI: efficiency, bias and mis-modeling. *Neuroimage* 45, S187–S198.

Lindquist, M., and Wager, T. (2007). Validity and power in hemodynamic response modeling: a comparison study and a new approach. *Hum. Brain Mapp.* 28, 764–784.

Locascio, J. J., Peggy, J., and Corkin, S. (1997). A method of adjusting for temporal and spatial correlations in analysis of mean fMRI signal intensity changes. *Neuroimage* 3, S76.

Logothetis, N. (2008). What we can do and what we cannot do with fMRI. *Nature* 453, 869–878.

Logothetis, N., and Wandell, B. (2004). Interpreting the BOLD signal. *Physiology* 66, 735–769.

Logothetis, N. K. (2002). The neural basis of the blood-oxygen-level-dependent functional magnetic resonance imaging signal. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* 357, 1003–1037.

Logothetis, N. K. (2003). The underpinnings of the bold functional magnetic resonance imaging signal. *J. Neurosci.* 23, 3963–3971.

Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., and Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature* 412, 150–157.

Loh, J. M., Lindquist, M. A., and Wager, T. D. (2008). Residual analysis for detecting mis-modeling in fMRI. *Stat. Sin.* 18, 1421–1448.

Lund, T. E., Madsen, K. H., Sidaros, K., Luo, W. L., and Nichols, T. E. (2006). Non-white noise in fMRI: does modelling have an impact? *Neuroimage* 29, 54–66.

Luo, W.-L., and Nichols, T. E. (2002). *Diagnosis and Exploration of Massively Univariate fMRI Models*. Technical Report, Department of Biostatistics, University of Michigan, Ann Harbor, MI.

Luo, W.-L., and Nichols, T. E. (2003). Diagnosis and exploration of massively univariate neuroimaging models. *Neuroimage* 19, 1014–1032.

Marchini, J. L., and Ripley, B. D. (2000). A new statistical approach to detecting significant activation in functional MRI. *Neuroimage* 12, 366–380.

Miezin, F. M., Maccotta, L., Ollinger, J. M., Petersen, S. E., and Buckner, R. L. (2000). Characterizing the hemodynamic response: effects of presentation rate, sampling procedure, and the possibility of ordering brain activity based on relative timing. *Neuroimage* 11, 735–759.

Mumford, J. A., and Nichols, T. (2009). Simple group fMRI modeling and inference. *Neuroimage* 47, 1469–1475.

Mumford, J. A., and Nichols, T. E. (2006). Modeling and inference of multisubject fMRI data. *IEEE Eng. Med. Biol. Mag.* 25, 42–51.

Mumford, J. A., and Nichols, T. E. (2008). Power calculation for group fMRI studies accounting for arbitrary design and temporal autocorrelation. *Neuroimage* 39, 261–268.

Mumford, J. A., and Poldrack, R. A. (2007). Modeling group fMRI data. *Soc. Cogn. Affect. Neurosci.* 2, 251–257.

Nichols, T. E., and Holmes, A. P. (2001). Non-parametric tests for functional neuroimaging: a primer with examples. *Hum. Brain Mapp.* 15, 1–25.

Ogawa, S., Lee, T. M., Kay, A. R., and Tank, D. W. (1990). Functional brain mapping by blood oxygenation level dependent contrast magnetic resonance imaging: a comparison of signal characters with a biophysical model. *Proc. Natl. Acad. Sci. U.S.A.* 87, 9868–9872.

Ogawa, S., Tank, D. W., Menon, R. S., Ellerman, J. M., Kim, S. G., Merkle, H., and Ugurbil, K. (1992). Intrinsic signal changes accompanying sensory stimulation: functional brain mapping with magnetic-resonance-imaging. *Proc. Natl. Acad. Sci. U.S.A.* 89, 5951–5955.

Ollinger, J., Shulman, G., and Corbetta, M. (2001). Separating processes within a trial in event-related functional MRI: I. The method. *Neuroimage* 13, 210–217.

Palmer, H. (2010). Optogenetic fMRI sheds light on the neural basis of the BOLD signal. *J. Neurophysiol.* 104, 1838.

Penny, W. D., and Holmes, A. P. (2007). "Random-effects analysis," in *Statistical Parametric Mapping: The Analysis of Functional Brain Images*, eds K. Friston, J. Ashburner, S. Kiebel, T. Nichols, and W. Penny (London: Academic Press), 156–165.

Petersson, K. M., Nichols, T. E., Poline, J. B., and Holmes, A. P. (1999). Statistical limitations in functional neuroimaging. I. non inferential methods and statistical models. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* 354, 1239–1260.

Purdon, P., Solo, V., Weisskoff, R. M., and Brown, E. N. (2001). Locally regularized spatiotemporal modeling and model comparison for functional MRI. *Neuroimage* 14, 912–923.

Purdon, P., and Weisskoff, R. (1998). Effect of temporal autocorrelation due to physiological noise and stimulus paradigm on voxel-based false-positive rates. *Hum. Brain Mapp.* 6, 239–249.

Razavi, M., Grabowski, T. J., Vispoel, W. P., Monahan, P., Mehta, S., Eaton, B., and Bolinger, L. (2003). Model assessment and model building in fMRI. *Hum. Brain Mapp.* 20, 227–238.

Robson, M. D., Dorosz, J. L., and Gore, J. C. (1998). Measurements of the temporal fMRI response of the human auditory cortex to trains of tones. *Neuroimage* 7, 185–198.

Roy, C. S., and Sherington, C. S. (1890). The regulation of the blood supply of the brain. *J. Physiol.* 11, 85–108.

Smith, M., Beckmann, C. F., Ramnani, N., Woolrich, M. W., Bannister, P. R., Jenkinson, M., Matthews, P. M., and McGonigle, D. J. (2005). Variability in fMRI: a re-examination of inter-session differences. *Hum. Brain Mapp.* 24, 248–257.

Smith, S., Jenkinson, M., Beckmann, C., Miller, K., and Woolrich, M. (2007). Meaningful design and contrast estimability in fMRI. *Neuroimage* 34, 127–136.

Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E., Johansen-Berg, H., Bannister, P. R., Luca, M. D., Drobnjak, I., Flitney, D. E., Niazy, R. K., Saunders, J., Vickers, J., Zhang, Y., Stefano, N. D., Brady, J.-M., and Matthews, P. M. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* 23, S208–S219.

Soltysik, D. A., Peck, K. K., White, K. D., Crosson, B., and Briggs, R. W. (2004). Comparison of hemodynamic response nonlinearity across primary cortical areas. *Neuroimage* 22, 1117–1127.

Stefanovic, B., Warnking, J., and Pike, G. (2004). Hemodynamic and metabolic responses to neuronal inhibition. *Neuroimage* 22, 771–778.

Steffener, J., Tabert, M., Reuben, A., and Stern, Y. (2010). Investigating hemodynamic response variability at the group level using basis functions. *Neuroimage* 49, 2113–2122.

Stephan, K., Harrison, L., Penny, W., and Friston, K. (2004). Biophysical models of fMRI responses. *Curr. Opin. Neurobiol.* 14, 629–635.

Talairach, J., and Tournoux, P. (eds). (1988). *Co-plannar Stereotaxis Atlas of the Human Brain*. New York, NY: Thieme.

Thirion, B., Pinel, P., Mériaux, S., Roche, A., Dehaene, S., and Poline, J.-B. (2007). Analysis of a large fMRI cohort: statistical and methodological issues for group analyses. *Neuroimage* 35, 105–120.

Thomsen, K., Offenhauser, N., and Lauritzen, M. (2004). Principal neuron spiking: neither necessary nor sufficient for cerebral blood flow in rat cerebellum. *J. Physiol.* 560, 181–189.

Ugurbil, K., Adriany, G., Andersen, P., Chen, W., Garwood, M., Gruetter, R., Henry, P., Kim, S., Lieu, H., Tkac, I., Vaughan, T., Van De Moortele, P. F., Yacoub, E., and Zhu, X. H. (2003). Ultrahigh field magnetic resonance imaging and spectroscopy. *Magn. Reson. Imaging* 21, 1263–1281.

Vazquez, A. L., and Noll, D. C. (1998). Nonlinear aspects of the BOLD

response in functional MRI. *Neuroimage* 7, 108–118.

Wager, T. D., and Smith, E. E. (2003). Neuroimaging studies of working memory: a meta-analysis. *Cogn. Affect. Behav. Neurosci.* 3, 255–274.

Wager, T. D., Vazquez, A., Hernandez, L., and Noll, D. C. (2005). Accounting for nonlinear BOLD effects in fMRI: parameter estimates and a model for prediction in rapid event-related studies. *Neuroimage* 25, 206–218.

Waldorp, L. (2009). Robust and unbiased variance of GLM coefficients for misspecified autocorrelation and hemodynamic response models in fMRI. *Int. J. Biomed. Imaging* 2009, 1–11.

Weisskoff, R. M., Baker, J., Belliveau, J., Davis, T., Kwong, K. K., Cohen, M. S., and Rosen, B. (1993). Power spectrum analysis of functionally-weighted MR data: what's in the noise? *Proc. Soc. Magn. Reson. Med.* 1, 7.

Woolrich, M., Behrens, T., and Smith, S. (2004). Constrained linear basis sets for HRF modelling using variational Bayes. *Neuroimage* 21, 1748–1761.

Woolrich, M. W., Ripley, B. D., Brady, J. M., and Smith, S. M. (2001). Temporal autocorrelation in univariate linear modelling of fMRI data. *Neuroimage* 14, 1370–1386.

Worsley, K. J., and Friston, K. J. (1995). Analysis of time-series revisited – again. *Neuroimage* 2, 173–181.

Worsley, K. J., Liao, C. H., Aston, J., Petre, V., Duncan, G. H., Morales, F., and Evans, A. C. (2002). A general statistical analysis for fMRI data. *Neuroimage* 15, 1–15.

Zhang, N., Zhu, X., and Chen, W. (2008). Investigating the source of nonlinear BOLD response to ultra-short visual stimulus. *Neuroimage* 43, 204–212.

**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 June 2010; accepted: 06 March 2011; published online: 18 March 2011.*
*Citation: Monti MM (2011) Statistical analysis of fMRI time-series: a critical review of the GLM approach. Front. Hum. Neurosci.* ***5****:28. doi: 10.3389/fnhum.2011.00028*