PERCEPTION, ACTION, AND COGNITION

EDITED BY : Snehlata Jaswal
PUBLISHED IN : Frontiers in Psychology

frontiers Research Topics

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**

PERCEPTION, ACTION, AND COGNITION

Topic Editor:
Snehlata Jaswal, Indian Institute of Technology Jodhpur, India

Even as simple a task as quenching thirst with a glass of water involves a sequence of perceptions and actions woven together by expectations and experience. What are the myriad links between perception and action, and what does cognition have to do with them? Intuitively we think that perception precedes action, but we also know that action moulds perception. The reciprocal links between perception and action are now accepted almost universally. The discovery of mirror neurons that encode observed actions has further emphasized the coupling of perception and action.

Ball games epitomize the close connections between perception, action, and cognition. Image by Snehlata Jaswal

The real aim of this research topic is to go beyond identifying the evidence for perception-action coupling, and study the cognitive entities and processes that influence the perception-action link. For example, the internal representations of perceived and produced events are created and modified through experience. Yet the perception action link is considered relatively automatic. To what extent is the perception-action link affected by representations and their manipulations by cognitive processes? Does selective attention modify the perception action coupling? How, and to what extent, does the context provide sources of cognitive control? The developmental trajectory of the perception-action link and the influence of cognition at various stages of development could be another line of important evidence. The responses to these and other such questions contribute to our understanding of this research area with significant implications for perception-action coupling.

Table of Contents

Editorial: The Balanced Triad of Perception, Action, and Cognition

Snehlata Jaswal *

Psychology, Cognitive Science, Indian Institute of Technology Jodhpur, Jodhpur, India

The Editorial on the Research Topic

Perception, Action, and Cognition

Conceptual distinctions between perception, action, and cognition are convoluted in real world proceedings, as we meander through life, tied to past thoughts and actions, oriented toward future goals, guided by current perceptions. The research topic "Perception, Action, and Cognition" aimed to go beyond the already established perception-action links to study the role of cognitive mechanisms in this triad.

A popular way of conceptualizing the triad of perception, action, and cognition is the idea of embodied cognition. The first section entitled "Embodied cognition" presents seven articles related to this concept. In the leading article, Hommel develops an efficient way to conceptualize embodied cognition using the Theory of Event Coding (TEC) framework. Arguing against the anti-cognitivist stance of many embodiment theorists, he firmly puts cognition back in the perception-action link, by maintaining the importance and involvement of internal representations in the production of actions. Vernon et al. propose that perception-action coupling is not only manifest in the behavioral arena, but also shows up in the internal processes of the agents, particularly those related to the self. The self-organizing, self-producing, and self-maintaining processes explain the reciprocity of perception and action. The sense of presence, which is the focus of the article by Triberti and Riva actually implies the presence of the "self" in the environment where perception-action coupling is manifest. Brizio and Tirassa go a step further to expound on how the mind is rooted in the self, particularly the biological self. They offer a taxonomy of control systems based on whether they are intentional, non-intentional, or meta-intentional, drawing on arguments from diverse disciplines ranging from psychology to biology to philosophy to artificial intelligence. The next two articles provide evidence for embodied cognition from clinical samples. Dreyer et al. report evidence from two patients with focal lesions, regarding sensorimotor systems in the cortex being crucial for the processing of semantic concepts, suggesting that without intact perceptual-motor systems, meaningful cognition is all but impossible. Wolpe et al. show that Parkinsons' Disease (PD) patients with higher levodopa dose equivalent show an abnormally high awareness of their actions and their positive outcomes, providing evidence for intact action systems feeding perceptual awareness. Finally, evidence regarding action being important for cognition also comes from a cross cultural study by Wang et al. who studied children's use of imitation in learning how to categorize objects. They propose that imitation of an action leads to direct experience, which in turn stimulates rule learning/categorization. This process is the same across the two cultures studied—Chinese and US.

The rest of the articles are ingenious forays into the socio-cultural realm. The second section of the research topic is, therefore, titled "The external world and inner reality." We begin the section with mirror mechanisms in the brain, the biological underpinnings of social perception and interactions. Volta et al. report an fMRI study that showed similar brain areas to be activated whether participants were actually walking or merely observing someone else walking,

thus supporting the idea of mirror mechanisms for walking. However, differences in brain areas activated were evident between two conditions—one in which a video of open countryside was played and the other in which the closed space of a corridor was depicted, indicating the strong influence of external context on brain processing. Desanghere and Marotta show that the physical shape of an object and its center of mass, are distinct influences, in how we look at it, and grasp it. This study independently manipulated physical shape and center of mass of an object, and thus used irregularly shaped objects to study the effect of these physical features on gaze and grasp. This study is a good illustration of how the same percept may lead to different reactions in different parts of the body perhaps because the relevance of the percept to the body part is different. In another study demonstrating the effect of perceptual cues on Simon task performance, Umemura demonstrates that stimulus locations in 2D and in 3D space give independent cues for task performance to the participants. This seems understandable; for pictorial cues are sufficient for 2D locations, but 3D locations can be precisely pinpointed only by the convergence of the eyes. Studying an important variable in our social interactions, Kuraguchi and Ashida focus on detection of beauty and cuteness. They found that whilst beauty can be equally discerned in central as well as peripheral vision, the detection of cuteness declines in peripheral vision as compared to central vision, especially in males. In another study of face processing, it becomes clear why autism is associated with difficulties in the socio-emotional sphere. Ujiie et al. found that university students scoring high on Autism Spectrum Quotient, a measure of autistic traits, show abnormal audio-visual speech integration, probably because they do not show adequate processing of global facial configurations. All these studies show how changes in the external environment result in behavioral/ motor changes.

In turn, the influence of actions or their outcomes on socio-cognitive processing is clear from the next two studies. Sun et al. compared decision times on a simple task presented on the winners' or the losers' side of the screen. Results showed an attention bias in the onlookers for the losers' side of the screen. Could this be the basis of empathy/sympathy in social situations? Using a Rapid Serial Visual Presentation paradigm, Kihara et al. demonstrate that when participants were asked to change a digit stream to a letter stream by pressing a button and identify four successive targets, successful target identification was linked to pupil dilation (a measure of the involvement of the locus coeruleus-noradrenaline system at the neural level). Thus, voluntary action initiated this neural substrate of transient visual attention.

The "self" emerges as a strong player in this section as well. Spape et al. present a new paradigm utilizing "avatars" or representations of the self in virtual reality, which can be used by researchers in cognition, social psychology, and human computer interaction. The study also augments our conceptualization of executive control, by showing how changes in the gender of the "avatars" can disrupt conflict adaptation in the virtual reality version of the Simon task. Zhang et al. use the virtual hand illusion to demonstrate that the bodily self is not a definite exact entity, rather, body representations are dynamic and are constantly updated by current inputs from the context. Imbir demonstrates the effect of subjective significance on the rational mind, and arousal on the experiential mind, using the emotional stroop task. The control of "reason," therefore, is dependent on whether we perceive a situation to be of significant or not. Finally, the review by Kavanagh and Winkielman suggests mimicry as an implicit but adaptive mechanism that underlies affiliation between the model and the mimic, and also signals affiliation with the group.

All these articles are heartening examples of work that goes much beyond establishing the perception-action link. Most contributions specify the mechanisms for this link, or the antecedent or consequent conditions of the link. For example, the articles in the first section propose the self as the unifying entity for perception and action. In the second, whereas some articles study the effect of perceptual factors on behavior, others specify how action influences attention/perception, and still others focus on the mediator. i.e., person involved. A perusal of all articles in this research topic reveals that perception, action, and cognition, are an unbreakable triad. Many manifest links have been established, and many other latent ones need to be explored in future research.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and approved it for publication.

The theory of event coding (TEC) as embodied-cognition framework

Bernhard Hommel*

Cognitive Psychology Unit, Institute of Psychology, Leiden Institute for Brain and Cognition, Leiden University, Leiden, Netherlands

The concept of embodied cognition attracts enormous interest but neither is the concept particularly well-defined nor is the related research guided by systematic theorizing. To improve this situation the theory of event coding (TEC) is suggested as a suitable theoretical framework for theorizing about cognitive embodiment—which, however, presupposes giving up the anti-cognitivistic attitude inherent in many embodiment approaches. The article discusses the embodiment-related potential of TEC, and the way and degree to which it addresses Wilson's (2002) six meanings of the embodiment concept. In particular, it is discussed how TEC considers human cognition to be situated, distributed, and body-based, how it deals with time pressure, how it delegates work to the environment, and in which sense it subserves action.

Keywords: embodied cognition theory, cognitivism, perception and action, distributed cognition, human cognition, theory, perception for action

The general intuition that human cognition is somehow "embodied" is widely shared and has stimulated various new brands of research. And yet, the concept of embodied cognition is ill-defined and a general, testable theory of its underlying mechanisms has not yet been presented. This has rendered tests of the concept difficult and often metaphorical in nature, which stands in the way of broader acceptance. The main reason for that, so I claim, is the anti-cognitivistic attitude of most embodiment approaches. Not only is this attitude more rooted in ideology than in empirical data, but it also prevents embodiment approaches from using cognitivistic tools to build mechanistic theories that provide the badly needed testable hypotheses. In the following, I shall argue that some cognitivistic approaches are well-equipped to capture the essence of cognitive embodiment. In particular, I shall argue that the theory of event coding (TEC; Hommel et al., 2001a) provides almost all that embodiment theories need, and that it therefore provides a mechanistic approach to the embodied-cognition concept—which TEC fully embraces. I shall demonstrate that by going through the six different meanings of embodied cognition that Wilson (2002) has identified in the literature, and explain how TEC fits with (the unideological aspects of) these six meanings.

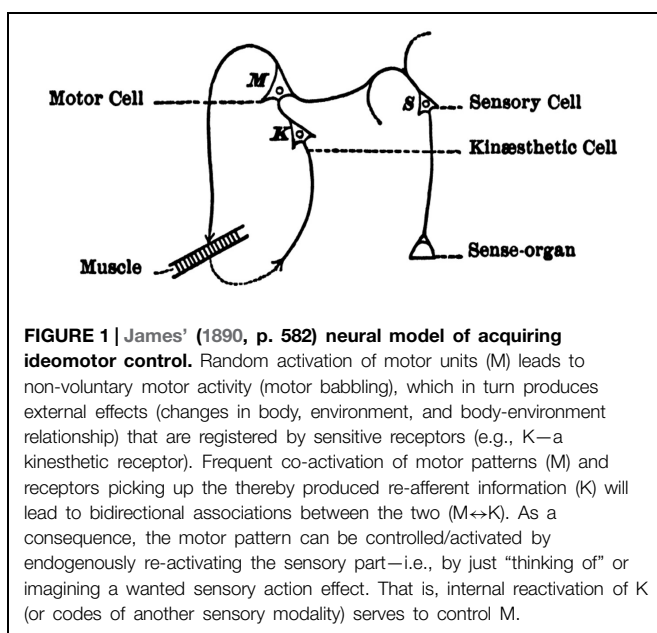## The Theory of Event Coding

Theory of event coding is rooted in the cognitivistic ideomotor approaches of Lotze (1852), Harless (1861), and James (1890), and yet embraces the idea that human cognition emerges from sensorimotor processing. In contrast to behavioristic or information-processing approaches, ideomotor theory considers humans as active agents that perform actions to reach particular

goals. Accordingly, the theoretical analysis does not start with stimuli but with goals (intended action effects), which are assumed to trigger the execution of movements suited to reach them. Goals are acquired by actively exploring the environment, which creates associations between motor activities and representations of their perceptual consequences (Elsner and Hommel, 2001). These action-effect bindings provide the basis for voluntary action: the agent only needs to "think of" the representation of a wanted action effect to activate the motor pattern needed to produce it (for overviews, see Hommel, 2009; Shin et al., 2010)—see **Figure 1**. Indeed, planning to produce particular action outcomes (e.g., hand movements or facial expressions) activates the neural codes of these consequences (areas EBR and FFA, respectively) before execution begins (Kühn et al., 2011).

Theory of event coding combines the ideomotor mechanism with assumptions about how perceptual and action events are represented. In particular, TEC comprises of four basic assumptions (Hommel et al., 2001a): (I) Perceptual events and planned actions are cognitively represented by *event codes*; (II) which are integrated assemblies of *feature codes*; (III) which in turn are cognitive/brain states correlated with external (perceived or self-generated) features (*distal coding*); (IV) so that the basic units of both perception and action (assemblies of feature codes) are *sensorimotor* entities, in the sense that they are activated by sensory input (=perception) and controlling motor output (=action).

What does it mean, from a TEC point of view, that cognition is embodied? As mentioned already, this depends on which sense of cognitive embodiment one refers to. According to Wilson (2002), at least six different meanings can be distinguished. In the following, I shall briefly introduce each of these criteria, carve out their less ideological essence, and show how they are met by TEC. For a broader, more detailed discussion, see Hommel (2015).

FIGURE 1 | James' (1890, p. 582) neural model of acquiring ideomotor control. Random activation of motor units (M) leads to non-voluntary motor activity (motor babbling), which in turn produces external effects (changes in body, environment, and body-environment relationship) that are registered by sensitive receptors (e.g., K—a kinesthetic receptor). Frequent co-activation of motor patterns (M) and receptors picking up the thereby produced re-afferent information (K) will lead to bidirectional associations between the two (M↔K). As a consequence, the motor pattern can be controlled/activated by endogenously re-activating the sensory part—i.e., by just "thinking of" or imagining a wanted sensory action effect. That is, internal reactivation of K (or codes of another sensory modality) serves to control M.

Situated Cognition

The idea that cognitive activity is always situated and taking place in a particular context comes in two different flavors (Wilson, 2002). One relates to the philosophical/pedagogical claim that knowing is inseparable from doing, which is why education should favor learning-by-doing over the passive accumulation of knowledge (e.g., Greeno, 1998). This approach emphasizes the role of active agency in knowledge acquisition and internal representation—exactly as proposed by ideomotor theory and TEC. The ideomotor approach does not merely presume the existence of action goals or particular knowledge to account for voluntary action, but explains how goals are acquired through concrete hands-on exploration of one's own body and its interactions with the environment (Verschoor et al., 2010). TEC extends this mechanism to a general theory of human cognition and claims that humans do not just acquire action goals through sensorimotor experience but that most or all knowledge is rooted in sensorimotor experience. It therefore seems to fit perfectly with the philosophical/pedagogical embodiment concept.

The other flavor of situated cognition comes from cognitive robotics (e.g., Clancey, 1997) and claims that agents may often not require internal information but can simply pick it up from their environment—which obviously borrows from Gibsonian Ecological Psychology and the assumption that environments provide action-related affordances for the active perceiver (Gibson, 1979). On the one hand, it is uncontroversial that humans and other primates possess (dorsal) online information-processing channels that feed more or less directly into action systems (Milner and Goodale, 2006), which may be considered a system that processes affordances (Michaels, 2000). On the other hand, however, there is no evidence that voluntary actions can be planned and carried out successfully without the help of other (ventral) off-line channels devoted to object identification, planning, and action evaluation—processes that rely on memory, which even in the case of procedural knowledge can be considered to represent past experiences.

It is this second off-line system responsible for setting up and planning voluntary actions that TEC is concerned with, while the theory has little to say about the subserving online channels (Hommel et al., 2001b). However, almost all of our everyday actions rely on previously acquired knowledge about how to use the available tools to reach our goals—just think of using a computer or coffee machine, or engaging in verbal communication and socially appropriate behavior. And they are often planned ahead in the absence of situational cues—which falls outside of the scope of (neo-) classical affordance-based approaches. As many actions require, or benefit from, the integration of knowledge-dependent off-line processes, and environmentally driven online processes, they can be considered goal-directed and context-sensitive at the same time. Accordingly, it makes sense to try integrating cognitivistic ideomotor theory and affordance-based approaches (Glover, 2004) rather than putting them into opposition—as some situated-cognition approaches tend to do (e.g., Pfeifer and Bongard, 2006).

Cognition under Time Pressure

This brand of the embodied-cognition concept assumes that engaging in cognitive activities is particularly time costly and therefore unlikely to be the basis of everyday action. In cognitive robotics, this idea has been taken to suggest dropping the cognitive overhead so to allow robots to meet real-time constraints (e.g., Pfeifer and Scheier, 1999), and reasoning theorists (e.g., Gigerenzer et al., 1999) have used it to argue that people prefer cognitive shortcuts over full-fledged cognitive analyses of a problem.

One can argue whether time pressure is a real problem in humans: not only are there few everyday situations that would not allow taking the time for fuller cognitive analysis, but nature has also equipped us with reflexes that allow engaging in fight or flight long before grasping the actual situational demands. Hence, survival seems possible with the human mix of slow cognition and fast reflexes.

But what is more, theoretical analysis and empirical evidence suggest that slow cognition does not intervene between perception and action even if time allows (Hommel, 2013). Indeed, the main purpose of cognition seems to consist in anticipatory (offline) preparation: selecting a goal, configuring the system for a particular task, priming goal-relevant action systems, and preparing for the processing of possible trigger stimuli. It is these preparatory processes that TEC is targeting. Once an action has been selected and sufficiently prepared, not much cognitive activity seems to go on and environmental information will commonly suffice to drive the action to completion—a kind of prepared reflex (Hommel, 2000). If, thus, cognition is not for online control but for off-line preparation, the possible slowness of cognitive processes does not serve as an argument against cognitivistic approaches and not even against (properly programmed) cognitive robots.

Offloading Cognitive Work onto the Environment

This blend of the embodiment approach tries to reduce the amount of knowledge that agents need to process. It assumes that the environment can serve as its own memory (e.g., Brooks, 1991; O'Regan and Noe, 2001), so that agents do not need to develop internal world models. Interestingly, supporting evidence comes exclusively from spatial tasks (Wilson, 2002), and it certainly makes sense that spatial decisions consider available spatial information. It is less clear how offloading might work with actions like talking, dancing, or writing an article, but this may be the reason why the available evidence is restricted to less knowledge-heavy spatial actions. In any case, the model-less embodiment approach is fully consistent with TEC, which does not assume that actions rely on world models. While TEC aims to explain how people are planning goal-directed actions (based on procedural, implicit knowledge gathered through active experience), it does not claim that all aspects of actions are predetermined by planning. It rather assumes that planning is restricted to the specification of goal-relevant action

outcomes (e.g., of the cup to be grasped for drinking), while the specification of goal-unrelated action features (e.g., the particular kinematics) is left to environmentally driven online channels that continuously feed in information during action execution (Hommel, 2010).

Distributed Cognition

The claim that human cognition is not restricted to an individual's mind and brain but involves the environment as well (e.g., Wilson and Golonka, 2013) can be considered another revival of a theme with a much older history (e.g., interactionism in personality psychology). According to Wilson (2002) the claim actually consists of two parts: that including the environment in analyses of human cognition provides more information than excluding it—which is true but too trivial to be controversial— and that excluding the environment from the analysis does not allow for any interesting insight into human cognition in principle.

It must be said that the distributed-cognition criterion has not yet been supported by any specific evidence. Neither has it been defined which aspects of the environment actually count (e.g., if a symbol on the monitor is enough, almost all cognitive research does take the environment into account), nor has it been demonstrated that, and in which sense the available cognitive/neurocognitive research has failed to produce meaningful results, nor has the approach itself produced any specific evidence to its own support—all the evidence that proponents tend to discuss was motivated by other than the distributed-cognition approach (e.g., see Wilson and Golonka, 2013). Hence, even the few observations in the literature that distributed-cognition proponents do find relevant did not require the distributed-cognition approach to make them.

However, a more liberal interpretation of the approach might get close to the situated-cognition criterion, and thus rightly attract attention to the relevance of concrete sensorimotor experience of the agent with his or her environment. This approach would make TEC and its reliance on sensorimotor experience a valuable tool to go beyond abstract complaints, and allow for the empirical test of concrete hypotheses about concrete phenomena.

Cognition Subserves Action

The claim that human cognition evolved to subserve action has considerable Darwinistic face value: evolution operates on actions, not on ideas. It is therefore not surprising that cognition-for-action has been a dominant theme in many approaches, including American pragmatism, behaviorism, Russian activity theory, the motor theory of speech, and the mirror neuron approach. The concept also represents the very core of TEC, which partly reflects its intellectual heritage. And yet, the architectural and process-related assumptions underlying TEC make it unique in the field in a number of ways.

Theory of event coding shares the idea of cognition-for-action with respect to phylogenetic and ontogenetic considerations:

the neural/functional architecture underlying the distribution of labor between dorsal and ventral information-processing streams is likely to reflect the importance of action in evolution and, as explained already, ontogenetic cognitive development relies on sensorimotor experience. However, TEC differs from other approaches in denying that every single use of cognitive skill or content must be accompanied by sensorimotor activity—as embodied-cognition proponents like Gallese and Goldman (1998) or Barsalou (2008) suggest (which is not to deny that these approaches share many other aspects with TEC).

One reason why cognition without sensorimotor activity or mental simulation should be possible relates to TECs intentional-weighting principle (Memelink and Hommel, 2013). TEC assumes that perceived or to-be-produced events are represented by means of event files—integrated bindings of the codes of distal event features (Hommel, 2004). The contribution of each component of such a binding is assumed to be weighted according to its situational relevance, so that codes related to the ringing sound of a telephone will be weighted more strongly than codes related to its color when waiting for a call. This means that event representations are tailored to the goal and task at hand, which implies that not all components of the representation are sufficiently activated to contribute to cognition and action. This means that, even though most cognitive representations are likely to comprise of both perceptual and action components, cognitive operations using these representations are possible without above-threshold activation of some components. If, thus, the cognitive operations do not require overt action (e.g., when processing words for silent reading), it is possible that they are carried out without measurable motor activity. In other words, sensorimotor activity is important for creating cognitive representations but not necessarily for using them.

The other reason is that grounding basic cognitive units in sensorimotor experience, as TEC assumes, does not necessarily prevent the creation of other representations that refer to combinations of, or relations between, such basic units. For instance, there is no reason to exclude that people are able to combine the (sensorimotorically grounded) representation of a horse with the (sensorimotorically grounded) representation of a horn to create the representation of a unicorn without ever having sensorimotorically experienced one. According to TEC, the resulting representations may be considered abstract but not necessarily symbolic or arbitrary, and they still can be considered as being grounded in sensorimotor experience.

Body-Based Cognition for Off-Line Use

The claim refers to the idea that cognitive structures or skills that emerged through sensorimotor interactions with the environment could be used off-line—in the absence of overt behavior—to subserve cognitive activities (e.g., Glenberg, 1997). This claim also has a rather long history, especially in Russian activity theory (e.g., Vygotsky, 1962) and approaches

that conceive of cognition as interiorized action. TEC fails to provide a systematic scenario of how interiorization might work in detail (and there is in fact no such theory available), but it does provide the necessary cognitive infrastructure. As explained already, TEC assumes that overt sensorimotor action leads to the binding of motor patterns and codes representing their consequences. Acquiring these bindings allows the agent to run them internally to simulate the action without actually activating the motor patterns (if intentional weighting deactivates the motor components). The acquisition of multiple sensorimotor events allows the agent to construct more complex event sequences, such as for making coffee (Kachergis et al., 2014). These representations provide information about how to move from one situation to another to reach a distant goal, which can be used to simulate and to compare alternative problem-solving strategies.

Conclusion

Many embodied-cognition approaches have been put forward with a strong anti-cognitivistic attitude that they share with, and in some cases borrow from, behaviorism, ecological psychology, and evolutionary psychology, and indeed many of the ecological, and evolutionary arguments have resurfaced in the embodied-cognition debate (e.g., see Wilson and Golonka, 2013). This is unfortunate for two reasons. For one, most of these arguments are simply misdirected: they mainly challenge the symbol-heavy good old-fashioned artificial intelligence, which, however, had only negligible impact on modern cognitive psychology/neuroscience. And, for another, rejecting cognitivism prevents embodied-cognition theorists to develop mechanistic and therefore testable models that allow them to explore how fruitful the embodiment approach actually is. Without such models, no direct comparison with non-embodied approaches is possible, and thus no competition for the better explanation on cognitive psychology's to-do list. Cherry-picking examples that minimize cognitive processes, knowledge, and internal preparation may help to illustrate basic principles, but scientific approaches to human cognition that are unable to account for everyday actions like making a phone call, preparing coffee, or asking for directions are unlikely to succeed.

I propose that a more constructive approach to embodied cognition is possible and probably more successful. As I have argued, a cognitivistic approach to the relationship between perception and action is likely to be useful for that purpose. In particular, the TEC is theoretically commensurable with all six of Wilson's (2002) meanings of the embodiment concept. TEC assumes the existence of internal representations and claims that such representations are involved in producing actions, which makes it a cognitivistic approach. At the same time, it does not only assume that human cognition is situated, distributed, and body-based, but also explains how; it also explains how cognition deals with time pressure, how it is delegated to the environment, and which sense it subserves action. Extensions of the theory are

possible and wanted on the way to a truly comprehensive framework of human cognition, and combining TEC with mechanistic models of online/affordance-based control and with assumptions about interiorization seems particularly promising in this respect.

References

Barsalou, L. W. (2008). Grounded cognition. *Annu. Rev. Psychol.* 59, 617–645. doi: 10.1146/annurev.psych.59.103006.093639

Brooks, R. A. (1991). Intelligence without representation. *Artif. Intell.* 47, 139–159. doi: 10.1016/0004-3702(91)90053-M

Clancey, W. J. (1997). *Situated Cognition: On Human Knowledge and Computer Representation*. New York, NY: Cambridge University Press.

Elsner, B., and Hommel, B. (2001). Effect anticipation and action control. *J. Exp. Psychol. Hum. Percept. Perform.* 27, 229–240. doi: 10.1037/0096-1523.27.1.229

Gallese, V., and Goldman, A. I. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends Cogn. Sci.* 2, 493–551. doi: 10.1016/S1364-6613(98)01262-5

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Boston, MA: Houghton Mifflin.

Gigerenzer, G., Todd, P. M., and the ABC Research Group. (1999). *Simple Heuristics that Make us Smart*. New York, NY: Oxford University Press.

Glenberg, A. M. (1997). What memory is for. *Behav. Brain Sci.* 20, 1–55. doi: 10.1017/s0140525x97000010

Glover, S. (2004). Separate visual representations in the planning and control of action. *Behav. Brain Sci.* 27, 3–24. doi: 10.1017/s0140525x04000020

Greeno, J. G. (1998). The situativity of knowing, learning, and research. *Am. Psychol.* 53, 5–26. doi: 10.1037/0003-066X.53.1.5

Harless, E. (1861). Der apparat des Willens. *Z. Philos. Philos. Kritik* 38, 50–73.

Hommel, B. (2000). "The prepared reflex: automaticity and control in stimulus-response translation," in *Control of Cognitive Processes: Attention and Performance*, Vol. 18, eds S. Monsell and J. Driver (Cambridge: MIT Press), 247–273.

Hommel, B. (2004). Event files: feature binding in and across perception and action. *Trends Cogn. Sci.* 8, 494–500. doi: 10.1016/j.tics.2004.08.007

Hommel, B. (2009). Action control according to TEC (theory of event coding). *Psychol. Res.* 73, 512–526. doi: 10.1007/s00426-009-0234-2

Hommel, B. (2010). "Grounding attention in action control: the intentional control of selection," in *Effortless Attention: A New Perspective in the Cognitive Science of Attention and Action*, ed. B. J. Bruya (Cambridge, MA: MIT Press), 121–140. doi: 10.7551/mitpress/9780262013840.003.0006

Hommel, B. (2013). Dancing in the dark: no role for consciousness in action control. *Front. Psychol.* 4:380. doi: 10.3389/fpsyg.2013.00380

Hommel, B. (2015). "Embodied cognition according to TEC," in *Foundations of Embodied Cognition*, eds Y. Coello and M. Fischer (New York, NY: Psychology Press).

Hommel, B., Müsseler, J., Aschersleben, G., and Prinz, W. (2001a). The theory of event coding (TEC): a framework for perception and action planning. *Behav. Brain Sci.* 24, 849–937. doi: 10.1017/S0140525X01000103

Hommel, B., Müsseler, J., Aschersleben, G., and Prinz, W. (2001b). Codes and their vicissitudes. *Behav. Brain Sci.* 24, 910–937. doi: 10.1017/S0140525X01520105

James, W. (1890). *The Principles of Psychology*, Vol. 2. New York, NY: Dover Publications. doi: 10.1037/10538-000

Kachergis, G., Wyatte, D., O'Reilly, R. C., de Kleijn, R., and Hommel, B. (2014). A continuous time neural model for sequential action. *Philos. Trans. R. Soc. B* 369, 20130623. doi: 10.1098/rstb.2013.0623

Kühn, S., Keizer, A., Rombouts, S. A. R. B., and Hommel, B. (2011). The functional and neural mechanism of action preparation: roles of EBA and FFA in voluntary action control. *J. Cogn. Neurosci.* 23, 214–220. doi: 10.1162/jocn.2010.21418

Lotze, R. H. (1852). *Medicinische Psychologie Oder die Physiologie der Seele*. Leipzig: Weidmann'sche Buchhandlung.

Memelink, J., and Hommel, B. (2013). Intentional weighting: a basic principle in cognitive control. *Psychol. Res.* 77, 249–259. doi: 10.1007/s00426-012-0435-y

Michaels, C. F. (2000). Information, perception, and action: what should ecological psychologists learn from Milner and Goodale (1995)? *Ecol. Psychol.* 12, 241–258. doi: 10.1207/s15326969eco1203_4

Milner, A. D., and Goodale, M. A. (2006). *The Visual Brain in Action*, 2nd Edn. Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780198524724.001.0001

O'Regan, J. K., and Noe, A. (2001). A sensorimotor account of vision and visual consciousness. *Behav. Brain Sci.* 24, 939–1031. doi: 10.1017/S0140525X01000115

Pfeifer, R., and Bongard, J. (2006). *How the Body Shapes the Way we Think: A New View of Intelligence*. Cambridge MA: MIT Press.

Pfeifer, R., and Scheier, C. (1999). *Understanding Intelligence*. Cambridge, MA: MIT Press.

Shin, Y. K., Proctor, R. W., and Capaldi, E. J. (2010). A review of contemporary ideomotor theory. *Psychol. Bull.* 136, 943–974. doi: 10.1037/a0020541

Verschoor, S. A., Weidema, M., Biro, S., and Hommel, B. (2010). Where do action goals come from? Evidence for spontaneous action-effect binding in infants. *Front. Psychol.* 1:201. doi: 10.3389/fpsyg.2010.00201

Vygotsky, L. S. (1962). *Thought and Language*. Cambridge MA: MIT Press. doi: 10.1037/11193-000

Wilson, A. D., and Golonka, S. (2013). Embodied cognition is not what you think it is. *Front. Psychol.* 4:58. doi: 10.3389/fpsyg.2013.00058

Wilson, M. (2002). Six views of embodied cognition. *Psychon. Bull. Rev.* 9, 625–636. doi: 10.3758/BF03196322

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

Embodied cognition and circular causality: on the role of constitutive autonomy in the reciprocal coupling of perception and action

David Vernon[1]*, Robert Lowe[1,2], Serge Thill[1] and Tom Ziemke[1,3]

[1] Interaction Lab, School of Informatics, University of Skövde, Skövde, Sweden, [2] Division of Cognition and Communication, University of Gothenburg, Gothenburg, Sweden, [3] Human-Centered Systems, Department of Computer and Information Science, Linköping University, Linköping, Sweden

The reciprocal coupling of perception and action in cognitive agents has been firmly established: perceptions guide action but so too do actions influence what is perceived. While much has been said on the implications of this for the agent's external behavior, less attention has been paid to what it means for the internal bodily mechanisms which underpin cognitive behavior. In this article, we wish to redress this by reasserting that the relationship between cognition, perception, and action involves a constitutive element as well as a behavioral element, emphasizing that the reciprocal link between perception and action in cognition merits a renewed focus on the system dynamics inherent in constitutive biological autonomy. Our argument centers on the idea that cognition, perception, and action are all dependent on processes focussed primarily on the maintenance of the agent's autonomy. These processes have an inherently circular nature—self-organizing, self-producing, and self-maintaining—and our goal is to explore these processes and suggest how they can explain the reciprocity of perception and action. Specifically, we argue that the reciprocal coupling is founded primarily on their endogenous roles in the constitutive autonomy of the agent and an associated circular causality of global and local processes of self-regulation, rather than being a mutual sensory-motor contingency that derives from exogenous behavior. Furthermore, the coupling occurs first and foremost via the internal milieu realized by the agent's organismic embodiment. Finally, we consider how homeostasis and the related concept of allostasis contribute to this circular self-regulation.

Keywords: embodied cognition, autonomy, agency, circular causality, homeostasis, allostasis

1. INTRODUCTION

The reciprocal coupling of perception and action in cognitive agents[1] is now well accepted and there are many examples from neuroscience and psychology, e.g., canonical visuo-motor neurons (Rizzolatti and Fadiga, 1998), mirror neurons (Rizzolatti et al., 1996; Rizzolatti and Craighero, 2004; Thill et al., 2013), and a variety of ways in which embodiment influences perceptual, motor,

---

[1] We here use the term *agent* to refer to any system that displays a cognitive capacity, whether it is a human or and artificial cognitive system, such as a cognitive robot.

and cognitive performance (Varela et al., 1991; Barsalou et al., 2003). However, cognition is more than a collection of perceptuo-motor contingencies. In Varela's words, cognition is *effective action* (Maturana and Varela, 1987; Varela et al., 1991): action that preserves the agent's autonomy, maintaining the agent and its ontogeny, i.e., its continued development. Prospection, i.e., prediction or anticipation, is one of the two hallmarks of a cognitive agent, the second being the ability to learn new knowledge by making sense of its interactions with the world around it and, in the process, enlarging its repertoire of effective actions (Vernon, 2010; Vernon et al., 2010). Cognition entails being able to anticipate the need for action and being able to anticipate the outcome of that action. According to some theories, e.g., Hesslow's simulation hypothesis (Hesslow, 2002, 2012), this can be achieved with a form of internal simulation focusing on goal-directed prospective action selection and adaptation. Perhaps the best encapsulation of this prospective goal-directed perceptuo-motor approach is ideomotor theory (Stock and Stock, 2004).

Such a characterization, however, while necessary and useful, ascribes only behavioral attributes to cognition. We wish to reassert here that cognition also has a constitutive aspect, one that complements the behavioral aspect. The constitutive/behavioral distinction derives from the constitutive autonomy and behavioral autonomy of biological agents (Froese et al., 2007; Froese and Ziemke, 2009), especially those systems that exhibit the characteristic of recursive self-maintenance (Bickhard, 2000), capacities that reflect Varela's and Maturana's concepts of autopoiesis (Maturana, 1970, 1975; Maturana and Varela, 1980), organizational closure (Varela, 1979; Maturana and Varela, 1987), and operational closure (Froese and Ziemke, 2009; Stewart et al., 2010); see **Figure 1**. In this view, cognition, perception, and action serve to support the autonomy of the agent, both in a constitutive sense and in a behavioral sense. The constitutive aspect of autonomy focusses on the dynamic self-organization of the agent as an embodied system, maintaining itself as a viable structure that can support the organizational processes in the first place. The behavioral aspect of autonomy, on the other hand, targets the interaction of that embodied system with the environment in which it is embedded, maintaining the external conditions which are necessary for constitutive autonomy. We develop further the issue of constitutive autonomy in Section 2 on autonomy and deal with it more definitively in Section 3 on constitutive processes.

The goal of this article is to argue the case that the reciprocal coupling of action and perception is founded primarily on their roles in the constitutive autonomy of the agent and an associated circular causality of global and local processes of self-regulation, rather than being a mutual sensory-motor contingency that derives from exogenous behavior. Our objective is the synthesis of three strands of thinking into a cohesive picture: (a) the distinction between constitutive and behavioral autonomy and related processes, (b) the dynamics of circular causality, and (c) allostatic self-regulation (in contradistinction to homeostatic self-regulation). While all of these strands can indeed be traced to previous work, this paper is, to our knowledge, the first such synthesis.

FIGURE 1 | A characterization of autonomous agents situated in a two-dimensional space spanned in one dimension by behavioral autonomy and in the other by constitutive autonomy (based on **Froese et al., 2007**). The behavioral dimension focusses on the degree of independence of human assistance and the extent to which the system sets its own goals, and therefore corresponds loosely to what is referred to as the *degree of autonomy* in robotics. The constitutive autonomy dimension focusses on the organizational characteristics that allow the system to maintain itself as an identifiable autonomous entity. Since some systems don't exhibit the requisite organizational characteristics (e.g., organizational closure; see main text), they aren't constitutively autonomous. These occupy the white region at the bottom of the space. Those systems that *are* constitutively autonomous can make different levels of contribution to the maintenance of their autonomy and, thus, this dimension corresponds loosely to *strength of autonomy* and the *task entropy* in robotics (Sheridan and Verplank, 1978).

The article begins with a discussion of biological autonomy, clarifying the distinction between constitutive and behavioral autonomy. This sets the scene for the introduction of constitutive processes. We begin with an explanation of the difference between self-organization and emergence and then summarize the key processes of autopoiesis, organizational closure, and structural coupling. These processes exhibit the pivotal attributes of continuous reciprocal causation and circular causality. We consider how homeostasis and the related concept of allostasis contribute to this circular self-regulation. Finally, we explain how the reciprocal coupling of perception and action can be understood in this framework, arguing that the coupling happens first and foremost via the internal milieu realized by the agent's embodiment. In presenting this synthesis, the article integrates and builds on many quite disparate concepts, not all of which will be familiar to every reader; an introduction to many of these ideas can be found in tutorial texts (e.g., Vernon, 2014).

2. AUTONOMY

Autonomy is a difficult concept to tie down (Boden, 2008) and there are several perspectives on what it means (Froese et al., 2007). For convenience, we will adopt a definition of autonomy as the degree of self-determination of a system, i.e., the degree to which a system's behavior is not determined by the environment

and, thus, the degree to which a system determines its own goals (Ziemke, 1997, 1998; Bertschinger et al., 2008; Seth, 2010; Vernon, 2014). Implicit in this definition is the idea that, in addition to selecting its goals, the agent can then choose how best to achieve them and that it can then act to do so. For biological autonomous entities, the issue of autonomy becomes one of survival, typically in the face of precarious conditions, i.e., environmental conditions in which the entity has to work to keep itself alive as an autonomous system, both physically and organizationally as a dynamic self-sustaining entity.

Living systems face two problems: they are delicate and they are dissipative. Being delicate means that they are easily disrupted and possibly destroyed by the stronger physical forces present in their environment (including other biological agents). Consequently, living systems have to avoid these disruptions and repair or heal them when they occur. Dissipation arises from the fact that living systems are comprised of far-from-equilibrium processes (Bickhard, 2000). This means that the system must have some external source of energy or matter if they are to avoid lapsing into a state of thermodynamic equilibrium. If they do succumb to this, they come to rest and cease to be able to change in response to or in anticipation of any external factors that would threaten their autonomy or their existence[2]. Again, as with the delicacy of living systems, the dissipation inherent in far-from-equilibrium stability means that the system has to continually acquire resources and repair damage to itself. All of this has to be done by the agent itself. Of course, it is better if the agent can avoid damage in the first place and cognition as a prospective modulator of perception and action is one of the primary mechanisms at the agent's disposal (Barandiaran and Moreno, 2008). By bringing to bear a capacity for prospection, cognition compensates for the fact that perception is bound to the here-and-now and allows the agent to anticipate the need for action and the outcome of that action.

From this perspective, autonomy, aided by cognition, *is* the self-maintaining organizational characteristic of living creatures that enables them to use their own capacities to manage their interactions with the world in order to remain viable: i.e., compensate for dissipation, avoid disruption, and self-repair when necessary (Christensen and Hooker, 2000). In other words, autonomy is the process by which a system manages—self-regulates—to maintain itself as a viable entity in the face of the precarious circumstances with which the environment continually confronts it. In Bickhard's words "the grounds of cognition are adaptive far-from-equilibrium autonomy—recursively self-maintenant autonomy" (Bickhard, 2000).

One can distinguish two types of autonomy: *behavioral autonomy* and *constitutive autonomy* (Froese et al., 2007; Barandiaran and Moreno, 2008; again, refer to **Figure 1**). Behavioral autonomy focusses on the external characteristics of the system: the extent to which the agent sets its own goals and its robustness and flexibility in dealing with an uncertain and possibly precarious environment. On the other hand,

constitutive autonomy focusses on the internal organization and the organizational processes that keep the system viable and maintain itself as an identifiable autonomous entity. Agents that are constitutively autonomous can make different levels of contribution to the maintenance of their autonomy, making them less or more effective in dealing with the uncertainty and precariousness of the environment in which the system is embedded and in which it has to survive. Behavioral and constitutive autonomy are linked: an agent can't deal with uncertainty and danger if it is not organizationally—constitutively—equipped to do so. Its behavior depends on internal preparedness, achieved both through the neural mechanisms of the central and peripheral nervous system and the hormonal mechanisms of the endocrine system. On the other hand, in precarious circumstances, the agent needs behavioral autonomy to allow it to achieve the requisite environmental conditions—through interaction—for constitutive autonomy to be able to operate at all. This complementarity of the constitutive and the behavioral reflects two different sides of the characteristic of recursive self-maintenant systems to deploy different processes of self-maintenance depending on environmental conditions: one—constitutive autonomy—is the internal endogenous aspect of that adaptive capacity and the other—behavioral autonomy—is the external exogenous aspect of that adaptive capacity.

The constitutive-behavioral distinction is sometimes cast as a difference between *constitutive* processes and *interactive* processes (Froese and Ziemke, 2009). As we have said, constitutive processes deal with the agent itself, its organization, and its maintenance as an agent through on-going processes of self-construction and self-repair. On the other hand, interactive processes deal with the interaction of the agent with its environment. Both processes play complementary roles in autonomous operation of the agents. Constitutive processes are more fundamental to the autonomy of the agent but both are required.

3. CONSTITUTIVE PROCESSES

3.1. Self-organization vs. Emergence

Autonomy is closely linked to self-organization. Boden notes that autonomy, self-organization, and freedom are three notoriously slippery notions and none of them can be properly understood without considering the others (Boden, 2008). One definition of self-organization goes as follows.

> "A process in which pattern at the global level of a system emerges solely from numerous interactions among the lower-level components of the system. The rules specifying interactions among the system's components are executed using only local information, without reference to the global pattern"
>
> (Camazine, 2006).

Emergence also refers to a process involving interacting components in a system and the consequent generation of a global pattern. However, in this case, the global pattern emerges as something qualitatively different from the underlying assembly of components and, most significantly, is not simply

---

[2]Self-healing structure-maintaining properties are also exhibited by some non-living non-equilibrium dissipative entropy-producing complex systems (Kondepudi et al., 2015).

a consequence of the superposition of the contributions of the individual components[3].

The form of self-organization in emergence gives rises to systems that have a clear identity or behavior that results from two factors: (a) local-to-global determination and (b) global-to-local determination. In local-to-global determination, the emergent process has its global identity constituted and constrained by local interactions. In global-to-local determination, the global identity and its interaction with the system environment constrain the local interactions (Thompson and Varela, 2001; Froese and Ziemke, 2009; Di Paolo et al., 2010). This is sometimes referred to as *emergent self-organization*. Such self-organization has also been defined as "the spontaneous emergence (and maintenance) of order, out of an origin that is ordered to a lesser degree" (Boden, 2008). This definition provides the key link between self-organization, emergence, and autonomy: that self-organization results from the intrinsic spontaneous character of the system (possibly involving interaction with the environment) rather than being imposed by some external force or agent. In other words, emergent self-organization is autonomous and, *vice versa*, autonomous systems typically involve some form of emergent self-organization.

Emergent self-organization gives rise to a special view of autonomy, a view that is also characterized by *self-production*. Not only is there a reciprocal local-global and global-local determination but the nature of the determination is to re-create the local components from which the global system arises. This is constitutive autonomy (Froese and Ziemke, 2009). A system which exhibits constitutive autonomy actively generates and sustains its existence and systemic identity under precarious conditions, i.e., conditions that are antagonistic to the delicate and dissipative nature of the cognitive agent and which, in the absence of some appropriate form of emergent self-organization and associated behavior, would cause the system to cease to exist and cause its identity to be destroyed.

3.2. Self-production and Self-construction: Autopoiesis and Organizational Closure

Constitutive autonomy is related to the concept of *organizational closure*. Varela famously equates autonomy with organizational closure:

"Autonomous systems are mechanistic (dynamic) systems defined as a unity by their organization. *We shall say that autonomous systems are organizationally closed. That is, their organization is characterized by processes such that (1) the processes are related as a network, so that they recursively depend on each*

other in the generation and realization of the processes themselves, and (2) they constitute the system as a unity recognizable in the space (domain) in which the processes exist"

(Varela, 1979, p. 55; emphasis in the original).

Maturana and Varela subsequently define autonomy as "the condition of subordinating all changes to the maintenance of the organization" (Maturana and Varela, 1980).

Organizational closure is a necessary characteristic of a particular form of self-producing self-organization called *autopoiesis* (Maturana, 1970, 1975) that operates at the bio-chemical level, e.g., in cellular systems. Autopoietic systems are quite literally self-organizing systems that self-produce. Maturana and Varela later expanded the concept to deal with autonomous systems in general and refer to it in this context as *operational closure*, rather than autopoiesis which is specific to the bio-chemical domain. The operational closure vs. organizational closure terminology can be confusing because in some earlier publications (e.g., Varela, 1979), Varela refers to *organizational closure* but in later works (by Maturana and Varela themselves, e.g., Maturana and Varela, 1987, and by others, e.g., Stewart et al., 2010) this term was subsequently replaced in favor of *operational closure*. However, the term operational closure is appropriate when one wants to identify any system that is identified by an observer to be self-contained and parametrically coupled with its environment but not controlled by the environment. On the other hand, organizational closure characterizes an operationally-closed system that exhibits some form of self-production or self-construction (Froese and Ziemke, 2009).

These organizational principles are also reflected in the concepts of Bickhard's *self-maintenance* and *recursive self-maintenance* in far-from-equilibrium systems (Bickhard, 2000). Arguably, these two concepts represent a generalization of the ideas of self-construction and self-production introduced by Maturana and Varela in their processes of autopoiesis, organizational closure, and operational closure. Self-maintenant systems contribute to the conditions which are necessary to maintain it, i.e., to keep it going. In contrast, recursive self-maintenant systems exhibit a stronger form of autonomy in that they can deploy different processes of self-maintenance depending on environmental conditions, recruiting different self-maintenant processes as conditions in the environment require. Self-maintenance and recursive self-maintenance align well with the concepts of self-organization and emergent self-organization (both constitutive and behavioral autonomy), respectively.

3.3. Continuous Reciprocal Causation and Circular Causality

In the foregoing, there has been a recurring theme: a circular relationship between part and whole: between local factors and global factors. It appears that the characteristics of emergence and emergent self-organization are dependent on dynamic re-entrant structures. This is related to the concept of *continuous reciprocal causation* (CRC; Clark, 1997) which occurs when some system is both continuously affecting and simultaneously being affected

---

[3]Seth distinguishes between nominal, weak, and strong emergence (Seth, 2010). Nominal emergence is simply the idea that emergence is some property that can be exhibited by a complete macro-level object or system but not by its constituent parts. Strong emergence claims that macro-level properties are in principle not deducible from observation of the micro-level components and they have causal powers that are irreducible, i.e., they arise only because of the existence of the emergent behaviors. These causal powers are directed at the behavior of the components from which the emergent pattern emerges. Weak emergence sits somewhere in between nominal and strong emergence. It doesn't commit to the principled irreducibility of macroscopic patterns or behavior to microscopic activity but asserts that the relationship between the two levels is complex.

by activity in some other system (Clark, 1998)[4]. In effect, one system causes an effect in a second system which then causes an effect in the first, reinforcing the dynamic and causing the process to continue. CRC can also occur in a single system. In this case, the causal contribution of each systemic component partially determines, and is partially determined by, the causal contributions of large numbers of other systemic components. Wheeler puts it like this: "CRC is causation that involves multiple simultaneous interactions and complex dynamic feedback loops, such that (a) the causal contribution of each systemic component partially determines, and is partially determined by, the causal contributions of large numbers of other systemic components, and, moreover, (b) those contributions may change radically over time" (Wheeler, 2008).

This single-system CRC is often referred to as *circular causality* or *circular causation* (Varela, 1979). While circular causality can occur between distinct sub-systems in this overall system, it more usually reflects the interaction between global system dynamics (the whole) and local system dynamics (the parts). For example, Kelso uses the term circular causality to describe the situation in dynamical systems where the cooperation of the individual parts of the system determines the global system behavior which, in turn, governs the behavior of these individual parts (Kelso, 1995). Thus, circular causality exists *between* levels of a hierarchy of system and sub-system. This influence of macroscopic levels on microscopic levels in a system is captured in the term *downward causation* i.e., that global-to-local or macroscopic-to-microscopic aspect of circular causality whereby the global system behavior causally influences the individual system components (Thompson and Varela, 2001; Seth, 2010). In circularly causal systems, global system behavior influences the local behavior of the system components and yet it is the local interaction between the components that determines the global behavior. Thus, in biological autonomy, the degree of participation of the components of a system is determined by the global behavior which, in turn, is determined by the interactions among the components through causal reciprocal feedback loops. Again, these ideas are also echoed in Bickhard's concept of recursive self-maintenance (Bickhard, 2000).

The idea of circular causality is also related to the notion of entrainment where the global macro state entrains the micro-constituent processes of which it comprises in order to maintain that macro state. This has been applied to higher level forms of constitutive processing, e.g., in human emotions, where the macro states provide a substrate for learning (Lewis, 2005). Similarly, circular causality is a central feature of an information-theoretic model of self-sustainability, i.e., autonomy, in ecosystem networks (Ulanowicz, 1998, 2000, 2011). The same model has been used to characterize the emergence and development of beliefs in human cognition (Castillo et al., in press). This network-centric perspective aligns with the interaction-dominant view of cognitive dynamics which

highlights that cognition is characterized by interactions among multiple spatial and temporal scales of organization and among nested structures, rather than by relationships between simpler components at a single scale (i.e., the component-dominant view; Ihlen and Vereijken, 2010; Dixon et al., 2012).

4. REALIZING CIRCULAR CAUSALITY

How might circular causality be manifest in a cognitive agent and, more generally, in a system that exhibits constitutive autonomy? In this section, we consider how homeostasis and the related concept of allostasis contribute to circular global-to-local and local-to-global self-regulation.

4.1. Homeostasis

The process of self-regulation is central to constitutive autonomy. In biological systems, the automatic regulation of physiological functions is referred to as *homeostasis*, a term coined by Cannon (1929) formalizing the idea advanced in the nineteenth century by Claude Bernard that "all the vital mechanisms, however varied they may be, have only one object, that of preserving constant the conditions of life in the internal environment" (Bernard, 1878). Put simply, homeostatic processes regulate the operation of a system in order to keep the value of some system variables constant or within acceptable bounds, e.g., body temperature and blood glucose level. It does this by sensing any deviation from the desired value and feeding this error back to the control mechanism to correct the error. The desired value is called the *setpoint* in control theory and the use of the deviation from the desired value is called feedback.

We have in previous work (Morse et al., 2008; Ziemke and Lowe, 2009) suggested that the autonomy of an agent is effected through a hierarchy of homeostatic self-regulatory processes, exploiting a spectrum of associated affective (i.e., emotional or feeling) states, ranging from basic reflexes linked to metabolic regulation, through drives and motives, and on to the emotions and feelings often linked to higher cognitive functions. The progression of processes of homeostasis from basic reflexes and metabolic regulation, through drives and motives, to emotions and feelings is described in a schema for a cognitive architecture that places affect on an equal footing with more conventional cognitive processes. This progression follows closely Damasio's hierarchy of levels of homeostatic regulation (Damasio, 2003) and is based on a relatively broad notion of homeostasis as "the process of maintaining the internal milieu physiological parameters (such as temperature, pH and nutrient levels) of a biological system within the range that facilitates survival and optimal function" (Damasio, 2003; Damasio and Carvalho, 2013). Different homeostatic processes regulate different system properties.

Typically, the autonomous agent is perturbed during interactions with the world with the result that the organizational dynamics have to be adjusted. This process of adjustment is exactly what is meant by homeostasis— self-regulation—and the motives at every level of this hierarchy of homeostatic processes are effectively the drives that are required to return the agent to a state where its autonomy is no longer threatened. In the

---

[4]Reciprocal causation (Laland et al., 2012) also exists between proximate and ultimate mechanisms of evolutionary biology (Mayr, 1961; Tinbergen, 1963; Scott-Phillips et al., 2011).

interaction with the world around it, the perturbations of the agent by the environment have no intrinsic value in their own right—they are just the stuff that happens to the agent as it goes about its business of survival—but for the agent this stuff, these interactions and perturbations, have a perceived value in that they act to endanger or support its autonomy. This value is conveyed through the affective aspect of these homeostatic processes and consequently the agent then attaches some value to what is an otherwise neutral world (even if it is a precarious one; Di Paolo, 2005). The implications for perception and action are significant and this brings us to the crucial issue regarding the reciprocal coupling of action and perception in cognition.

First, perceptions and actions form a complementary set of environment-agent/agent-environment perturbations that are related not as extrinsic stimulus-response perceptuo-motor contingencies but as intrinsic processes that lead to the regulation of the system and autonomy preservation through emergent self-organization. The processes of perception and action are mutually dependent because they are both modulated by the system—globally-determined—through downward causation and, together with other homeostatic processes, they give rise to the global constitutive autonomy-preserving system behavior.

Second, perception and action are reciprocally coupled and mutually dependent because, from the perspective of enactive cognitive science, perception and action form a joint process of making sense of the world in which the agent is embedded (Maturana and Varela, 1987; Varela et al., 1991; Vernon, 2010). This "sense" captures the lawfulness of the agent's environment as it relates to the agent's constitutive and behavioral autonomy. Since the agent is an organizationally closed system, perception and action are perturbing forces rather than system inputs and outputs. This process of mutual perturbation of the agent and environment in which it is embedded, facilitating the on-going operational identity of the agent and its autonomous self-maintenance, is known as *structural coupling*. The process of structural coupling produces an embodiment-specific congruence between the system and its environment. For this reason, we say that the system and the environment are *co-determined*. In enactive cognitive science, this is also referred to as *structural determination* to emphasize the dependence of an agent's space of viable environmentally-triggered changes on the agent's structure, i.e., its particular embodiment, and its internal dynamics (Maturana and Varela, 1987; Varela et al., 1991).

4.2. Allostasis

While many autonomous agents are self-governing in the sense that they adjust automatically to events in the environment and self-correct when necessary (e.g., by way of homeostasis), other autonomous agents begin to adjust *before* the event actually occurs. This form of autonomy requires a continual preparation for what might be coming next. It means that an autonomous system anticipates what events might occur in its environment and actively prepares for them so that it is capable of dealing with them if they do occur. From this perspective, autonomy requires

pre-emptive action, not just reactive action, and predictive self-regulation, not just reactive self-regulation. These autonomous systems ready themselves for multiple contingencies, i.e., possible events, and have several strategies for dealing with them. They deploy them while pursuing some goal or other that the system has defined for itself. To an extent, this characteristic can be viewed as predictive self-regulation and is known as *allostasis* (Sterling, 2004; Schulkin, 2011; Sterling, 2012).

Allostasis encourages a rethink of the classical control theoretic perspective on homeostasis revolving around feedback loops respecting set points that demarcate ideal states. According to Sterling (Sterling, 2004), allostasis can be conceived in terms of prediction where brain areas implicated in planning and decision making are viewed as supplying inputs that may override other inputs that signal errors from ideal homeostatic balance. Such global overriding of "basal" homeostasis operates in the service of supplying the organism with the resources previously learned to be necessary to meet predicted environmental pressures. Sterling considers allostasis as a means of permitting adaptive bodily regulation according to "stability through change" which accounts for both internal needs and external pressures (or opportunities) in contrast to the Bernard notion of "stability through constancy." Thus, allostasis is concerned with adapting to change in order to achieve the goal of stability in the face of uncertain circumstances. Efficient regulation requires the anticipation of needs and preparation to satisfy them before they arise: "The brain monitors a very large number of external and internal parameters to anticipate changing needs, evaluate priorities, and prepare the organism to satisfy them *before* they lead to errors. The brain even anticipates its own local needs, increasing flow to certain regions—before there is an error signal" (Sterling, 2012). For example, human behavior in adapting to pain involves such predictive regulation, rather than mere reaction to tissue damage; that means "the nervous system is organized to anticipate potential pain and to adjust behavior before the risk of tissue damage becomes critical" (Morrison et al., 2013).

Allostasis, rather than being based on a reciprocal sharing of resources among systems (classic homeostasis), entails a degree of centralized control over sub-systems (Sterling, 2012). This can also be viewed in terms of downward causation: while Lewis (2005) references the macroscopic state that entrains its "emotion"-based microscopic constituents in the service of learning, Sterling (2012), p. 14, refers to "[memory] retrieval involv[ing] elaborate connections within "limbic" structures ... that ... project in cascades to prefrontal cortex." The limbic (emotion) system's constituents subserve constitutive organization (they relay signals related to sustaining the viability of the organism) and may be entrained by prefrontal cortex (which also reciprocally connects to neocortex) to facilitate adaptive behavior over a goal-directed sequence.

The focus on predictive regulation in allostasis mirrors strongly the anticipatory nature of cognition. Seth (2013) emphasizes this by developing the role that interoceptive predictive coding (as distinct from the more usual view of prospection in exteroceptive predictive coding) plays in the experience of "body ownership and conscious selfhood," viewing

emotions—subjective feeling states—as emerging from cognitive evaluations of physiological changes. As such, he targets a larger quarry in cognitive science: consciousness and neuropsychiatric illness. He hypothesizes that predictive coding arises through "an extended autonomic neural substrate" (Seth, 2013), p. 565, taking the principle on which allostasis is based—the causal role played by prediction in biological regulation—to the next level. Specifically, he highlights the role of active inference, an extension of predictive coding, whereby the interoceptive prediction errors can be suppressed not by updating the generative model that gave rise to the predictions but by internal action, translating the predictions into reference points for autonomic regulatory processes, e.g., physiological homeostasis. He notes that attention can then be viewed as a way of balancing active inference and model update, (referred to as precision weighting). Seth reinforces the idea that "an organism should maintain well-adapted predictive models of its own physical body ... and of its internal physiological condition" (Seth, 2013), p. 567, and in Seth (2015) he develops this further, grounding predictive coding, active infererence, and the principle of free energy in cybernetics and allostatic mechanisms for the maintenance of internal organization. Barrett and Simmons (2015) emphasize the same point with their *Embodied Predictive Interoception Coding* (EPIC) model, pointing out the direct link between active inference in the cortex and interceptive predictions of the internal milieu of the body, i.e., its physiological state relating to, e.g., heart rate, glucose levels, carbon dioxide in the bloodstream, and temperature.

In summary, allostasis differs from homeostasis in its predictive character and in its ability to anticipate and adapt to change rather than resist it. Significantly, allostasis is effected at a higher level of organization, involving greater number of sub-systems acting together in a coordinated manner with global processes modulating local ones, reflecting the character of circular causality. In contrast, mechanisms for homeostasis operate at a simpler level of negative feedback control (Sterling, 2004; Muntean and Wright, 2007; Sterling, 2012). Although you can view allostasis as a complementary mechanism to homeostasis, Sterling notes that it was introduced as a potential replacement for homeostasis as the core model of physiological regulation (Sterling, 2004, 2012).

In the next section, we look briefly at one example of how the principles of homeostasis and allostasis can be used to describe how cognition arises through constitutive autonomy.

4.3. Toward a Constitutive Autonomy Cognitive Architecture

Based on Damasio's view of the architecture and physiology of the mammalian brain, we have in previous work proposed two schemas for an enactive cognitive architecture that explicitly embraces the constitutive-behavioral distinction (Morse et al., 2008; Ziemke and Lowe, 2009). They are schemas in the sense that they identify the principal characteristics of the architecture without providing a detailed design of the component parts of the architecture and the dynamics of their interaction. A design approach called *holistic-reductionism* complements the schemas, focussing on the interdependencies of the components rather than on the identification of independent functional modules, as is normally the case with computational cognitive architecture design. Any modularity in the system emerges from the interdependence of the embodied cognitive processes rather than by phylogenetic pre-specificiation.

The first version of the architecture schema traverses two dimensions: (i) Constitutive Organization and (ii) Behavioral

FIGURE 2 | The three strands of thought being advanced in this paper are (A) the distinction between constitutive and behavioral autonomy and related processes, (B) the dynamics of circular causality, and (C) predictive allostatic self-regulation. The Cognitive-Affective Architecture Schematic in **(A)** is an example of the first aspect. It exhibits a spectrum of constitutive organization brought about by the recruitment of a progression of emotions, from reflexes, through drives and motivations, to emotions-proper and feelings. Each level in the constitutive organization is associated on the Internal Organization axis with an increasing level of homeostatic autonomy-preserving self-maintenance, ranging from basic metabolic processes through reactive sensorimotor activity (pre-somatic effects), associative learning and prediction (somatic modulation), to interoception and internal simulation of behavior prior to action. Equally, each level in the constitutive organization is associated on the Behavioral Organization axis with an increasing level of complexity in behavior, ranging from approach-avoidance, sequenced behaviors, and multi-sequenced behaviors. A more complete cognitive architecture that fully embraces constitutive autonomy would also incorporate processes for circular causality and allostasis.

Organization (Morse et al., 2008). The former refers to the system's internal dynamics as it maintains its integrity—its autonomy—in the face of perturbation by various stimuli. At the core of this space there is metabolic homeostatic self-regulation. This extends to stimulus valence evaluation, somatic state response, and content evaluation, each level offering increasing organizational complexity, an increasing degree of decoupling between stimulus and response, and an increasing degree of appraisal and associated adaptivity. Each level in the constitutive organization dimension is matched by an associated level in the behavioral organization dimension: approach-avoidance, sequenced behavior, and multi-sequenced behavior, respectively. Thus, the behavioral organization dimension is coupled by sensorimotor perception to the constitutive organizational dimension.

A later version of the architecture (see **Figure 2**) reflects this coupling by referring to a single space of constitutive organization which is viewed from two perspectives: internal organization and behavioral organization (Ziemke and Lowe, 2009). The spectrum of constitutive organization is realized by the recruitment of a progression of emotions, from reflexes, through drives and motivations, to emotions-proper and feelings. Each level in constitutive organization is associated on the internal organization axis with an increasing level of homeostatic autonomy-preserving self-maintenance, ranging from basic metabolic processes through reactive sensorimotor activity (pre-somatic effects), associative learning and prediction (somatic modulation), to interoception and internal simulation of behavior prior to action. Equally, each level in constitutive organization is associated on the behavioral organization axis with an increasing level of complexity in behavior, ranging from approach-avoidance, sequenced behaviors, and multi-sequenced behaviors.

The key idea is that different levels of cognitive function and behavioral complexity are associated with, and are brought about by, different levels of emotion, each linked to affective homeostatic processes ranging from reflexes right through to internal simulation. An extension of this schematic that augmented the homeostatic processes with allostatic ones would embrace more fully the concepts of constitutive autonomy advanced in this paper, including, as we have mentioned above, circular causality.

5. CONCLUSION

Cognition is commonly cast as a prospective process of adaptation, growth, and development (Vernon, 2010; Vernon et al., 2010), often focussing on behavioral autonomy. Here, however, we have recast cognition in a different light that emphasizes the importance of constitutive autonomy. Specifically, we have argued for a form of predictive regulation—allostasis—that is intrinsic to adaptive agents and exhibits a type of circular causality that naturally gives rise to reciprocal coupling of perception and action in embodied agents. The resulting synthesis of these different lines of thinking lead us to argue that perceptions and actions form a complementary set of environment-agent/agent-environment perturbations that are related not only as extrinsic stimulus-response perceptuo-motor contingencies but also as intrinsic processes that lead to the regulation of the system and autonomy preservation through emergent self-organization. The perceptions and actions are mutually dependent because they are both modulated by the system—globally-determined—through downward causation. Together they form a process of mutual perturbation of the agent and environment in which it is embedded, i.e., structural coupling, that facilitates both constitutive and behavioral autonomy. It remains as a significant research challenge to uncover the specific mechanisms by which circular causality and allostasis arise in natural agents and how—and to what degree—they might be replicated in artificial systems.

ACKNOWLEDGMENTS

REFERENCES

Barandiaran, X., and Moreno, A. (2008). Adaptivity: from metabolism to behavior. *Adapt. Behav.* 16, 325–344. doi: 10.1177/1059712308093868

Barrett, L. F., and Simmons, W. K. (2015). Interoceptive predictions in the brain. *Nat. Rev. Neurosci.* 16, 419–429. doi: 10.1038/nrn3950

Barsalou, L. W., Niedenthal, P. M., Barbey, A., and Ruppert, J. (2003). "Social embodiment," in *The Psychology of Learning and Motivation,* Vol. 43, ed B. Ross (San Diego, CA: Academic Press), 43–92.

Bernard, C. (1878). *Leçons Sur Les Phénomènes de la vie Communs Aux Animaux Aux et Végétaux.* Paris: J.-B. Baillière.

Bertschinger, N., Olbrich, E., Ay, N., and Jost, J. (2008). Autonomy: an information theoretic perspective. *Biosystems* 91, 331–345. doi: 10.1016/j.biosystems.2007.05.018

Bickhard, M. H. (2000). Autonomy, function, and representation. *Commun. Control - Artif. Intell.* 17, 111–131.

Boden, M. A. (2008). Autonomy: what is it? *Biosystems* 91, 305–308. doi: 10.1016/j.biosystems.2007.07.003

Camazine, S. (2006). "Self-organizing systems," in *Encyclopedia of Cognitive Science* (John Wiley & Sons, Ltd.). doi: 10.1002/0470018860.s00644

Cannon, W. B. (1929). Organization of physiological homeostasis. *Physiol. Rev.* 9, 399–431.

Castillo, R. D., Kloos, H., Richardson, M. J., and Waltzer, T. (in press). Beliefs as self-sustaining networks: drawing parallels between networks of ecosystems and adults' predictions. *Front. Psychol.*

Christensen, W. D., and Hooker, C. A. (2000). An interactivist-constructivist approach to intelligence: self-directed anticipative learning. *Philos. Psychol.* 13, 5–45.

Clark, A. (1997). *Being There: Putting Brain, Body, and World Together Again.* Cambridge, MA: MIT Press.

Clark, A. (1998). Time and mind. *J. Philos.* 95, 354–376. doi: 10.2307/2564539

Damasio, A., and Carvalho, G. B. (2013). The nature of feelings: evolutionary and neurobiological origins. *Nat. Rev. Neurosci.* 14, 143–152. doi: 10.1038/nrn3403

Damasio, A. R. (2003). *Looking for Spinoza: Joy, Sorrow and the Feeling Brain.* Orlando, FL: Harcourt.

Di Paolo, E. A. (2005). Autopoiesis, adaptivity, teleology, agency. *Phenomenol. Cogn. Sci.* 4, 429–452. doi: 10.1007/s11097-005-9002-y

Di Paolo, E., Rohde, M., and De Jaegher, H. (2010). "Horizons for the enactive mind: values, social interaction, and play," in *Enaction: Towards a New*

Paradigm for Cognitive Science, eds J. Stewart, O. Gapenne, and E. Di Paolo (Cambridge, MA: MIT Press), 33–87.

Dixon, J. A., Holden, J. G., Mirman, D., and Stephen, D. G. (2012). Multifractal dynamics in the emergence of cognitive structure. *Top. Cogn. Sci.* 4, 51–62. doi: 10.1111/j.1756-8765.2011.01162.x

Froese, T., Virgo, N., and Izquierdo, E. (2007). "Autonomy: a review and a reappraisal," in *Proceedings of the 9th European Conference on Artificial Life: Advances in Artificial Life*, Vol. 4648, eds F. Almeida e Costa, L. M. Rocha, E. Costa, I. Harvey, and A. Coutinho (Berlin; Heidelberg: Springer), 455–464. doi: 10.1007/978-3-540-74913-4_46

Froese, T., and Ziemke, T. (2009). Enactive artificial intelligence: investigating the systemic organization of life and mind. *Artif. Intell.* 173, 466–500. doi: 10.1016/j.artint.2008.12.001

Hesslow, G. (2002). Conscious thought as simulation of behavior and perception. *Trends Cogn. Sci.* 6, 242–247. doi: 10.1016/S1364-6613(02)01913-7

Hesslow, G. (2012). The current status of the simulation theory of cognition. *Brain Res.* 1428, 71–79. doi: 10.1016/j.brainres.2011.06.026

Ihlen, E. A. F., and Vereijken, B. (2010). Interaction-dominant dynamics in human cognition: Beyond $1/f^\alpha$ fluctuation. *J. Exp. Psychol. Gen.* 139, 436–463. doi: 10.1037/a0019098

Kelso, J. A. S. (1995). *Dynamic Patterns – The Self-Organization of Brain and Behaviour, 3rd Edn*. Cambridge, MA: MIT Press.

Kondepudi, D., Kay, B., and Dixon, J. (2015). End-directed evolution and the emergence of energy-seeking behaviour in a complex system. *Phys. Rev. E* 91, 1–5. doi: 10.1103/PhysRevE.91.050902

Laland, K. N., Odling-Smee, J., Hoppitt, W., and Uller, T. (2012). More on how and why: cause and effect in biology revisited. *Biol. Philos.* 28, 719–745. doi: 10.1007/s10539-012-9335-1

Lewis, M. D. (2005). Bridging emotion theory and neurobiology through dynamic systems modelling. *Behav. Brain Sci.* 28, 169–245. doi: 10.1017/S0140525X0500004X

Maturana, H. (1970). *Biology of Cognition*. Research Report BCL 9.0, University of Illinois, Urbana, IL.

Maturana, H. (1975). The organization of the living: a theory of the living organization. *Int. J. Man Mach. Stud.* 7, 313–332. doi: 10.1016/S0020-7373(75)80015-0

Maturana, H., and Varela, F. (1987). *The Tree of Knowledge — The Biological Roots of Human Understanding*. Boston, MA; London: New Science Library.

Maturana, H. R., and Varela, F. J. (1980). *Autopoiesis and Cognition — The Realization of the Living*. Dordrecht: Boston Studies on the Philosophy of Science; D. Reidel Publishing Company.

Mayr, E. (1961). Cause and effect in biology. *Science* 134, 1501–1506. doi: 10.1126/science.134.3489.1501

Morrison, I., Perini, I., and Dunham, J. (2013). Facets and mechanisms of adaptive pain behavior: predictive regulation and action. *Front. Hum. Neurosci.* 7:755. doi: 10.3389/fnhum.2013.00755

Morse, A., Lowe, R., and Ziemke, T. (2008). "Towards an enactive cognitive architecture," in *Proceedings of the First International Conference on Cognitive Systems* (Karlsruhe).

Muntean, I., and Wright, C. D. (2007). Autonomous agency, AI, and allostasis — a biomimetic perspective. *Pragmat. Cogn.* 15, 485–513. doi: 10.1075/pc.15.3.07mun

Rizzolatti, G., and Craighero, L. (2004). The mirror neuron system. *Annu. Rev. Physiol.* 27, 169–192. doi: 10.1146/annurev.neuro.27.070203.144230

Rizzolatti, G., and Fadiga, L. (1998). "Grasping objects and grasping action meanings: the dual role of monkey rostroventral premotor cortex (area F5)," in *Sensory Guidance of Movement, Novartis Foundation Symposium 218*, eds G. R. Bock and J. A. Goode (Chichester: John Wiley and Sons), 81–103.

Rizzolatti, G., Fadiga, L., Gallese, V., and Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cogn. Brain Res.* 3, 131–141. doi: 10.1016/0926-6410(95)00038-0

Schulkin, J. (2011). Social allostasis: anticipatory regulation of the internal milieu. *Front. Evolut. Neurosci.* 2:111. doi: 10.3389/fnevo.2010.00111

Scott-Phillips, T. C., Dickins, T. E., and West, S. A. (2011). Evolutionary theory and the ultimate-proximate distinction in the human behavioural sciences. *Perspect. Psychol. Sci.* 6, 38–47. doi: 10.1177/1745691610393528

Seth, A. (2010). Measuring autonomy and emergence via Granger causality. *Artif. Life* 16, 179–196. doi: 10.1162/artl.2010.16.2.16204

Seth, A. K. (2013). Interoceptive inference, emotion, and the embodied self. *Trends Cogn. Sci.* 17, 565–573. doi: 10.1016/j.tics.2013.09.007

Seth, A. K. (2015). "The cybernetic Bayesian brain — from interoceptive inference to sensorimotor contingencies," in *Open MIND*, Vol. 35, eds T. Metzinger and J. M. Windt (Frankfurt: MIND Group), 1–24.

Sheridan, T. B., and Verplank, W. L. (1978). *Human and Computer Control for Undersea Teleoperators*. Technical report, MIT Man-Machine Systems Laboratory.

Sterling, P. (2004). "Principles of allostasis," in *Allostasis, Homeostasis, and the Costs of Adaptation*, ed J. Schulkin (Cambridge: Cambridge University Press), 17–64.

Sterling, P. (2012). Allostasis: a model of predictive regulation. *Physiol. Behav.* 106, 5–15. doi: 10.1016/j.physbeh.2011.06.004

Stewart, J., Gapenne, O., and Di Paolo, E. A. (2010). *Enaction: Toward a New Paradigm for Cognitive Science*. Cambridge MA: MIT Press.

Stock, A., and Stock, C. (2004). A short history of ideo-motor action. *Psychol. Res.* 68, 176–188. doi: 10.1007/s00426-003-0154-5

Thill, S., Caligiore, D., Borghi, A. M., Ziemke, T., and Baldassarre, G. (2013). Theories and computational models of affordance and mirror systems: an integrative review. *Neurosci. Biobehav. Rev.* 37, 491–521. doi: 10.1016/j.neubiorev.2013.01.012

Thompson, E., and Varela, F. (2001). Radical embodiment: neuronal dynamics and consciousness. *Trends Cogn. Sci.* 5, 418–425. doi: 10.1016/S1364-6613(00)01750-2

Tinbergen, N. (1963). On the aims and methods of ethology. *Z. Tierpsychol.* 20, 410–433. doi: 10.1111/j.1439-0310.1963.tb01161.x

Ulanowicz, R. E. (1998). A phenomenology of evolving networks. *Syst. Res. Behav. Sci.* 15, 373–383. doi: 10.1002/(SICI)1099-1743(1998090)15:5<373::AID-SRES271>3.0.CO;2-I

Ulanowicz, R. E. (2000). *Growth and Development; Ecosystems Phenomenology*. Lincoln, NE: toExcel Press.

Ulanowicz, R. E. (2011). "Quantitative methods for ecological network analysis and its application to coastal ecosystems," in *Treatise on Estuarine and Coastal Science*, Vol. 9, eds E. Wolanski and D. S. McLusky (Waltham, MA: Academic Press), 37–57.

Varela, F. (1979). *Principles of Biological Autonomy*. New York, NY: Elsevier North Holland.

Varela, F., Thompson, E., and Rosch, E. (1991). *The Embodied Mind*. Cambridge, MA: MIT Press.

Vernon, D. (2010). Enaction as a conceptual framework for development in cognitive robotics. *Paladyn J. Behav. Rob.* 1, 89–98. doi: 10.2478/s13230-010-0016-y

Vernon, D. (2014). *Artificial Cognitive Systems — A Primer*. Cambridge, MA: MIT Press.

Vernon, D., von Hofsten, C., and Fadiga, L. (2010). *A Roadmap for Cognitive Development in Humanoid Robots, Cognitive Systems Monographs (COSMOS)*, Vol. 11. Berlin: Springer.

Wheeler, M. (2008). Cognition in context: phenomenology, situated robotics and the frame problem. *Int. J. Philos. Stud.* 16, 323–349. doi: 10.1080/09672550802113235

Ziemke, T. (1997). "The 'environmental puppeteer' revisited: a connectionist perspective on 'autonomy'," in *Proceedings of the 6th European Workshop on Learning Robots* (Brighton, UK).

Ziemke, T. (1998). Adaptive behaviour in autonomous agents. *Presence* 7, 564–587. doi: 10.1162/105474698565947

Ziemke, T., and Lowe, R. (2009). On the role of emotion in embodied cognitive architectures: from organisms to robots. *Cogn. Comput.* 1, 104–117. doi: 10.1007/s12559-009-9012-0

Being Present in Action: A Theoretical Model About the "Interlocking" Between Intentions and Environmental Affordances

Stefano Triberti[1]* and Giuseppe Riva[1,2]

[1] Department of Psychology, Università Cattolica del Sacro Cuore, Milan, Italy, [2] Applied Technology for Neuro-Psychology Lab, Istituto Auxologico Italiano, Milan, Italy

Recent neuropsychological evidence suggest that a key role in linking perceptions and intentions is played by sense of presence. Despite this phenomenon having been studied primarily in the field of virtual reality (conceived as the illusion of being in the virtual space), recent research highlighted that it is a fundamental feature of everyday experience. Specifically, the function of presence as a cognitive process is to locate the Self in a physical space or situation, based on the perceived possibility to act in it; so, the variations in sense of presence allow one to continuously adapt his own action to the external environment. Indeed intentions, as the cognitive antecedents of action, are not static representations of the desired outcomes, but dynamic processes able to adjust their own representational content according to the opportunities/restrictions emerging in the environment. Focusing on the peculiar context of action mediated by interactive technologies, we here propose a theoretical model showing how each level of an intentional hierarchy (future-directed; present directed; and motor intentions) can "interlock" with environmental affordances in order to promote a continuous stream of action and activity.

Keywords: intentions, presence, action, agency, affordance

INTRODUCTION

Recently, Riva et al. (2011, 2015a), Riva and Mantovani (2012), Waterworth and Riva (2014) proposed that a fundamental role in coupling intentions, perception, and action is played by *sense of presence*, conceived as a specific cognitive process. The concept of sense of presence emerged around the Nineties in the field of interactive technology studies, in particular in Virtual Reality. Indeed, the first studies tried to understand what allowed people to feel present inside computer-simulated environments.

The so-called *Media Presence* theories (Loomis, 1992; Sheridan, 1992, 1994; Schloerb, 1995; Lombard and Ditton, 1997) consider sense of presence as the function of the experience of a given medium. These theories explain sense of presence on the basis of perception and attention. For example, according to Lombard and Ditton (1997) sense of presence appears when an "illusion of non-mediation" establishes, that is, the individual using virtual reality stops to pay attention to the technology in use (for example, the head mounted display) and focuses on the content of the virtual environment. On the one hand, these theories are useful to provide virtual reality design guidelines. On the other hand, these theories fail in explaining why something

like the sense of presence exists. Why, from an evolutionary point of view, should we need something like a cognitive process devoted to generate a sense of "being there" while interacting with simulation technologies? *Media Presence* theories do not provide answers to this question (Lee, 2004; Riva et al., 2011). The existence of sense of presence highlights that the perceived location of an individual is not a mere by-product of the individual actually being in a given place. Indeed, since the manipulation of environmental features (for example: digitally rendering a simulated environment via virtual reality) is able to alter the perceived location of the individual, it appears that a specific form of information processing is devoted to provide such outcome. We can consider critically the proposal by Lombard and Ditton (1997): if sense of presence depends on an illusion excluding the virtual reality technology from our attention, how can we feel a sense of presence in physical reality, where no technological mediations exist?

In contrast, *Inner Presence* theories (Zahorik and Jenison, 1998; Moore et al., 2002; Lee, 2004; Riva et al., 2011, 2015b; Riva and Waterworth, 2014) consider sense of presence as a fundamental component of our cognition, which plays a precise role in our everyday life and is not necessarily related to the fruition of interactive media. Riva et al. (2011, 2015b) proposed a complex model that describes sense of presence as a neuropsychological phenomenon whose central goal is the control of agency and activity, through the unconscious separation of "internal" and "external". In other words, the experience of presence can be described as the outcome of an intuitive meta-cognitive process that allows us to control our actions through the comparison between intentions and perceptions (Riva and Mantovani, 2012). Following this view, presence is a core neuropsychological phenomenon whose goal is to produce a sense of agency and control: I am present in a real or virtual space if I manage to put my intentions into action (enacting them). Feeling variations in the sense of presence, one can monitor his own actions and tune his activity accordingly.

According to this theory, the link between sense of presence and the enacting of intentions is strong and fundamental. If we consider again the field of virtual reality and new media, indeed, technical aspects of the virtual environments (such as, for example, pictorial realism or intensity of the sensory stimulation) have a weak impact on the sensation of "being there" if compared with the impression of being able to enact intentions. For example, a video game player can feel strongly present while playing a product with very simple graphics and basic animations. This could happen in that the game features:

(1) an usable and easy-to-master interface which is adequate to the player's capable movements.
(2) an easily understandable game structure (possible directions to take, objectives, game rules, etc.) which efficiently relates to the here-and-now desired outcomes for the player.
(3) a compelling, engaging and interesting storyline, which provides the player with long-term goals and ultimate objectives.

Indeed, numerous experiments demonstrated that self-reported sense of presence in virtual environments is strongly related to the usability and effectiveness of interactive features (Coelho et al., 2006) and to narration/storyline contents (Gorini et al., 2011; Triberti et al., 2014). According to *Inner Presence* theories, this aspects influence sense of presence also in everyday life. One can feel more or less present in a given situation depending on how much he has the impression of being able to enact his own intentions, recognizing and using environmental opportunities for action, and then monitoring the perceived action outcomes as more or less consistent with the representational content of intentions. But how does this happen? What does it mean to "feel present" in everyday life? How does exactly sense of presence relates to one's own Self?

According to Riva et al. (2011, 2015b), Waterworth and Riva (2014) sense of presence is a unitary feeling, but on the process side it can be divided into three phylogenetically different layers/subprocesses. On the side of Self, these layers are symmetrical to the layers of Self as described by Damasio (2011). According to him, the conscious Self is built, as a first step, on a collection of "primordial feelings" constituted by enteroceptive, proprioceptive and motor information coming from the body (proto-Self), which allow the organism to distinguish itself from the external environment. At the second level, core-Self is related to the perceptual differentiation between the Self and the recognized external object. Finally, the Autobiographical Self is related to the emergence of consciousness and symbolic/categorical knowledge: thanks to the use of language, we become able to represent the events in our personal story and, as a consequence, also to formulate abstract action plans oriented to the distant future. According to this layered conception of the Self, sense of presence can be represented as composed by three sub-processes. *Proto-presence* is the process of internal/external separation related to proprioception and motor control, since its object is the basic distinction between Self and non-Self, still without differentiating the characteristics of the external object(s). Basically, a sense of proto-presence allow us to monitor whether motor intentions are being correctly enacted by our own body, regardless of the external environment. Differently, *Core presence* is related to the sensorial experience of the environment. At this layer, the agent starts interacting with the objects. The "external" is specified at the level of affordances for actions, so, the agent monitor his own intentions as having or not the expected effects on the external environment which is around him at the present moment. Finally, *Extended presence* is to verify the significance to the Self of experienced events in the external world. The more the Self is present in significant experiences, the more it will be able to reach its goals, increasing the possibility of surviving. Extended presence requires intellectually and/or emotionally significant content. In other words, feeling extended presence means monitoring the enacting of abstract/general objectives into complex action plans. **Figure 1** shows how the three layers of presence relate to the Self as explained by Damasio (2011), and how intention enacting generates sense of presence through the

Layer of Presence	Consciousness	Layer of Self	Information
Proto-presence	Mostly unconscious	Proto-Self	Proprioceptive
Core presence	Conscious of here-and-now	Core Self	Perceptual
Extended presence	Conscious of Self in relation with the world	Autobiographical Self	Conceptual

Intention
(Op: prediction of the final state of the environment)

Action
(T2)

Cognition
(T1)

Enaction of the Representational Content

Presence:
Oa – Op = 0

Break:
Oa – Op ≠ 0

Intentional Agent

Perception
(T2)

Actual State of the External Environment (Oa)
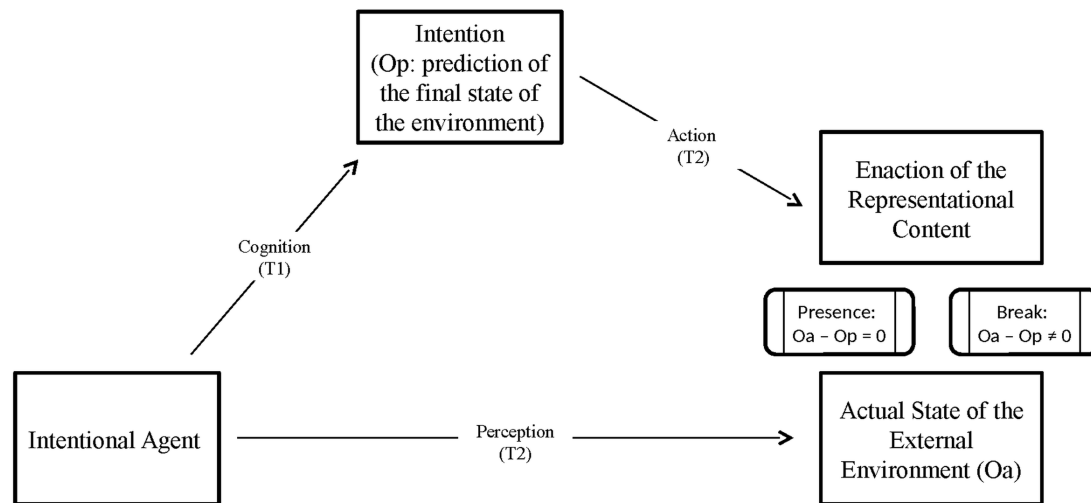
FIGURE 1 | Layers of presence and layers of Self, and a representation of the feeling of presence emerging from intention enacting and action monitoring (adapted from Waterworth and Riva, 2014).

confrontation between action and the final state of the external environment.

The theory of sense of presence highlights that we continuously monitor our own activity (in the form of intention enacting) in our own body, and in the external world. In this sense, intentions should be able to "interlock" with the opportunities for action coming from the environment, both at the level of the current situation and at the level of the extended, conceptual possibilities. On the one hand, this relates to activity monitoring when the action is initiated. For example, recent hypotheses coming from Neuroscience (Numan, 2015) sustain that the hippocampus and the prefrontal cortex play a critical role in the enacting of action plans and the formation of episodic, autonoetic memories; specifically, the prefrontal cortex is responsible for elaboration of goal-directed behavior and transmission of an efference copy of the action (or "corollary discharge") to the hippocampus, which serves as an intention-outcome comparator. Finally, the response of the hippocampal comparator returns to the prefrontal cortex where it is used to strengthen the current action plan (in case of success) or to reformulate it (in case of intention-outcome mismatch), this way fostering memory updating. On the other hand, in the present contribution we will try to show how a comparison between intentions and the external world happens even before the action. The "interlocking" metaphor highlights that,

as we will deepen in the next section, intentions should be conceived as layered structures themselves. Indeed, in order to an intention being enacted, any layer of an intentional structure (distal-conceptual; proximal-present; motor-micropresent) should find its own correspondence in the external world, in terms of feasible affordances. In this sense, "monitoring our own activity" does not mean to control for the consequences of actions only, but also, and more importantly, to perceive and recognize the affordances for action relying in the external environment before the action onset. How does this process actually happen? In order to explain this, we will deepen the concept of intention itself, showing how also intentions can be represented as hierarchies/layered structures; then, we will introduce a theoretical model about the interlocking between intentions and environmental affordances at the different levels of information processing.

INTENTIONS

Numerous philosophical conceptions, as well as common sense, posit that intentions guide actions. Classic experiments and theories challenged this apparently simple assumption, in that they have shown that neural activation related to the initiation of movement (readiness potential) seemed to appear independently

of conscious awareness (Libet, 1999, 2010). For this reason, according to Libet (1999, 2010) and other researchers (Wegner, 2002), intentions may not be the causal antecedents of action, but an illusion generated by the consciousness after the action onset.

The paradox highlighted by Libet's (1999, 2010) work can be resolved when considering a more sophisticated conception of intentions (Gallagher, 2006; Pacherie and Haggard, 2010). Indeed, intentions are not only motor representations guiding the motor components of an action and appearing just immediately before the movement itself, so that they can be fully associated with the readiness potential. On the contrary, they may develop at larger time scales (potentially, including almost a lifetime between the generation of the intention and its achievement) in that they entail conscious deliberative decisions, abstract, and descriptive representations, and "mental time travel" as a cognitive adaptation allowing humans to simulate contingencies and consequences of future actions (Suddendorf and Corballis, 2007; Eren, 2009; Corballis, 2013).

Recently, Pacherie (2006, 2008), Pacherie and Haggard (2010) introduced a dynamic theory of intentions which distinguishes among Distal Intentions, Proximal Intentions and Motor Intentions.

- Distal Intentions develop at potentially large time scales and represent abstract reasoning about means and plans (e.g.,: "I want to become a psychologist")
- Proximal Intentions constitute the conscious antecedents of a given action, developing at the level of the present time (e.g.,: "Now I will write a psychology essay to pass my university exam")
- Motor Intentions develop at a micro-present level and guide the motor components of the action (i.e., how the action should be physically performed); they are mostly unconscious (e.g.,: "I'm moving my hands on the keyboard *this way* to write").

As Pacherie (2008) says, the three layers of intentions do not simply coexist, but they form an "intentional cascade" with Distal Intentions generating Proximal Intentions, and Proximal Intentions generating Motor Intentions. However, it is not fully clear how can the different layers of intentions relate to the external world. Indeed, Motor Intentions are strictly dependent on the physical environment where the movement is about to take place; for example, the motor intention "moving my hands on the keyboard *this way* to write" should "take into consideration" distance between the body and the computer, the position of the keys on the keyboard, as well as the strength to use with the fingers in order to correctly activate the keys.

Similarly, the intention "now I will write an essay", as a proximal intention, has to "anchor the action plan in the current situation" (Pacherie, 2008, p. 188). Enacting a proximal intention means identifying the environmental affordances which permit the activation of the behavior. For example, the computer having a precise set of functions which allows one to write, cite and save his own work. In other words, at the level of proximal intentions, an agent should perceive and identify the opportunities for action existing within the environment. This happens not only at the mere motor level, but also identifying the functions of tools, the

limits imposed by obstacles, the possibility to move or not to a different environment, and so on.

From the point of view of intention enacting, the Distal Intention ("I want to become a psychologist") is the most elusive. Pacherie (2008, p. 188) sustains that distal intentions provide proximal intentions with an action plan that "may be still mostly descriptive and abstract". Indeed, they are not directly related to the context of action (one may intend to become a psychologist independently of what is happening around him at the present moment). But how do distal intentions relate to the external world?

Castelfranchi (2014, p. 107), who is interested in showing that intentions are a specific kind of goals, define intentions as "those goals that actually drive our voluntary actions or are ready/prepared to drive them". Also he focuses on trying to understand how abstract intentions can guide actions. He argues that abstract intentions need to be converted in "concrete cues", so that it becomes possible for the agent to control whether they have been achieved or not. For example, a distal and abstract intention such as "I want to take revenge for the offense" should be situated in a precise situation (or multiple situations) in which precise actions (insulting, manipulating, dueling) take place. Doing so, an agent can effectively control his own intentions and agency. However, the concept of "concrete cues" appears as somewhat elusive. What are "concrete cues" exactly? On the one hand, they are probably effects and micro-effects of the action, that the agent compare with the representational content of the intention to monitor whether the action is being performed as desired/expected, as it is argued by the Comparator Models of agency (Pacherie, 2008; Carruthers, 2012; Chambon et al., 2014; Numan, 2015). However, this kind of "concrete cues" (i.e., those coming from the detection of action consequences) seem not sufficient to us as explanation of intentions. In the next section, we will introduce a theoretical model about the "interlocking" between intentions and environmental affordances, in order to show how intentions can relate to the external world even in absence of "concrete cues" conceived as consequences of performed action.

INTRODUCING A MODEL ABOUT THE HIERARCHICAL INTERLOCKING OF ENACTED INTENTIONS

We argue that an agent should not "control" an intention just at the time the intention is already in the form of its physical enaction. In other words, and agent should know whether a given intention can be enacted or not, already at the time when the intention is distant-future directed, abstract and merely descriptive, still not specified into physical actions and micro-movements at the motor level. Specifically, the agent should know whether his own intentions satisfy or not criteria different from the ones the agent himself uses to monitor effectiveness of physical action. At each level of the intentional hierarchy, intentions are the object of a cognitive/intuitive evaluation which authorizes them to proceed down the cascade until the initiation and the monitoring of behavior. But what are the concrete

cues at the level of abstraction? How can we know whether a given cognitive process promoting behavior deserves the status of "intention" (e.g., goal guiding action or ready/prepared to guide it)? This kind of concrete cues allow the formation of an intention since they consist of information attesting intention's "enactability" in the external world. As Laurent (2003) observes, mental representations (and also intentions) do not represent the state of the external world, but the state of one's own engagement in the world. This means that intentions have to reflect the opportunities for enacting an action, this way becoming "simulated affordances" that illustrate what reality affords to enact behavior, at any level of information processing.

Indeed, each level of the intentional hierarchy is characterized by an external world-dependent requisite to be satisfied, in order to continue to guide action. So, we sustain that a given intentional hierarchy has to "interlock" with the external world, already at the time when the first movement(s) of the corresponding action are still not initiated.

- The motor intention is the intention that directly informs the movement. For this reason, it has to interlock with physical objects and properties which make possible, totally, or in part, the actual performance. Spatial and temporal limitations belong to this category, such as physical properties of objects that make them more or less adequate to the morphological structure of the agent. The condition for a cognitive process becoming a motor intention is having a representational content which is directly *tangible/manipulable*.
- The proximal intention consists in the product of the distal intention, as a general goal that is transformed into action plan(s). The proximal intention interlocks with environmental affordances, conceived as opportunities for action: it has to identify sets of functions, rules and provisions, in the form of possible courses of actions. For example, the intention "writing now an essay" requires the agent being able to recognize the environmental features which make possible the intention enacting, such as having the computer, turning it on and being able to write through the keyboard. To sum up, the condition for a cognitive process becoming a proximal intention is having a representational content which is *perceivable* as a concrete opportunity for action in the external world.
- The Distal Intention ("I want to become a psychologist"; "I want to take revenge of the offense"; "I want to win the gold medal") interlocks with socio-cultural conditions which maintain a relationship of resemblance with the "possible world" imagined by the agent, that is, those contents of his imagination featuring emotional and identity values. For example: "I want to become a psychologist" is not an abstract claim; it actually develops in a socio-cultural context where (1) some desirable outcomes for the agent (e.g.,: understanding people; studying behavior; treating psychopathologies) are incarnated in the "psychologist" figure, and (2) that socio-cultural context features a number of possible courses of action one could take to actually become a psychologist. The condition for a cognitive process becoming a distal intention is to have a representational content which is *thinkable*, basing on
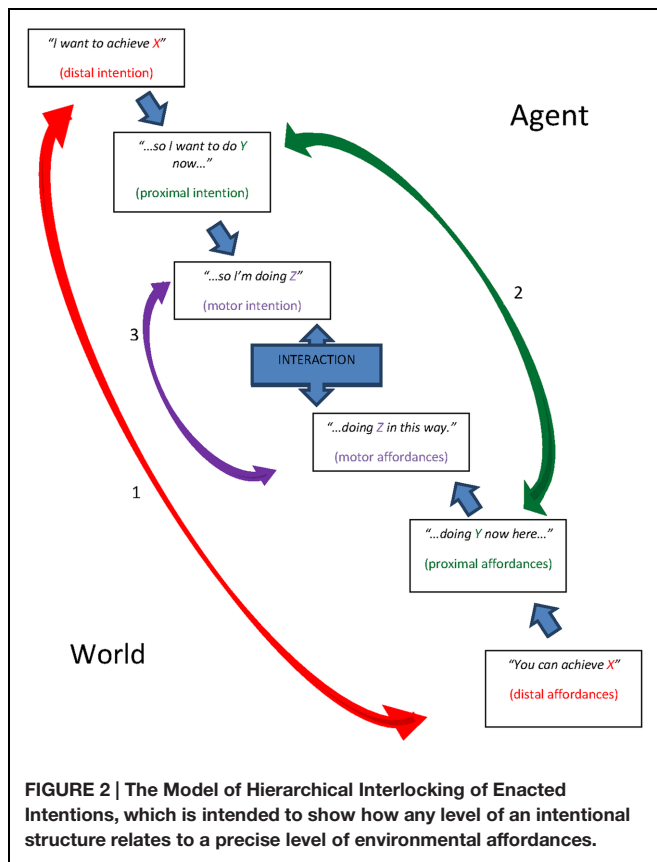
the characteristics of the socio-cultural context in the external world.

Intentions (from the motor level to the more abstract one) are already interlocked with the world even before they are transformed into actions, because they are fundamentally predictive. The human mind has the capacity to generate probabilistic models about the future, basing on the analysis of sensory inputs. According to the so-called "free energy" framework (Friston et al., 2006; Friston, 2009, 2010; Fotopoulou, 2014), the fundamental function of our brain is to reduce the inconsistency between predictions about the world and the world as it is actually perceived, or, monitoring the divergence between our motivations/needs and the "state of the coupling between the individual and his environment" (Laurent, 2003, p. 387). This inconsistency/divergence is the free energy, which has to be maintained at the lowest possible level to avoid surprise (Clark, 2013). The brain continuously generates prediction models based on noisy sensory inputs to represent future states of the body and the external world.

In our view, intentions work as prediction models with associated motivational value. Also a distal and abstract intention ("I want to become a psychologist") is constructed and sharpened over time to match incoming external/sensory inputs. This process entails two main sub-level processes related to one's own activity monitoring. The first entails the identification of affordances and opportunities for action (even in abstract terms) to understand whether and how the intention can be progressively accompanied to become an action. The second process consists in its progressive transformation in a more-and-more practical, concrete, and motor guide for action. In other words, the second process is the generation of the intentional cascade.

Let us consider an example. One person feels a motivational drive to study human behavior and to treat psychopathologies. For some reason, this has a positive emotional value for him and he considers these tasks as consistent with his own identity. So, he decides *he wants to be a psychologist*. This distal, abstract intention is not dependent on the here-and-now context. However, it does not start to guide behavior, nor it starts an intentional cascade, "out of nowhere". The agent has to control whether or not there are, in his perception of the world, general opportunities to reach his own purpose: going to the university, following courses, augmenting his own knowledge, obtaining a psychology degree which would be accepted and recognized by the society he lives in. The distal intention has to match with thinkable opportunities in the world, this way starting to reduce the inconsistency between the representational content of the intention ("I want to become a psychologist") and the current reality of the thinkable and perceivable world ("I am currently not a psychologist").

Of course, at this level the free energy resulting from the inconsistency between the volitional representation and the actual state of the world is very high; moreover, it is probably impossible to represent it, because both the intention and the desired state exist just in abstract, imaginative terms. For this reason, while general opportunities for actions start to

FIGURE 2 | The Model of Hierarchical Interlocking of Enacted Intentions, which is intended to show how any level of an intentional structure relates to a precise level of environmental affordances.

appear and to match with the intention, the intention itself has to be specified in here-and-now guides for action to adapt to more-and-more situated environmental affordances. This is "when" proximal intentions are generated, and have to be matched with the concrete opportunities for action existing in the current situation. Then, while proximal affordances are approached, motor affordances may appear informing how the physical movement should be performed. At this moment, the action can be initiated transforming proximal intentions in the best set of motor intentions for the situation. Doing this, the agent progressively fulfill his own distal intention and the entire intentional cascade, this way reducing the free energy resulting from the confrontation between the desired state (the representational content of the intention) and the actual state of the world.

Figure 2 shows a theoretical model we originally proposed in the field of technology and human computer interaction (Triberti et al., n.d.; Triberti and Riva, 2015), originally labeled "Perfect Interaction Model" because it virtually represented the interaction in which every intentional level of the user perfectly-interlocks with the characteristics of a technology. Here, we present it as representing general intentional agency. In other words, the model represents the Hierarchical Interlocking of Enacted Intentions.

The model has six levels, three representing the human agent part (distal intention, proximal intention, motor intention) and the other three representing the world part (distal affordances,

proximal affordances, motor affordances). The three arrows show how every intentional level interlocks with a precise level of the world's opportunities for action.

Considering our example: the first arrow relates to the agent who wants to become a psychologist. He starts his own action plan in that the world actually presents the possibility to become a psychologist, containing possible courses of actions and cultural representations associated with this figure (distal affordances). Using Laurent's (2003) terminology, the representational content of the distal intention resembles a "simulated affordance" in that it regards thinkable opportunities in the world to be achieved. The second arrow relates to the here-and-now intention to write an essay; it interlocks with the proximal affordances given by the set or "structure" (Garrett, 2010) of functions guaranteed by the computer the agent decides to use (a technology that allows one to write, cite, and save his own work); finally, arrow 3 represents the interlocking between the motor affordances of the computer (that is, the interface, or the physical representation of the structure of functions) and the motor intentions representing the movements to be performed (moving fingers *this way* to write a given letter).

In the field of Human Computer Interaction and User Experience, the present model is useful to identify the source of interaction failures at the level of intentional representation and/or technological features (Triberti et al., n.d.; Triberti and Riva, 2015) (for example: does the user ignore what to do, and so he doesn't know how to structure action plans of use, or is the technology that doesn't communicate well its own functions?). In this context, we argue that such a model may be useful to show how intentions relates to the external world already at the time when they are not transformed into physical action, through a process of continuous confrontation between representational content and the opportunities in the external world, devoted to progressive reduction of free energy.

As a conclusion, in accordance with both the theory of Intentions by Pacherie (2008) and the theory of Self by Damasio (2011), the presence theory from which we have started highlights that our own actions and intentions are enacted at a three-levels complexity; as motor behavior, based on proprioceptive information coming from our own bodies and their interaction with the physical properties of external objects; as proximal/contextual behavior, based on the perception of environmental functions and affordances; and as future behavior, based on the prefiguring/simulation of action plans. In this sense, any intentional hierarchy layer has to interlock with the respective environmental affordances, being them ready-to-hand physical properties (motor intentions), tools/obstacles actually present in the here-and-now environment (proximal intentions), or conceptualized action plans that are part of the society's cultural background (distal-abstract intentions). The highest the success of the interlocking process at any level, the highest is the sense of being present in a situation, as the result of the impression of being able to transform intentions into actions and controlling one's own agency in the world.

This contribution expands on the previous literature on the topic in two ways: on the one hand, it constitutes the first attempt to link the theory of presence to a modeling of intention

enacting; on the other hand, it deepens the concept of intention highlighting its relationship with the world prior to its enacting.

Indeed, the described process happens on the background of sense of presence, that is, the sensation to be in a given situation emerging from the impression to be able to enact intentions. Sense of presence is not an automatic outcome of the "simple fact"

that one find himself in a given place. On the contrary, we "know where we are" basing on our perceived possibility of being able to pursue our own objectives in distal, proximal, and motor terms.

Thanks to this fundamental process, we state the basis not only for the action plans related to the situated agency of motor behavior, but also of our own distal Self-projecting in future life.

REFERENCES

Carruthers, G. (2012). The case for the comparator model as an explanation of the sense of agency and its breakdowns. *Conscious. Cogn.* 21, 30–45. doi: 10.1016/j.concog.2010.08.005 ; discussion 55–8,

Castelfranchi, C. (2014). Intentions in the light of goals. *Topoi* 33, 103–116. doi: 10.1007/s11245-013-9218-3

Chambon, V., Sidarus, N., and Haggard, P. (2014). From action intentions to action effects: how does the sense of agency come about? *Front. Hum. Neurosci.* 8:320. doi: 10.3389/fnhum.2014.00320

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav. Brain Sci.* 36, 181–204. doi: 10.1017/S0140525X12000477

Coelho, C., Tichon, J. G., Hine, T. J., Wallis, G. M., and Riva, G. (2006). "Media presence and inner presence: the sense of presence in virtual reality technologies," in *From Communication to Presence: Cognition, Emotions and Culture towards the Ultimate Communicative Experience. Festschrift in honor of Luigi Anolli*, Vol. 18, eds G. Riva, M. T. Anguera, B. K. Wiederhold, and F. Mantovani (Amsterdam: IOS Press), 25–45.

Corballis, M. C. (2013). Mental time travel: a case for evolutionary continuity. *Trends Cogn. Sci.* 17, 5–6. doi: 10.1016/j.tics.2012.10.009

Damasio, A. (2011). *Self Comes to Mind: Constructing the Conscious Brain.* New York, NY: Random House.

Eren, A. (2009). Self-efficacy beliefs and mental time travel ability: uncovering a hidden relationship in educational settings. *Learn. Individ. Differ.* 19, 513–517. doi: 10.1016/j.lindif.2009.06.002

Fotopoulou, A. (2014). The virtual bodily self: mentalisation of the body as revealed in anosognosia for hemiplegia. *Conscious. Cogn.* 33, 500–510. doi: 10.1016/j.concog.2014.09.018

Friston, K. (2009). The free-energy principle: a rough guide to the brain? *Trends Cogn. Sci.* 13, 293–301. doi: 10.1016/j.tics.2009.04.005

Friston, K. (2010). The free-energy principle: a unified brain theory? *Nat. Rev. Neurosci.* 11, 127–138. doi: 10.1038/nrn2787

Friston, K., Kilner, J., and Harrison, L. (2006). A free energy principle for the brain. *J. Physiol. Paris* 100, 70–87. doi: 10.1016/j.jphysparis.2006.10.001

Gallagher, S. (2006). "Where's the action? Epiphenomenalism and the problem of free will," in *Does Consciousness Cause Behavior?*, eds W. Banks, S. Pockett, and S. Gallagher (Cambridge: Cambridge University Press).

Garrett, J. J. (2010). *Elements of User Experience, The: User-Centered Design for the Web and Beyond.* New York City, NY: Pearson Education.

Gorini, A., Capideville, C. S., De Leo, G., Mantovani, F., and Riva, G. (2011). The role of immersion and narrative in mediated presence: the virtual hospital experience. *Cyberpsychol. Behav. Soc. Netw.* 14, 99–105. doi: 10.1089/cyber.2010.0100

Laurent, E. (2003). Mental representations as simulated affordances: not intrinsic, not so much functional, but intentionally-driven. *Intellectica* 3, 385–387.

Lee, K. M. (2004). Why presence occurs: evolutionary psychology, media equation, and presence. *Presence* 13, 494–505. doi: 10.1162/1054746041944830

Libet, B. (1999). Do we have free will? *J. Conscious. Stud.* 6, 47–57.

Libet, B. (2010). Unconscious cerebral initiative and the role of conscious will in voluntary action. *Behav. Brain Sci.* 8, 529. doi: 10.1017/S0140525X00044903

Lombard, M., and Ditton, T. (1997). At the heart of it all: the concept of presence. *J. Comput. Mediat. Commun.* 3, 20. doi: 10.1111/j.1083-6101.1997.tb00072.x

Loomis, J. M. (1992). Distal attribution and presence. *Presence* 1, 113–119. doi: 10.1162/pres.1992.1.1.113

Moore, K., Wiederhold, B. K., Wiederhold, M. D., and Riva, G. (2002). Panic and agoraphobia in a virtual world. *Cyberpsychol. Behav.* 5, 197–202. doi: 10.1089/109493102760147178

Numan, R. (2015). A prefrontal-hippocampal comparator for goal-directed behavior: the intentional self and episodic memory. *Front. Behav. Neurosci.* 9:323. doi: 10.3389/fnbeh.2015.00323

Pacherie, E. (2006). "Towards a dynamic theory of intentions," in *Does Consciousness Cause Behavior?*, eds S. Pockett, W. P. Banks, and S. Gallagher (Cambridge: MIT Press).

Pacherie, E. (2008). The phenomenology of action: a conceptual framework. *Cognition* 107, 179–217. doi: 10.1016/j.cognition.2007.09.003

Pacherie, E., and Haggard, P. (2010). "What are intentions?," in *Conscious Will and Responsibility. A tribute to Benjamin Libet*, eds L. Nadel and W. Sinnot-Armstrong (Oxford: Oxford University Press), 70–84.

Riva, G., Botella, C., Baños, R., Mantovani, F., García-Palacios, A., Quero, S., et al. (2015a). "Presence-inducing media for mental health applications," in *Immersed in Media: Telepresence Theory, Measurement & Technology*, eds M. Lombard, F. Biocca, J. Freeman, W. IJsselsteijn, and R. Schaevitz (New York, NY: Springer), doi: 10.1007/978-3-319-10190-3

Riva, G., Mantovani, F., Waterworth, E., and Waterworth, J. A. (2015b). "Intention, action, self and other: an evolutionary model of presence," in *Immersed in Media: Telepresence Theory, Measurement & Technology*, eds M. Lombard, F. Biocca, J. Freeman, W. IJsselsteijn, and R. J. Schaevitz (New York, NY: Springer), doi: 10.1007/978-3-319-10190-3

Riva, G., and Mantovani, F. (2012). From the body to the tools and back: a general framework for presence in mediated interactions. *Interact. Comput.* 24, 203–210. doi: 10.1016/j.intcom.2012.04.007

Riva, G., and Waterworth, J. A. (2014). "Being present in a virtual world," in *The Oxford Handbook of Virtuality*, ed. M. Grimshaw (Oxford: Oxford University Press).

Riva, G., Waterworth, J. A., Waterworth, E. L., and Mantovani, F. (2011). From intention to action: the role of presence. *New Ideas Psychol.* 29, 24–37. doi: 10.1016/j.newideapsych.2009.11.002

Schloerb, D. W. (1995). A quantitative measure of telepresence. *Presence* 4, 64–80. doi: 10.1162/pres.1995.4.1.64

Sheridan, T. B. (1992). Musings on telepresence and virtual presence. *Presence* 1, 120–126. doi: 10.1162/pres.1992.1.1.120

Sheridan, T. B. (1994). "Further musings on the psychophysics of presence," in *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics* (San Antonio, TX: IEEE), 2. doi: 10.1109/ICSMC.1994.399986

Suddendorf, T., and Corballis, M. C. (2007). The evolution of foresight: what is mental time travel, and is it unique to humans? *Behav. Brain Sci.* 30, 299–313; discussion 313–351. doi: 10.1017/S0140525X07001975

Triberti, S., Chirico, A., and Riva, G. (n.d.). "New technologies as opportunities for flow experience: a framework for the analysis," in *Psychological Flow: Empirical Research and Applications*, eds L. Harmat, F. Ø. Andersen, F. Ullén, and J. Wright (New York, NY: Springer).

Triberti, S., Repetto, C., and Riva, G. (2014). Psychological factors influencing the effectiveness of virtual reality-based analgesia: a systematic review. *Cyberpsychol. Behav. Soc. Netw.* 17, 335–345. doi: 10.1089/cyber.2014.0054

Triberti, S., and Riva, G. (2015). "Engaging users to design positive tecnologies for patient engagement: the perfect interaction model," in *Patient Engagement: A Consumer Centered Model to Innovate Healthcare*, eds G. Graffigna, S. Barello, and S. Triberti (Berlin: De Gruyter Open).

Waterworth, J., and Riva, G. (2014). *Feeling Present in the Physical World and in Computer-Mediated Environments*. London: Palgrave Macmillan.

Wegner, D. M. (2002). *The Illusion of Conscious Will.* Cambridge: MIT Press.

Zahorik, P., and Jenison, R. L. (1998). Presence as being-in-the-world. *Presence* 7, 78–89. doi: 10.1162/105474698565541

Biological Agency: Its Subjective Foundations and a Large-Scale Taxonomy

Adelina Brizio[1] and Maurizio Tirassa[2]*

[1] Faculty of Communication Science, Università della Svizzera Italiana, Lugano, Switzerland, [2] Department of Psychology and Centre for Cognitive Science, University of Turin, Torino, Italy

We will outline a theory of agency cast in theoretical psychology, viewed as a branch of a non-eliminativist biology. Our proposal will be based on an evolutionary view of the nature and functioning of the mind(s), reconsidered in a radically subjectivist, radically constructivist framework. We will argue that the activities of control systems should be studied in terms of interaction. Specifically, what an agent does belongs to the coupling of its internal dynamics with the dynamics of the external world. The internal dynamics, rooted in the species' phylogenetic history as well as in the individual's ontogenetic path, (a) determine which external dynamics are relevant to the organism, that is, they create the subjective ontology that the organism senses in the external world, and (b) determine what types of activities and actions the agent is able to conceive of and to adopt in the current situation. The external dynamics that the organism senses thus constitute its subjective environment. This notion of coupling is basically suitable for whichever organism one may want to consider. However, remarkable differences exist between the ways in which coupling may be realized, that is, between different natures and ways of functioning of control systems. We will describe agency at different phylogenetic levels: at the very least, it is necessary to discriminate between non-Intentional species, Intentional species, and a subtype of the latter called meta-Intentional. We will claim that agency can only be understood in a radically subjectivist perspective, which in turn is best grounded in a view of the mind as consciousness and experience. We will thus advance a radically constructivist view of agency and of several correlate notions (like meaning and ontology).

Keywords: agency, interaction, phylogeny, cognition theoretical neuroscience, theoretical psychology

INTRODUCTION

Most, if not all, research paradigms in psychology and in the cognitive sciences agree that the mind is a control system. If behaviorists acknowledged that minds exist at all, they would probably say that they are such systems. Classic information-processing and computational psychologists have often talked explicitly of the mind in such terms. Allen Newell, for example, one of the leading figures in classical cognitive science, wrote that the mind is "the control system that guides the behaving organism in its complex interactions with the dynamic real world" (Newell, 1990, p. 43). Artificial intelligence, Artificial life, and autonomous robotics have a necessary focus on control systems, and the same holds for the most recent trend in scientific psychology, namely the attempt at the integration and cross-fertilization of psychology and the neurosciences.

We basically agree that the main or exclusive function of the mind is to overview and control an organism's activities. However, the notion of *control system* is not easily defined, nor are the nature and the functioning of the particular control system that is the *mind*.

Newell's definition is cast in terms of *behaving organisms*. This definition is interesting under several respects. One is that it adopts a third-person perspective on the topic of action: generally speaking, no agent would conceive of its own activities and actions in terms of *behaviors*. Again, this is typical not only of classical cognitive science, but of most approaches to the study of control systems as well.

Second, Newell's definition entails that any control system that guides a "behaving organism" is a mind, that is that any animal endowed with a nervous system, from the Cnidaria or the Ctenophora (Hatschek, 1888; Moroz et al., 2014) to mammals has a mind, and that the main or only product of an organism's mind is behavior.

Our position diverges from Newell's on both issues. As regards the former, we will adopt a subjectivist perspective of agency, namely that agency can only be understood in the first person, and that the first person has causal powers, that is, it is not a mere epiphenomenon of non-subjective machinery endowed with "real" powers. However, the debate about subjectivity and its role is at least as old as psychology is, and we do not think that proponents of either position might be convinced to change their mind by a "silver bullet" argument of sorts. Thus, rather than arguing in favor of a re-evaluation of subjectivity, we will skip to the subsequent step and try to develop a possible description of cognitive architectures that may follow from such re-evaluation.

We will also object to Newell's implication that any control system from the simplest to the most complex is a mind. The problem with this notion is that it misses too many crucial features of what a mind is: in particular, the issues of meaning and Intentionality. We will propose that all agents are not equal, and that some differentiation be made between different types of biological control systems.

Still another crucial point of Newell's definition is that it talks of behaving *organisms*. It has often been claimed in classical cognitive science that minds may be embodied not only in biological bodies, but also in thermostats, computers, and generally in any physical and computational mechanism capable of supporting the basic operations of intelligence. This is a consequence of the *multiple realizability thesis* that follows from the computational postulate on the nature of the mind (see e.g., Turing, 1950; Haugeland, 1981; Pylyshyn, 1984).

While the ideas that we will discuss in this paper are incompatible with computationalism, we do not have the space here to engage in an examination and criticism of it (for which see e.g., Searle, 1980, 1992; Tirassa, 1994; Manera and Tirassa, 2010; Tirassa and Vallana, 2010). We will circumscribe our discussion to biological entities; it may be interesting, however, to notice that Newell (1990) himself, in apparent contrast with his own previous work (e.g., Newell and Simon, 1976), talks of behaving *organisms*.

Actually, one of the goals of this paper is to provide for a biologically based conception of interaction and agency, and to argue that this requires biology to take a step from its implicitly or explicitly eliminativist positions concerning the existence and the causal roles of subjectivity. Instead, we will claim that agency can only be understood in a radically subjectivist perspective, which in turn is best grounded in a view of the mind as consciousness and experience. We will also advance a radically constructivist view of agency and of several correlate notions (like *meaning* and *ontology*).

WORLD, ENVIRONMENT, AND INTERACTION

Living organisms have several interesting properties that differentiate them from other entities. These can be resumed in the notion that living beings do not passively exist in the world, but actively interact with and within it.

The most apparent manifestation of this property is that they are capable of maintaining, at least within certain boundaries, their own coherence and autonomy in the face of a world which does not take particular care of them. This is not to say that the world is hostile toward them; sometimes it is, of course, but most of the time it is just indifferent.

Coherence means that living organisms have a substantially harmonious anatomic and functional structure, each part of which, under normal conditions, concurs to keep them alive and healthy and participates more or less congruously to the relations that they entertain with the surrounding world. *Autonomy* means that living organisms create and maintain an internal environment which follows dynamics of its own, that are neither completely separated from nor totally determined by the dynamics of the external environment.

Coherence and autonomy are not independent of each other; on the contrary, they shape each other in a dynamic circular relation which lies at the very foundation of life. Together, they provide for the adaptivity of organisms, that is, for their capability of creating and maintaining a dynamic compatibility with the environment in which they are immersed. When such capability falls under certain thresholds, the organism dies.

There are several ways in which living beings stay coherent and autonomous. One is by creating and maintaining a permeable separation between their internal environment and the external milieu. Indeed, it is because such separation exists that talk of an internal environment becomes possible at all. The internal environment differs from the external one in that it is structured (coherent) as well as in several physical and chemical features, which may range from temperature to the concentration of various substances like structural proteins, enzymes, nucleic acids, metabolic products and by-products, and so on. Permeability allows for an adaptive management of such differences, without the need for the organism to become a totally self-contained and self-sufficient universe, which would of course be impossible. In other words, organisms rely for their survival on precisely those world dynamics from which they have to keep, under other respects, a certain degree of separation.

Thus, the second way of maintaining coherence and autonomy is by exploiting the features of the environment that

are relevant to the organism. Most living species, for example, are positively or negatively sensitive to degrees of light, temperature, or the concentration of various nutrients and toxins. Again, organisms have to rely on the external world in order to secure their autonomy, survival, and welfare.

A third way of maintaining coherence and autonomy is by exploiting the biological, behavioral, etc. features of other individuals, belonging to the same species or to others. This process may be viewed as just a special case of the previous one; however, given the importance that it gains in the species that adopt it, it deserves special consideration. Predation, symbiosis, or parasitism are obvious examples of how organisms may interact with each other. As interestingly, many organisms depend on each other for their own survival, as well as for the perpetuation of some parts of themselves, via sexual reproduction and related behaviors. Sexual reproduction in its turn imposes specific constraints on several features of the organisms, ranging from sexual dimorphism to each individual's need to be a desirable sexual mate. The possibility of relying on the others for one's own survival may become so prominent that the individual members of some species may lose their reproductive capabilities—as is the case, for example, with eusociality in social insects and other species where the biological unit is, under certain respects, the community—or end up sacrificing their own life so that other conspecifics may instead keep theirs.

We said above that living beings maintain their coherence and autonomy not thanks to isolation from the external world, but thanks to a delicate and dynamic interaction with it. They actually are immersed in a world with which they need to cope, if they are to survive and prosper.

The notion of *coping with the world*, however, is tricky. No organism could ever cope with all the features of the world. The world is too complex, too rich of dynamics that are more or less independent of each other, for any organism to keep track of the whole of it. It would anyway be extravagant to keep track of all the incidents and occurrences in the universe. Each organism can limit itself to cope with some of the external dynamics that are relevant to its internal environment, its survival and welfare, and possibly its interests and goals, at least as far as organisms with interests and goals are concerned.

This is not merely an issue of physical distance. It is trivially true that what happens in the Andromeda galaxy is practically irrelevant to the survival and welfare of a shrimp in the Atlantic ocean. The real point instead is that most of what "objectively" happens even in the shrimp's immediately proximal world is exactly as irrelevant to it.

The only happenings that do count for an organism are those that potentially affect its well-being, its survival, its reproduction, and its interests and goals. No organism could ever cope with all such occurrences, of course (if it did, it would never die); however, the more occurrences it can cope with, and the better it does so, the better its chances of prosperity are.

Each organism is thus capable of interacting with the occurrences and incidents of the world that are relevant to it; circularly, it is precisely *because* an organism can cope with such occurrences and incidents that they are relevant to it. What happens out of this set of world dynamics may affect the organism

in several ways, at least from the viewpoint of an external observer capable of noticing the ongoing events, but will not be relevant from the organism's viewpoint. Think, for example, of radiations: since an organism does not interact with them, they simply do not exist to it, even if they kill it.

What is the nature of the occurrences and incidents that are relevant to an organism? Apart from a few physical and chemical parameters, most of them have nothing to do with what we consider the world's fundamental dynamics. We are accustomed to believe that the universe is made up of protons, neutrons, and electrons, or of even smaller particles, of electromagnetic waves and other forms of energy, and so on, and that world events are made up of movements, variations, and transformations of such entities and aggregates thereof. However, no control system of a living organism deals directly with electrons, atoms, or molecules (except, of course, for that of human researchers in fundamental physics, on working hours). What living beings deal with is air, food, preys and predators, rivers, trees and mountains, sexual mates, water, parasites, paths, obstacles, and dangers: all entities that, whatever their true or ultimate structure and composition, are nonetheless interesting as entities in themselves, characterized by dynamics of their own that are not reducible to those of fundamental physics and chemistry. And, of course, from the viewpoint of most organisms there just are no such things as *fundamental physics and chemistry*: as far as we know, these notions are characteristic of a comparatively small, historically given subset of the human species.

This means, in practice, that it is the organism itself that selects which, of all the happenings in the universe, are relevant to it: which peculiar configurations of atoms and energy are food, water, threats, opportunities, mates. In this sense, the environment is, to each organism, subjective: not because no external reality exists, independent of the organism, or because such reality cannot affect its welfare or its very existence, but because each organism is only capable of interacting with certain dynamics of the universe, with which it interacts according to the ways and ends that its nature allows.

Thus, a certain dynamics of the external world can be inconsequential to an organism, or influent but not relevant (because the organism is not equipped to interact with it—something that only an external observer might appreciate), or relevant in one way or another (being a threat, an opportunity, something edible, something with which to reproduce, and so on).

In other words, it is the internal dynamics of an organism that create the external ones and, in making so, give them meaning. Meaning is, therefore, subjective.

Each organism thus creates a certain set of external dynamical meanings with which it then interacts. The specific ways in which it does so are dictated by the biology of the species to which it belongs as well as, in part, by its own ontogenetic trajectory. It is the features of the organism's biology that establish its physical and chemical requirements, what kinds of geographical environment it will find more suitable, what food it needs and how it can recognize and obtain it, what dangers it has to defend from, whether, why and how it has to interact with its conspecifics, and so on. Within the same species, there always

is room for individual variations. These may be due to small differences between the genotypes or between the phenotypes to which they give rise in their interaction with the environment (Lewontin, 1998), to learning, to individual preferences, and so on. However, of course, the similarities between the members of a certain species are much greater than the differences.

There can instead be remarkable differences between species. In evolutionary time, phylogeny has generated and selected many different ways to be in the world: all of them are equally subjective, all equally legitimate, all more or less equally compatible with the real world, all more or less equally capable of taking care of the organisms' explicit or implicit interests.

The nature of the relation occurring between an organism and its environment thus depends, in the last analysis, on the interaction between its genotype (which results from evolution), its phenotype (which emerges from the interaction between the genotype and the ontogenic environment), its idiosyncratic developmental paths, and the environment itself.

THE EVOLUTION OF CONTROL SYSTEMS

The range of processes that may fall under the label of "an organism's interaction with its subjective environment" is very broad. Multicellular organisms typically interact with the environment on many levels: from the exchange of water, ions, and metabolites, to processes aimed at recognizing and destructing non-self molecules and microorganisms, to the production of substances that may affect the metabolism, the growth or the behavior of a conspecific, an aggressor or a prey, and so on. While the final operative details of each of these functions are typically assigned to a specialized organ, apparatus, or system, all of an organism's parts participate, in one way or another, to the organism's interactions with the external environment.

In the phylogeny of animals, furthermore, an organ—or, better still, a whole system: the nervous system—has developed which specializes in the peculiar function of centralizing, governing, and coordinating, at least to a certain extent, certain features of the interaction with the environment.

The initial appearance and evolutionary success of this system were in the service of movement. To be capable of moving from a sunny area to a shadowy one and vice versa, or of actively searching for water and food or fleeing from a danger is extremely useful to an organism. This requires a behavioral coordination that can only be achieved by a specialized management of the relevant processes.

Circularly, the nervous system originated from sensory-motor circuitry for phototaxis and chemotaxis. Cells endowed with electric properties enabling them to react to light or chemicals progressively differentiated into what will become a system of specialized sensors and effectors and, later on, a system in charge of coordinating and mediating the management of the various contingencies and opportunities that the animal meets in the environment.

This type of circular relation is typical of biology (Mayr, 1997; Gould, 2002). Organisms do not evolve because the world forces them to do so: that is, they do not progressively become more adapt to an objectively given environment which "poses problems" that they must "face" by evolving (see also Gould and Lewontin, 1979).

Something very different happens instead: each organism creates, through its peculiar ways of interacting, a subjective environment of its own (an *Umwelt*, in Uexküll, 1934, words), and it is to this that it, circularly, is adapted. The "problems" that living beings must "face" are not in the world, but in their interaction with the world; analogously, the "solutions" that evolve are not in the organisms, but in the interaction that they have with their subjective environment. Thus, there are neither real "problems" nor real "solutions": only more or less successful ways of creating and maintaining sustainable *Umwelten*. A significant genetic mutation can, under certain conditions, give rise to a new species, characterized by different ways of interacting with a different subjective environment.

Of course, each organism's subjective environment has to be compatible with the real world, whatever its ultimate nature may be. This is, indeed, a matter of compatibility, not of problem solving. Evolution is not a movement on the gradient of adaptation toward optimality.

Thus, as far as nervous systems are concerned, their appearance in phylogeny is by no means necessary: the "real world" does not "pose problems" requiring coordination and control to an organism which, having no such capabilities, finds itself in the need to evolve them. Actually, countless living species have no control system whatsoever, least of all one made of neurons. On the other hand, nervous systems, once appeared, significantly alter the types of interaction that organisms can generate. Coordination and control become salient features of such interaction, which gives rise to still different types of interaction, and so on. As a result, a whole evolutionary lineage takes a wholly new course.

By the same line of reasoning, not all control systems need be similar both in structure (which is obvious) and in the type of interaction they create and maintain with the world. We will propose an outline of two very large-scale types of such systems, which we will call *non-Intentional* and *Intentional*, based on an analysis of the possible types of interactions with the world. A particularly interesting subtype of Intentional architectures called *meta-Intentional* will also be described.

Intentionality, or aboutness, is a mind's property of being able to entertain semantic (meaningful) relationships with the world. In philosophy (e.g., Searle, 1983), artificial intelligence (e.g., Rao and Georgeff, 1992), ethology (e.g., Prato Previde et al., 1992), and cognitive science (e.g., Airenti et al., 1993; Tirassa, 1999b), Intentionality[1] often is characterized in terms of mental states like various types of desires, beliefs, and intentions, with propositional content. This is often couched within the computational postulate about the nature of the mind, whereby cognition consists in the syntactic manipulation of symbols (Manera and Tirassa, 2010; Tirassa and Vallana, 2010).

---

[1]Conventionally (e.g., Searle, 1983), Intentionality as *aboutness* is written with a capital initial, so to immediately distinguish it from the acceptation of *having an intention* or *doing something intentionally*.

The computational postulate, however, is far from being unanimously accepted, and other areas of literature tend to equate Intentionality with consciousness or phenomenal experience or, as it is often said, to view representations as happening at the interaction of the conscious mind/brain and the external world (see e.g., Heidegger, 1927; Merleau-Ponty, 1945; Nagel, 1986; Varela et al., 1991; Searle, 1992; Varela, 1996). This means, among the rest, that the mere neural coding of sensory stimuli does not count as representation (e.g., Clark, 2001).

This is also our view. For the analysis that follows, a simple definition of Intentionality as synonym with aboutness, semantics, and phenomenal experience will suffice.

TYPES OF CONTROL SYSTEMS

What nervous systems do is to mediate in wholly new ways between the animal's internal dynamics and the external ones. Simple mobile animals are characterized by taxes, that is, movements in space along physical-chemical gradients such as light, temperature or the concentration of certain molecules. When nervous systems appear in phylogeny, taxes are substituted for by locomotion, that is, active movements, endogenously generated from internal states like the variation of certain physiological parameters, the perception of relevant entities in the external world as well as, a few millions of years later, desires and opinions.

An animal's internal dynamics end up being the very center of its interaction with its subjective environment. This leads to the differentiation of several types of internal dynamics, and therefore to a progressive increase in the structural complexity of the interaction that the animal is capable of generating.

Of course, the control systems of animals are not all alike. The very anatomy and physiology of control systems are strikingly different across species. Differences in the types of interaction generated follow accordingly.

Non-Intentional Control Systems

The evolution of nervous systems may be described in several ways. In principle, each such description, if correct, should match the others; however, current knowledge is far from providing maps of such precision. What is available is, on the one side, a great deal of punctiform analyses of single behavioral or cognitive functions in single species and, on the other side, a few rough decompositions as tentative identifications of wide phases in the phylogeny of control systems. Our discussion will fall into the latter area: based in part on previous work (Tirassa et al., 2000) we will propose an extremely large-scale classification of control systems, from the viewpoint of their (hypothetical) subjective functioning.

After the appearance of control systems, the next crucial step corresponds to the transition from invertebrates to vertebrates, characterized among the rest by the appearance of encephalic and cortical structures (Gans and Northcutt, 1983). Such anatomic transition is likely to find a functional correspondence in the appearance of phenomenal experience in the proper sense, that is, of consciousness, or awareness, or at least of object-based cognition, which is its primary manifestation. Despite several studies (e.g., Menzel and Giurfa, 2001; Menzel et al., 2006;

Bateson et al., 2011; Mendl et al., 2011; Gibson et al., 2015), we still do not know whether and how invertebrates are conscious. Even if they were, however, there would probably be nothing that they would be conscious of; at least, nothing in their behavior lets us think so (see e.g., the discussion of phonotaxis in crickets in Clark, 2001; Hedwig and Poulet, 2005; Hedwig, 2006; Hennig, 2009). Even in the apparently most complicated cases (such as the famous "dance" of honeybees) their behavior can invariably be explained in terms of comparatively simple transformations of sensory signals into motor commands (see e.g., Kesner and Olton, 1990; Wehner, 2003; Poulet and Hedwig, 2005).

Winged insects, for example, prepare for landing as a reaction to the visual expansion of a texture from below, signaling a surface rapidly getting closer. Analogously, they prepare for flight as a reaction to the contraction of a texture from below (signaling a surface rapidly getting farther) as well as to the expansion of a texture from above (signaling a potential danger approaching; Lindemann and Egelhaaf, 2012). This non-object-based mechanism is comparatively simple and extraordinarily effective; so simple and effective, indeed, that it has not undergone significant evolution over the last several million years.

However, nothing in the behavior of winged insects makes us think that they have any semantics for, or conscious experience of, surfaces for landing or of dangers approaching. To them, a texture is worth another, provided it activates the takeoff or landing mechanisms. They appear to be completely unable to discriminate between a rose petal falling and the newspaper with which an exasperated human is striving to kill them.

If these animals are conscious, we are unable to understand *what* they are conscious of, and, therefore, *why* they should be conscious at all. One hypothesis that could be made is that consciousness might be a necessary property of any neuronal system (but then, even of a neuron in isolation?). Such hypothesis, however, is not necessarily better than others, nor would it explain what such animals could be conscious of.

Intentional Control Systems

Intentionality is the property of entertaining semantic, that is meaningful, relations with the world (Brentano, 1874). Its appearance is a big turn in phylogeny: while the case of invertebrates is uncertain, it can safely be claimed that vertebrates are conscious because they experience the world, that is they give meaning to it. Intentionality can only be conscious (Searle, 1992): experience requires a point of view (in a sense, experience is a point of view) and a point of view has to be *someone's* point of view, therefore subjective (Nagel, 1986). Thus, there is no experience without subjectivity and, of course, no subjectivity without consciousness.

This conception is different to the general consensus in mainstream psychology. Most scientific paradigms from the nineteenth century to the 21st build on the assumption that consciousness is substantially irrelevant, or only marginally or occasionally relevant, and that what really matters is unconscious, non-subjective, non-meaningful knowledge and processes. Although a thorough discussion of these issues would fall outside the scope of this work, it may be useful to spend a few words.

In the perspective we are trying to outline, the mind is neither an epiphenomenon (as it is in behaviorism, neural reductionism, and related forms of eliminativism) nor a set of descriptions (as it is in classical cognitive science and computational psychology). Both such accounts are untenable for philosophical (e.g., Searle, 1980, 1992; Nagel, 1986; Johnson, 1987; Varela et al., 1991) and biological (e.g., Edelman, 1992; Varela, 1996) reasons; furthermore, they are, in a sense, equivalent, in that both are rooted in, and possible consequences of, dualism (Tirassa, 1999a). Instead, the mind is a material property of the brain, which means that cognitive causation does not go from brain to brain and from brain to mind, but from mind/brain to mind/brain; better still, since brains do not grow in vases, from mind/body to mind/body.

That we have no hint at how this is possible, that is, at how a few kilograms of seemingly undistinguished matter may have the property of being subjective, does not make it less real. Actually, all existing theories are as obscure on this point: how does subjectivity occur as an epiphenomenon? how does it emerge from computation? Therefore, instead of endlessly arguing in favor or against the various views, we will just try to go on, trying instead to develop a few consequences of our approach, from which it might be judged more aptly.

To equate mind, consciousness, subjectivity, meaning, and experience also means that there is no Self, if the word is taken to refer to a *homunculus* or mental entity which is abstracted from, and exists independently of, space and time. Instead, the mind re-creates itself from instant to instant. The sense of continuity that we perceive depends on the fact that each "slice" of such re-creation is causally generated by the meshing of the preceding one and the current interaction with the world, and so on, back in time, up to the very first instant when our mind began to exist.

Each such "slice" in the functioning of a mind/body literally is the product of the previous history of that mind/body, plus its current interactional dynamics. The patterns with which the mind/body re-creates itself are rooted in the evolutionary history of the species to which the organism belongs. At least in certain species, such patterns also depend on and, circularly, generate individual differences whose roots are to be found in genetic variations between individuals as well as in the details of each individual's interaction with the world (that is, its ontogenetic history). In a metacognitive species like ours, the latter includes the mind's interaction with itself (that is, its autobiography) and with other minds and the artifacts they generate (that is, social and cultural history).

Another point worth remarking is that this conception of the mind requires that it be identified neither with attention (at least because we are conscious of several things on which we do not focus our attention) nor with abstract or formal reasoning, language, or self-awareness. The latter capabilities depend on the existence of consciousness but are not identifiable with it, both for analytical reasons and because most Intentional species do not appear to possess them.

INTERACTION AND BEHAVIOR

Nervous systems, we said, are control systems. What they control is not, as in Newell's definition cited in the introduction, the organism's *behavior*, but its *interaction*; or, at least, certain features of the overall interaction. There are many reasons why this remark is crucial.

The first is that behavior is a third-person term. Behavior only exists in the eye of an observer who pursues interests of its own, not in that of the "behaving" organism. Control systems do not behave: they work in the first person.

Actually, non-Intentional control systems only work in the first person in a very peculiar sense, because, as we saw, there is no reason to think that "there is anybody home." Notwithstanding, what such systems produce is interaction anyway, but one of a kind that consists in the mere activation of motor patterns starting from the meshing of sensory patterns with relevant physiological internal states.

Another reason why control systems are better said to produce interaction than behavior is that the nervous system does several things beside "producing behavior" (whatever acceptation the term is given) or even "reasoning." Many bodily functions fall, wholly or partly, under its jurisdiction: it controls, for example, the activity of the cardiovascular, the respiratory, the digestive, and the endocrine systems.

These activities are not independent of, or separate from, the generation of interaction. Notwithstanding their anatomical, physiological, and functional variety, the components of the nervous system work in strict synergy, which, of course, is precisely why it is a *system*. All of them are directly or indirectly interconnected, so that the nervous system is best described as a single interaction-producing network of cells. While some parts of it are more involved than others in the various aspects of interaction, it certainly is not an assemblage of "encapsulated modules" working in isolation from one another.

Consider, for example, what happens to a mammal who perceives a predator—say, a gazelle and a lion. For a start, such "perception" is mental activity: what we mean when we say that the gazelle perceives the lion is that it views a certain entity in the world as a specific type of potential threat to its security. To say that this experience is subjective does not imply that the lion literally is a creation of the gazelle's mind, but that its meaning is. "Predator" is a semantic relation between the two animals in the current situation, not an intrinsic property of one of them: to an elephant, a cat is much less a predator than it is to a mouse; to the mouse, a lion is much less a predator than it is for a gazelle; to the gazelle, a lion on the horizon is much less a predator than it is a lion 100 m away. To a hypothetical gazelle equipped with an armor plate and pump-action rifle, the lion would be nothing more than an occasional nuisance[2].

When an animal perceives such a danger, its blood pressure rises, as the result of a variation in heart activity and in the total

---

[2]It could be objected that "predator" is a word, and that gazelles have no language. Of course we do not think that the gazelle says to itself "Alas, another goddamn lion… alright: time to run." We have already argued against the identification of semantics and language. It could also be objected that the lion will actually eat the gazelle if it catches it, and that there is nothing semantic in this, and therefore that the lion is objectively a predator. Again, true, but we are concerned here with the gazelle's mind, not with ethology as construed by human scientists: it is unlikely that the gazelle views the ongoing situation in terms of the nutritional habits of *Felis leo* as they are portrayed on educational TV channels.

caliber of its blood vessels. Furthermore, there is a redistribution of the blood flow away from certain districts (like the digestive system) and toward others (like the brain and the locomotor system). Respiratory frequency increases. Several hormones and other substances are released in the blood and go to affect the functioning of an array of organs and apparatuses. As a result of these complex modifications in its mind/body, the animal will become prepared to a fight-or-flight activity. Such condition will modify in turn the subsequent flow of the animal's mental dynamics; e.g., it will be frightened, but also more ready, compared to what would have been in a different situation, to search its subjective environment for certain relevant affordances. The gazelle might, for example, act so to draw the lion's attention away from its offspring; it might recognize a river not as a reservoir of drinkable water, but as an obstacle for its enemy, who might be reluctant to cross it; it might see the herd as a source of salvation and safety, and so on.

Talk of instinct here would be correct and misleading at the same time: correct, because the reconceptualizations that the gazelle does of its subjective environment can hardly be viewed as the result of sophisticated reasoning, or of preceding experience with similar situations (although, of course, the latter may certainly play a role). Misleading, because it is extremely unlikely that the control system of the gazelle is hardwired to do things like "looking for salvation beyond the river if a lion is behind me, the herd is too far removed, and the river is wide enough to be a problem for it; but only if the offspring is safe." This would only be possible in a giant lookup table like those of computational psychology, and mind/bodies simply are no such tables (if only because there is no *homunculus* inside who could look them up). Without representations, no component of the gazelle's subjective situation (the lion, the herd, the river, and so on) could be present to its control system; and, without a however small leap of intuition and creativity, no acknowledgment of the possible moves and their comparative chances of success, and therefore no situated decision, would be possible.

Thus, the notion of instinct is of little help. What happens is instead that the animal continuously reconceptualizes the surrounding environment. This is certainly made possible—indeed, generated—by the gazelle's specific biology, comprised of its phylogeny and ontogeny, but no less mental for this. What is misleading about the notion of instinct is the impossibility for such label to capture the actual nature of Intentional control systems. The advantage that such systems offer, compared to non-Intentional ones, is precisely that they work on dynamic flows of meanings, not on hardwired sensory/motor relations.

Reconceptualization means that a tight semantic coupling to the world is maintained. Meanings are not in the world, but in the animal's experience in each moment. The animal's past history is crucial in the generation of its current experience, not because it has been coded and stored for future reuse, but because it is what has led the animal to the particular state in which it currently is. The mind exists exclusively in the present, but such present is the child of the past, and results from the integration of the past with the world as it is now (Glenberg, 1997).

An agent's cognitive dynamics across time results from the interaction of its mind/body with the surrounding (mental,

bodily, physical, and social) environment. Interaction at any instant t_i is causally generated by the state in which the mind/body was in the instant t_{i-1} that immediately preceded, together with co-occurring factors that may affect its functioning, like the activity of sensory receptors, emotional and thinking processes (whatever they are in each species), the effect of various blood chemicals, and so on.

Each "slice" of a mind/body dynamics thus is the product of that dynamics so far, meshed with the meanings found in the interaction in which the animal is currently immersed. This way, it is neither a mindless body nor a disembodied mind that causes the overall dynamics: the state of an agent's mind/body at any slice of time plays a causal role in the state of that mind/body at the subsequent slice of time.

In its turn, interaction at t_i will contribute to generate the state of the mind/body in the instant t_{i+1} that will immediately follow. Thus, in each instant interaction results from all the interactions at t_{i-n}. Memory and learning are to be understood as modifications of each possible future experience, rather than independent, switch-on/switch-off "cognitive functions." The pattern of development of such history results from the biology of that particular organism, so that it definitely is not a matter of nature vs. nurture (or of rationalism vs. empiricism) that we have here. Viewed from a biological vantage point, these are false dichotomies (see, for example, Lorenz, 1965).

Intentional systems thus are in no way "less biological" than non-Intentional ones, unless one believes that biology needs to be eliminativist, and such eliminativist biology is then opposed to a mentalist psychology supposed to have nothing to do with biology. That this conception has ruled both disciplines for several decades is the consequence of the acceptance, on both sides, of Cartesian dualism and its legacy. There is no reason to accept such position; indeed, there are several reasons to reject it.

Let us now go back to Serengeti. The gazelle that is fleeing from the lion is not "behaving": it is giving an overall meaning to the subjective environment in which it finds itself, and reconceptualizing in its light the environment itself, looking for affordances relevant to such meaning.

Of course, the lion who is chasing the gazelle does, in its own way, the very same.

AGENTS

We can now define an agent as an Intentional, conscious organism who lives in a situation, and strives continuously to make it more to her liking[3]. What we call the *situation* is a subjective, dynamic, and open map of the world.

This definition allows to exclude several entities that, albeit self-propelled, do not act in any sense of the term (if not, possibly, in a metaphorical one): household appliances like the thermostat

---

[3]Our definition is akin to Pollock's (1993); however, his was cast within a computational perspective instead of a first-person or consciousness-based one, and thus developed in wholly different directions. Analogously, our use of the word situation has nothing to do with situation semantics (Barwise and Perry, 1983). Also, we are aware that a clinical psychologist, a psychotherapist, or a sociologist might object that the locution "…to her liking" is extremely ambiguous under most circumstances.

that operates the air conditioning in this room, the computer on which we are writing this paper, or, in a very different fashion, the mosquito that, aware of nothing, is flying around us, trailing the heat of our body and the carbon dioxide that it produces. Yet, each of these entities or living beings has been characterized as an agent proper in other paradigms within the cognitive sciences.

In our definition, an agent can only exist in biology if it entertains with the world the kind of relation that Maturana and Varela (1980) call *structural coupling*. The Intentional features of this relation were discussed in a previous section, as well as the remark that they must satisfy a constraint of compatibility, not correctness, with respect to the real world.

One implication of the latter consideration is that several different types of agents can exist in principle, and do indeed exist on this planet. Actually, there exist as many types of agents as representational species, and smaller differences occur between the various individuals that belong to each. Each (type of) agent will see its own set of world dynamics and possibilities for action.

Human beings assume that they can entertain objective knowledge because they are capable of generating descriptions of the world and exploiting them for action as well as for intersubjective, publicly shared, and agreed-upon consideration. However, this is only our specific way of knowing, compatible with the ultimate truth (Kant's *Noumenon*, 1781) but neither closer to it nor more objective than that of other species (Nagel, 1974, 1986).

Let us consider again the coupling between an agent and its world. What is coupled is the agent's internal dynamic and the external ones. The world has dynamics of its own, which depend on its properties at the various levels that can be considered: chemical, physical, geological, meteorological, astronomical, biological, and so on. For the scope of this paper, however, it will not be necessary to discriminate between these diverse entities: we will just gather them all under the label "external dynamics."

An animal's internal dynamics include the meanings that it finds in the external ones; circularly, the external dynamics may be said to be generated by the internal ones. Each external dynamics thus corresponds to one of the entities that are interesting for the animal and, while such entity is present to the animal's mind (that is, while it subjectively exists), it is a continuous flow of mutable, self-modifying meanings.

It is necessary to conceive of such dynamics as *flows* because the world is mutable. To the gazelle, a small dot which is rapidly getting closer can suddenly turn into a lion, but then it may begin to chase another member of the herd, or it may reach too close to the place where the younglings are. The river may be too rapid to cross, but then open into a slower bend that permits safe wading; but, with the lion getting closer, the perceived dangerousness of the rapid trait may decrease to the point that attempting to cross it becomes preferable to being killed. A control system has to view the world as flows because the world is a flow, and even more so is the subjective environment in which an agent lives.

It is necessary to conceive of such dynamics as flows of *meaning* because what counts is the meaning of the various entities, that is, the role that they play in the agent's overall situation and the actions they afford. Control systems are not there to dispassionately, disinterestedly compile inventories of the entities that exist in the universe, but to do something with them: eat them, fight them, take care of them, ignore them, have sex with them, and also—why not, for a species like ours?—put them into inventories, but meaningful ones. The situated roles and affordances that characterize each entity are not separate from the entity itself, or a later attachment to an otherwise objective, neutral knowledge: instead, they are the very reason why the mind exists.

It is necessary to conceive of such dynamics as flows of meaning *within an overall situation* because each flow exists not in isolation, but relative to the others that the animal "views" in each moment. The meanings that each flow has in each moment depend on the current state of the overall situation and in their turn contribute to establishing it. The gazelle who is running from the lion will view the river as an opportunity for salvation, rather than as a good place for a rest and a drink, not because its mouth is not dry, or because the water in the river ceases to be drinkable when there are lions in the neighborhood, but because the point of the situation has nothing to do with mouth dryness and drinking. Thus, it is the dynamic meaning of the overall situation that gives the river its meaning as possible salvation, and running toward the river modifies the overall situation. It may, for example, give a sense of imminent salvation that has the gazelle double its efforts, while finding itself in the wide, open land might give it a sense that salvation is out of reach, and thus induce it to accept an opportunity for fighting instead.

ACTIONS

An agent lives in a complex situation, made up of dynamic flows of meanings.

Each such flow may subjectively be more or less pleasurable; or it may be neutral, which only means that it is neither particularly pleasurable nor particularly unpleasant. Most of the times, a flow will be more pleasurable under certain respects and less under others. The greater or lesser pleasantness of each dynamics depends both on the dynamics itself and on how it fits into the overall situation[4].

To act is to alter such flows so to make the overall situation more pleasurable. The change in the animal's overall situation depends on specific interventions upon specific features of the subjective environment. Each flow of meaning that contributes to making up the overall situation may offer opportunities for action; to intervene upon one or another such flow depends on their respective pleasantness, on their respective contributions to the pleasantness of the overall situation, on the apparent possibilities for successful action and, all in all, on the balance of contingencies and opportunities that the agent views in the world.

Since the world is dynamic, and follows causal paths of its own, to act is to interfere with one or more such paths so to alter its spontaneous evolution. An action thus is an induced

---

[4]Of course, such a one-dimensional conception of emotions and motivations is definitely too rough; however, it may do for our current purposes.

modification of the dynamics that the subjective environment would otherwise undergo. Since the agent's capabilities for action are limited, the agent will focus on one such dynamics, or on a few, leaving the others to their natural course.

Furthermore, since the world is dynamic, to act requires monitoring its spontaneous evolution, interweaving one's own moves with it and managing to dynamically coordinate the relation between action and world. This requires at least a minimal capability of prediction of what the evolution of the world will be with or without the agent's interference, or with different possible interferences. Action is intrinsically situated: were it not so, there would simply be no action at all. Of course, these capabilities of monitoring, prediction, and coordination will be different between species (and, within each of them, between individuals). Each species lives in its own type of subjective environment, and acts within it.

Since the subjective situation is mutable, the agent will move from one flow of meaning to another, always trying to make the overall situation more pleasurable. This process is continuous and seamless. Exactly as the world is a continuous dynamic flow, so are the agent's mind and actions.

Actions, thus, have no beginning and no end other than the points in time when the agent sets their beginning and their end; and they are not chosen out of a repertoire which univocally defines their preconditions, effects warranted by default, and procedures of execution, as it happens instead in most classic theories of planning, both in psychology (e.g., Newell and Simon, 1972; Shallice, 1982) and in classic artificial intelligence (e.g., Fikes and Nilsson, 1971; Russell and Norvig, 2009), or even in ethology, with the notion of ethogram (Jennings, 1906; Makkink, 1936). The surfer who rides the ocean waves, exploiting their push, trying to keep her balance by simultaneously *following* the waves and *fighting* them, provides a better metaphor of an agent's life than the game of chess does, with its discrete and precisely defined moves carefully picked out of a closed repertoire and staged in a closed world, where nothing happens except the moves themselves, one at a time.

In other words, there is no intrinsic ontology of actions, except for the one that the agent will throw in at each moment. Similarities between situation/action couples, of course, allow an observer to generalize and abstract, but that does not mean that such generalizations capture a natural subjective ontology. The subjective ontology of action will depend, moment by moment, on the situation in which the agent finds itself and on the interests that it pursues in it.

This shows particularly well if we consider what might be called *the granularity of actions*, or, better still, *the minimal unit of action*. When we say that "an agent is doing something," what do we mean, precisely?

What the agent does is to alter, in a direction which it foresees as favorable, the spontaneous evolution of the world, by leveraging on one of its characteristics. The ontology of the representation that the agent has of the world dynamics are not predefined, but they are created, moment by moment, according to the agent's interests and to the contingencies and opportunities that it views in the world. The same holds for actions, which are the external counterparts of representations. The ontology of

action is created moment by moment, because that is also how what the agent represents is generated.

This may be viewed as a reformulation of the idea that what a representational animal does is to live within its situation, not to behave in greater or lesser accordance to the descriptions that an external observer might give. Furthermore, there can exist no repertoire of possible actions stored in an unconscious subsystem placed out of the here and now, if only because no stored recipe could be coupled to the current state of the world.

Thus, there also is no fixed minimal unit of action; at each time, the minimal unit of action will be what the agent decides it to be. To look inattentively at a landscape from which no danger is expected, while slowly grazing in the grass, is as much an "elementary" and "unitary" action as it is to focus in sudden alarm on a particular dot in the landscape, wondering whether it could be a lion.

Nor would it be a good idea to consider "elementary action" the minimal body movement possible to an individual. This would make no sense from the psychological or the physiological points of view, because we do not usually reason in terms of physical movements (except when we are learning a new movement, or when a breakdown occurs during action) and because to define such minimal movement would be impossible: the activation of a single motor neuron? and on what temporal scale?

The perspective that we are trying to outline relies on a radically non-dualist conception of the mind/body. The subjective situation in which the agent finds itself includes at least (for several species, *only*) visual, auditory, olfactory and other types of perceptions, as well as proprioceptive information concerning posture, what parts of the body are in touch with what, and so on. Such information may have variable degrees of granularity, according to the global properties of the situation and to the contingencies and opportunities that the agent views within it.

To act is to coordinate these information with the decision one is making, reconceptualizing at each moment the position that the head, the eyes, the limbs and the rest of the body should have. The realization of an action is thus the direct counterpart of the intention to perform it: it requires nothing more than that, nothing that is not already "contained" in the intention.

This makes no sense in a dualist "mind sends commands to body" perspective or in an eliminativist "mindless body moves according to instincts, neural firing, or reinforcements" one. However, it becomes reasonable in an Intentional view of the mind/body as one of the material properties of a control system which includes the whole nervous system, as well as its relations with the rest of the body and the surrounding environment[5].

---

[5]In any case, there seem to be no other possibilities. Computational psychology, cast in terms of libraries of operators (that is, of an abstract, objective, predefined ontology of actions), must at a certain point deliver the responsibility for their realization to non-representational capabilities "of a robotic kind" (see for example McDermott, 1987; Harnad, 1994). At the other end of the spectrum, Searle (1983; 1992) claims that a "nonrepresentational Background" is in charge of the task, but then his description thereof remains somewhat mysterious. One of the advantages of the position that we are advocating here is that it makes sense of the relations between phenomenal dynamics, body dynamics, and external dynamics.

Thus, what happens is not that the agent represents a goal and, while time magically stands still, searches an inner store for the action(s) that will provably realize such goal, and sends the decision through a descending hierarchy of "levels of abstraction" until it somehow is translated from the cognitive into the bodily, becoming a sequence of commands delivered to the effectors for execution. What happens is instead that the agent singles out, in its subjective environment, a certain dynamics which offers some desirable opportunity (what might be called an attractor), and in so doing it reconceptualizes the whole of its own mind/body system in the realization of the relevant intervention, remaining at each moment coupled to the dynamic subjective environment. Perception, decision, action, and feedback are not different phases, possibly assigned to different subsystems, but different viewpoints that an observer may take of the mind/body, while the mind/body simply coordinates with the world.

This conception owes a lot to Gibson's (1977, 1979) notion of *affordance*. In a possible reading of his work, the entities of the world present themselves as attractors, variously positive or negative, that are afforded (hence the neologism) to the animal. Affordances are neither in the world nor in the animal: they reside in the interaction between the two. What the world puts in the interaction (and makes Gibson talk of *direct perception*) is the resources and the constraints to which the animal's control systems must conform, such as the invariants of the optical flow. What the animal puts in the interaction is its own nature, which makes a certain configuration of light, a certain texture, and so on, take the subjective shape of a certain affordance[6].

META-INTENTIONAL ARCHITECTURES

We defined an Intentional agent as a conscious organism who lives in a subjective, open, and continually revised interpretation of an ultimately unknowable environment—what we call the agent's *situation*—and strives to make it more to its liking. The agent's mind is the experience of a complex flow of meanings, and meanings are dynamical affordances.

Let us now go back to the control systems of animals. In our extremely large-scale theory of the phylogeny of control systems, the first big transition which they underwent is the appearance of representations. The second is the appearance, in one or few evolutionary lineages within mammals, of what we will call *meta-Intentional* control systems.

Let us start from a related notion, that of *metacognition*. This term was first introduced by Flavell (1979), who defined it as the ability to think about thinking, and has since been used mostly in the area of human social cognition and communication (e.g., Tirassa and Bosco, 2008). However, it is misleading insofar as it seems to refer to a set of capabilities that make up a supplementary, "upper cognitive layer" that adds to a "base cognitive layer" without actually changing the meaning of the latter, but simply manipulating it or exploiting it when needed. This is the case, for example, with logical and formal metalanguages, upon which the notion of

---

[6]We are not claiming that Gibson would agree with our proposals, but only that we have been influenced by our understanding of his work.

metacognition is framed. Such conceptions, however, cannot be applied to psychology or biology, precisely because they rely on a propositional (that is, formal, syntactic, and recursive) notion of mind, which could only work under the assumption that there is a *homunculus* inside who is in charge of operating the system, knowing when and how to nest the propositions, how to manipulate them, and so on, meanwhile losing meaning (Searle, 1980).

The mind is one; it is not composed of layers, least of all layers of nested computations. "Metacognition" can in no way be independent of, separate from, or placed above an alleged "rest of cognition"; on the contrary, it is intrinsic to the human way of knowing the world, that is, to the internal dynamics of the human mind. The dynamics that we see in the world, their pleasantness, and the affordances that they offer are immediately and intrinsically made different by such capabilities.

In our proposal, meta-Intentionality refers to a whole constellation of interwoven capabilities that can be resumed as the idea that the individual itself—including its body, its mind, its history, and so on—becomes part of what is represented. Roughly, while Intentional control systems can be said to experience the world, meta-Intentional ones can be said to experience themselves in the world.

Meta-Intentional minds actually are a subset of the Intentional minds, but one that is enough interesting to warrant a separate discussion. Because our comprehension of the minds of the other primates and of the cetaceans still is so unsatisfactory, the only known species that can safely be said to belong to this class is ours; however, that at least one meta-Intentional species exists is enough to require consideration.

The mind of a meta-Intentional agent has as its object of experience the agent itself, immersed in and interacting with its subjective environment as it is, was, or could be. This flow is based upon a narrative infrastructure which includes "islands" of description and explanation of the meaning themselves. Such descriptions and explanations offer further opportunities for actions, or affordances.

Suppose you are arriving at a friend's party. As you enter, the host introduces a person to you; she smiles and pronounces her own name; you pronounce yours and shake hands with her. After a brief exchange, you go on to meet the other guests. Some of them you know, some you don't, so you probably get introduced to a few other persons. After a while you chance upon the person whom you met on your arrival. Under normal circumstances, you recognize her face and start the conversation from where it had been interrupted, maybe searching your memory for her name. The next day you meet your friend for lunch, and she is again in the company of that person. If memory does not fail you, this time you effortlessly remember her name and begin a friendly conversation, reminiscing about the main events of the party and letting the exchange go wherever the three of you let it go. If you and she become acquainted, you will end up recognizing her at a distance, from her way of walking and the general shape of her figure.

What has happened? When you first met this person, at the common friend's place, you paid attention to (at least) basically two things: her look—particularly, her face—and her

name. Like all human beings, you have a specific faculty of face perception and general person recognition (Paller et al., 2003; Peterson and Rhodes, 2003), which will of course take due notice of that person's lineaments and name. In more detail, when she is introduced to you you *experience* her look and her name; this experience becomes a *description* of that person, which in its turn *shapes your experience* of that person the next time(s) you meet her, allowing you to recognize her with increasing certainty. What we have here is a circular (better still, spiral) relation between "base-level" experience and "upper-level" metacognition.

Meta-Intentionality is not necessarily reasoning: it is just the intertwining and coevolution of experience and description that allows for the re-enactment of the cognitive performance (Guidano, 1987, 1991). Experience gives rise to descriptions in the form of narratives, explanations, maps, and so on; these become reincorporated into experience giving it new forms, structures and meanings.

Of course, this process may occasionally become more deliberate and ratiomorph, e.g., if the second time you meet that person at the party you have forgotten her name, you might actively conjure up a way to have her say her name again; or, if you find yourself interested in her, you might actively try to build an understanding of her ways of looking at the world, her interests, and so on. The point is not that these activities are not possible, but that they are not necessary. In time, the very infrastructure of your experience of that person will be shaped by the maps of her that you have built, so that under many, or most, circumstances, you will know, with no particular attentional or reasoning efforts, how you ought to behave with her, how she would react to something that you might say or do, what you can expect from her, and so on. Your descriptions will have melted into your experience, changing its shape and allowing for new, more complex maps to emerge.

Something very similar happens when we substitute the notion of explanation for that of description we have just used. A meta-Intentional mind, or at least that which characterizes the human species, is structured so to look for explanations of the events it perceives. Such explanations may be couched e.g., in folk psychology (e.g., Tirassa, 1999b; Tirassa and Bosco, 2008; Bosco et al., 2014; Brizio et al., 2015), naïve physics (e.g., Hayes, 1979; Smith and Casati, 1994; Spelke, 1994; Smith, 1995), and related ways of looking at the world. Here, again we have the same spiral relation between experience and explanation; of course, it would be reasonable to argue that explanation is indeed but a type of description.

Thus, when we sit in a car and turn the key we have a whole, complex set of expectations and experience. If the engine does not start our mind begins to formulate dynamics of possible explanations. This happens because we have had different kinds of experience with engines that start and do not start. Such experience, which begins partial and scattered, progressively becomes more unitary and, interwoven with fragments of descriptions and explanation, goes to shape the future experience, which will present itself already laden with partially ready-made cognitive structures that in their turn allow a more sophisticated set of possible action. When the engine does not start, an expert

driver will immediately look at the dashboard to check if it is turned on, she will try to remember when it was that she had the battery checked last time, if there is petrol in the tank, and so on. Each such affordance is made possible by the narrative structuring of experience which is provided by description and explanation.

Another crucial feature of human meta-Intentional agency is plan construction and use. A plan is a resource for action (Agre and Chapman, 1990), a description that we use to guide our management of the situation. To build a plan is to imagine an alternate situation, or several alternate situations, and to keep it present to our attention while we look for a way to realize it.

At first sight, plans are the meta-Intentional version of what simple actions are to a standard Intentional mind. However, there is something more to this point than when we said that action always includes a prediction of how the world will evolve spontaneously and as a result of action. The latter capability, however sophisticated, only requires a *forward* projection of the situation currently at sight (and, as far as memories are concerned, meshing this with a *backward* projection). Planning requires instead wholly alternative situations to be conjured up—it is, in the same metaphor, a *lateral* projection.

The example of long-term planning is particularly telling. When we plan where to spend our holidays next year, we use our knowledge of the present: we may assume, for example, that we will be able to count on a certain income, that we will still have the friends we have today, and so on. This is more the identification of certain pillars with which to build the alternate situations than a real prediction of how our life will evolve over the next several months. It is only after building such alternate situations that we begin planning, that is, imagining how *those* situations might evolve.

Thus, the current situation may contain, as if it were transparent, its future development; but it does not contain alternate situations, which have instead to be conjured up from scratch by projecting possible interferences with actual or potential dynamics. Such alternate situations will be dynamic in their turn, which makes the whole operation comparatively difficult. Planning thus requires more than just "basic" Intentional capabilities.

Planning is not something that occurs every now and then, or once in a while. Once a species has such capability, it will play a role in all of its mental activities. So, we are *always* making plans. Everything we do is only understandable as part of a plan: it is because we always live in several situations, only one of which is the "real" one, that we can decide, for example, to sit at a desk writing a paper on a sunny Sunday afternoon. The writer's current situation includes being a professional researcher, pursuing certain intellectual interests, having to earn salary so to be able to keep a certain way of life, and so on; therefore, writing a scientific article is just part of the affordances that she views in the situation. It would make no sense to keep these knowledge, desires, attitudes, and so on, separate from the writer's representation of the situation, or to conceive of them as add-on, "meta-layer" features plugged into an otherwise simpler system. They are just part of the writer's current flows of

meaning: to be meta-Intentional requires no additional effort to a meta-Intentional animal.

It follows from this description that a meta-Intentional agent can also try and modify what we have called pillars and see what would happen, like a writer would. Indeed, this is the starting point for fiction, pretend play, story-telling, and the ability to put oneself in somebody else's shoes.

Finally, meta-Intentional descriptions allow an agent to imagine how she would look from the outside. Our control system is, at least within certain limits, an observer of its own interactions. One of the result of such observation is a narrative description of how we would appear from an external standpoint. This capability plays an immediate role in our agency. Many features and dynamics of our social life depend on the internalization of such external descriptions: think for example, of our capability of obeying abstract social rules, of experiencing shame or remorse, of thinking we are overweight, or wondering whether we are sexually attractive.

Furthermore, since our control system uses these observations from the exterior as a feedback on the interaction that is going on under its supervision, with the same spiral dynamics we described above, it can be said to produce, in a sense, behavior, that is a third-person description of our activity. This, again, would not be a separate faculty, but an immediate feature of our Intentionality. Thus, if behavior is conceived of as a function of the observer, and not of the observed organism, then the only animals that really do *behave* are, paradoxically, humans. Which, of course, is what our commonsense knowledge has always taken for granted.

CONCLUSION

We are aware that this paper is structured in an unusual way, so let us try again to make our intentions clear. There is a general (albeit, of course, not unanimous) consensus in the cognitive sciences on the very nature of cognition and action. This consensus relies on the substantial irrelevance of consciousness and experience, based on the adoption of either the computational postulate or of various forms of eliminativism. We include neural reductionism in the latter.

Then there are islands and whole archipelagos of dissenting, heterodoxical, and truly heretical positions (for large-scale reviews see e.g., Osbeck, 2009; Manera and Tirassa, 2010; Tirassa and Vallana, 2010). However, while there are reasons to reject what we have called the consensus, the alternatives (still?) show a low tendency to merge into a unitary paradigm, or anyway to give rise to one.

This has been going on for several years now. It is not necessarily worrying: after all, most disputes in philosophy, in psychology, in economics, or in the social sciences appear to be as old as human culture is, and this does not detract from their interest or usefulness. In psychology, which is the most important of our cultural and scientific matrices, eliminativists cohabit more or less happily with phenomenologists, behaviorists with psychoanalysts and computational neuroscientists, and so on.

The downside of this situation, however, is a sense that no dispute can ever be settled, that the same arguments are reiterated over and over like ready-made tokens or gambling chips. What we tried to do with this paper is to look for a way out of this seeming stalemate by avoiding, as much as possible, the umpteenth discussion about the premises and instead trying to just develop the consequences of a certain set of premises. Instead of getting stuck into well-rehearsed arguments, we brought together a variety of literature, from psychology to biology, from philosophy to artificial intelligence and tried to see what would happen if we took certain ideas seriously and tried to develop some of their consequences.

Of course the attempt is only partially successful, to be optimistic; the limits and possible objections of this work are obvious, ranging from the heterogeneity of the sources used (and sometimes the ambiguity of our interpretations thereof) to the difficulty of imagining the empirical counterparts and consequences of our perspective. We might have chosen one single issue, as circumscribed as possible, and tried to develop it in relative isolation, but this would have brought us back to the starting point. Furthermore, we believe that the real interesting topics—the nature of the mind, the nature and relations of perception and action, and so on—cannot be decomposed without losing too much of their significance and import.

We had to start somewhere, and we definitely are not certain to have reached anywhere. Yet, we hope that the reader has found the attempt decently interesting and fruitful.

AUTHOR CONTRIBUTIONS

ACKNOWLEDGMENTS

REFERENCES

Agre, P. E., and Chapman, D. (1990). "What are plans for? robotics and autonomous systems," in *Designing Autonomous Agents: Theory and Practice from Biology to Engineering and Back,* Vol. 6, ed P. Maes (Cambridge, MA: MIT Press), 17–34.

Airenti, G., Bara, B. G., and Colombetti, M. (1993). Conversation and behavior games in the pragmatics of dialogue. *Cogn. Sci.* 17, 197–256. doi: 10.1207/s15516709cog 1702_2

Barwise, J., and Perry, J. (1983). *Situations and Attitudes.* Cambridge, MA: MIT Press.

Bateson, M., Desire, S., Gartside, S. E., and Wright, G. A. (2011). Agitated honeybees exhibit pessimistic cognitive biases. *Curr. Biol.* 21, 1070–1073. doi: 10.1016/j.cub.2011.05.017

Bosco, F. M., Gabbatore, I., and Tirassa, M. (2014). A broad assessment of theory of mind in adolescence: the complexity of mindreading. *Conscious. Cogn.* 24, 84–97. doi: 10.1016/j.concog.2014.01.003

Brentano, F. (1874). *Psychologie vom Empirischen Standpunkte [Psychology from an Empirical Standpoint]*. Transl. by A. C. Rancurello, D. B. Teller, and L. L. McAlister (1995) (London; New York, NY; Routledge) in English. Leipzig: Duncker & Humbolt.

Brizio, A., Gabbatore, I., Tirassa, M., and Bosco, F. M. (2015). "No more a child, not yet an adult": studying social cognition in adolescence. *Front. Psychol.* 6:1011. doi: 10.3389/fpsyg.2015.01011

Clark, A. (2001). *Mindware. An Introduction to the Philosophy of Cognitive Science*. New York, NY; Oxford: Oxford University Press.

Edelman, G. M. (1992). *Bright Air, Brilliant Fire: On the Matter of the Mind*. New York, NY: Basic Books.

Fikes, R. E., and Nilsson, N. J. (1971). STRIPS: a new approach to the application of theorem proving to problem solving. *Artif. Intell.* 2, 189–208. doi: 10.1016/0004-3702(71)90010-5

Flavell, J. H. (1979). Metacognition and cognitive monitoring: a new area of cognitive-developmental inquiry. *Am. Psychol.* 34, 906–911. doi: 10.1037/0003-066X.34.10.906

Gans, C., and Northcutt, R. G. (1983). Neural crest and the origin of vertebrates: a new head. *Science* 220, 268–273. doi: 10.1126/science.220.4594.268

Gibson, J. J. (1977). "The theory of affordances," in *Perceiving, Acting, and Knowing*, eds R. E. Shaw and J. Bransford (Hillsdale, NJ: Erlbaum), 127–143.

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Boston, MA: Houghton Mifflin.

Gibson, W. T., Gonzalez, C. R., Fernandez, C., Ramasamy, L., Tabachnik, T., Du, R. R., et al. (2015). Behavioral responses to a repetitive visual threat stimulus express a persistent state of defensive arousal in Drosophila. *Curr. Biol.* 25, 1401–1415. doi: 10.1016/j.cub.2015.03.058

Glenberg, A. M. (1997). What memory is for. *Behav. Brain Sci.* 20, 1–55. doi: 10.1017/s0140525x97000010

Gould, S. J. (2002). *The Structure of Evolutionary Theory*. Harvard, MA: Harvard University Press.

Gould, S. J., and Lewontin, R. C. (1979). The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. *Proc. R. Soc. Lond. B Biol. Sci.* 205, 581–598. doi: 10.1098/rspb.1979.0086

Guidano, V. F. (1987). *Complexity of the Self: A Developmental Approach to Psychopathology and Therapy*. New York, NY: Guilford.

Guidano, V. F. (1991). *The Self in Process. Toward a Post-Rationalist Cognitive Therapy*. New York, NY: Guilford.

Harnad, S. (1994). "Does the mind piggy-back on robotic and symbolic capacity?" in *The Mind, the Brain, and Complex Adaptive Systems*, eds H. Morowitz and J. Singer (Santa Fe: Santa Fe Institute/Addison Wesley), 204–220.

Hatschek, B. (1888). *Lehrbuch der Zoologie: Eine Morphologische Übersicht des Thierreiches zur Einführung in das Studium dieser Wissenschaft*. Jena: Fischer.

Haugeland, J. (ed.) (1981). *Mind Design*. Cambridge, MA: MIT Press.

Hayes, P. J. (1979). "The naive physics manifesto," in *Expert Systems in the Micro-Electronic Age*, ed D. Michie (Edinburgh: Edinburgh University Press), 242–270.

Hedwig, B. (2006). Pulses, patterns and paths: neurobiology of acoustic behavior in crickets. *J. Comp. Physiol. A. Neuroethol. Sens. Neural Behav. Physiol.* 192, 677–689. doi: 10.1007/s00359-006-0115-8

Hedwig, B., and Poulet, J. F. (2005). Mechanisms underlying phonotactic steering in the cricket *Gryllus bimaculatus* revealed with a fast trackball system. *J. Exp. Biol.* 208, 915–927. doi: 10.1242/jeb.01452

Heidegger, M. (1927). *Sein Und Zeit [Being and Time]*. Transl. by J. Macquarrie and E. Robinson (1962) (London: SCM) in English. Tübingen: Mohr.

Hennig, R. M. (2009). Walking in Fourier's space: algorithms for the computation of periodicities in song patterns by the cricket *Gryllus bimaculatus*. *J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol.* 195, 971–987. doi: 10.1007/s00359-009-0473-0

Jennings, H. S. (1906). *Behavior of the Lower Organisms*. New York, NY: Columbia University Press.

Johnson, M. (1987). *The Body in the Mind: The Bodily Basis of Imagination, Reason and Meaning*. Chicago, IL: University of Chicago Press.

Kant, I. (1781). *Critik Der Reinen Vernunft [Critique of Pure Reason]*. Transl. by P. Guyer and A. W. Wood (1998) (Cambridge: Cambridge University Press) in English. Riga: Johann Friedrich Hartknoch.

Kesner, R. P., and Olton, D. S. (eds.) (1990). *Neurobiology of Comparative Cognition*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Lewontin, R. (1998). *Gene, Organismo e Ambiente [Gene, Organism, and Environment]*. Transl. by The triple helix (2000) (Cambridge, MA; London: Harvard University Press) in English. Roma e Bari: Laterza.

Lindemann, J. P., and Egelhaaf, M. (2012). Texture dependence of motion sensing and free flight behavior in blowflies. *Front. Behav. Neurosci.* 6:92. doi: 10.3389/fnbeh.2012.00092

Lorenz, K. (1965). *Evolution and Modification of Behavior*. Chicago, IL; London: The University of Chicago Press.

Makkink, G. F. (1936). An attempt at an ethogram of the European Avocet (*Recurvirostra avasetta L.*), with ethological and psychological remarks. *Ardea* 25, 1–62.

Manera, V., and Tirassa, M. (2010). "Cognitive science," in *The Pragmatics Encyclopedia*, ed L. Cummings (London; New York, NY: Routledge), 55–58.

Maturana, H. D., and Varela, F. J. (1980). *Autopoiesis and Cognition. The Realization of the Living*. Dordrecht: Reidel.

Mayr, E. (1997). *This is Biology - The Science of the Living World*. Cambridge, MA: Belknap Press.

McDermott, D. (1987). A critique of pure reason. *Comput. Intell.* 3, 151–160.

Mendl, M., Paul, E. S., and Chittka, L. (2011). Animal behavior: emotion in invertebrates? *Curr. Biol.* 21, R463–R465. doi: 10.1016/j.cub.2011.05.028

Menzel, R., and Giurfa, M. (2001). Cognitive architecture of a mini-brain: the honeybee. *Trends Cogn. Sci.* 5, 62–71. doi: 10.1016/S1364-6613(00)01601-6

Menzel, R., Leboulle, G., and Eisenhardt, D. (2006). Small brains, bright minds. *Cell* 124, 237–239. doi: 10.1016/j.cell.2006.01.011

Merleau-Ponty, M. (1945). *Phénoménologie de la Perception [Phenomenology of Perception]*. Transl. by C. Smith (1981) (London: Routledge) in English. Paris: Gallimard.

Moroz, L. L., Kocot, K. M., Citarella, M. R., Dosung, S., Norekian, T. P., Povolotskaya, I. S., et al. (2014). The ctenophore genome and the evolutionary origins of neural systems. *Nature* 510, 109–114. doi: 10.1038/nature13400

Nagel, T. (1974). What is it like to be a bat? *Philos. Rev.* 83, 435–450. doi: 10.2307/2183914

Nagel, T. (1986). *The View from Nowhere*. Oxford: Oxford University Press.

Newell, A. (1990). *Unified Theories of Cognition*. Harvard, MA: Harvard University Press.

Newell, A., and Simon, H. A. (1972). *Human Problem Solving*. Englewood Cliffs, NJ: Prentice-Hall.

Newell, A., and Simon, H. A. (1976). Computer science as empirical enquiry: symbols and search. *Commun. Assoc. Comput. Mach.* 19, 113–126. doi: 10.1145/360018.360022

Osbeck, L. M. (2009). Transformations in cognitive science: implications and issues posed. *J. Theor. Philos. Psychol.* 29, 16–33. doi: 10.1037/a0015454

Paller, K. A., Ranganath, C., Gonsalves, B., LaBar, K. S., Parrish, T. B., Gitelman, D. R., et al. (2003). Neural correlates of person recognition. *Learn. Mem.* 10, 253–260. doi: 10.1101/lm.57403

Peterson, M. A., and Rhodes, G. (eds.) (2003). *Perception of Faces, Objects, and Scenes: Analytic and Holistic Processes*. New York, NY: Oxford University Press.

Pollock, J. L. (1993). The phylogeny of rationality. *Cogn. Sci.* 17, 563–588. doi: 10.1207/s15516709cog1704_4

Poulet, J. F., and Hedwig, B. (2005). Auditory orientation in crickets: pattern recognition controls reactive steering. *Proc. Natl. Acad. Sci. U.S.A.* 102, 15665–15669. doi: 10.1073/pnas.0505282102

Prato Previde, E., Colombetti, M., Poli, M., and Cenami Spada, E. (1992). The mind of organisms: some issues about animal cognition. *Int. J. Comp. Psychol.* 6, 79–119.

Pylyshyn, Z. W. (1984). *Computation and Cognition*. Cambridge, MA: MIT Press.

Rao, A., and Georgeff, M. (1992). "An abstract architecture for rational agents," in *Proceedings of the 3rd International Conference on Knowledge Representation*

and Reasoning, eds B. Nebel, C. Rich, and W. Swartout (San Mateo, CA: Morgan Kaufmann), 439–449.

Russell, S., and Norvig, P. (2009). *Artificial Intelligence: A Modern Approach, 3rd Edn.* Upper Saddle River, NJ: Pearson.

Searle, J. R. (1980). Minds, brains, and programs. *Behav. Brain Sci.* 3, 417–456. doi: 10.1017/S0140525X00005756

Searle, J. R. (1983). *Intentionality: An Essay in the Philosophy of Mind.* Cambridge: Cambridge University Press.

Searle, J. R. (1992). *The Rediscovery of the Mind.* Cambridge, MA: MIT Press.

Shallice, T. (1982). Specific impairments of planning. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 298, 199–209. doi: 10.1098/rstb.1982.0082

Smith, B. (1995). Formal ontology, common sense and cognitive science. *Int. J. Hum. Comput. Stud.* 43, 641–667. doi: 10.1006/ijhc.1995.1067

Smith, B., and Casati, R. (1994). Naive physics: an essay in ontology. *Philos. Psychol.* 7, 225–244. doi: 10.1080/09515089408573121

Spelke, E. (1994). Initial knowledge: six suggestions. *Cognition* 50, 431–445. doi: 10.1016/0010-0277(94)90039-6

Tirassa, M. (1994). Is consciousness necessary to high-level control systems? *Psycoloquy* 5, 82.

Tirassa, M. (1999a). Taking the trivial doctrine seriously: functionalism, eliminativism, and materialism. *Behav. Brain Sci.* 22, 851–852.

Tirassa, M. (1999b). Communicative competence and the architecture of the mind/brain. *Brain Lang.* 68, 419–441.

Tirassa, M., and Bosco, F. M. (2008). On the nature and role of intersubjectivity in human communication. *Emerg. Commun. Stud. New Technol. Pract. Commun.* 10, 81–95.

Tirassa, M., Carassa, A., and Geminiani, G. (2000). "A theoretical framework for the study of spatial cognition," in *Spatial Cognition. Foundations and Applications*, ed S. Ó. Nualláin (Amsterdam; Philadelphia: Benjamins), 19–31.

Tirassa, M., and Vallana, M. (2010). "Representation and computation," in *The Pragmatics Encyclopedia*, ed L. Cummings (London; New York, NY: Routledge), 399–402.

Turing, A. M. (1950). Computing machinery and intelligence. *Mind* 59, 433–460. doi: 10.1093/mind/LIX.236.433

Uexküll, J. (1934). *Streifzüge Durch die Umwelten von Tieren und Menschen. Ein Bilderbuch Unsichtbarer Welten [A Foray into the Worlds of Animals and Humans (with "A theory of meaning")].* Transl. by J. D. O'Neil (2010) (Minneapolis, MN; London: University of Minnesota PressBerlin) in English. Berlin: J. Springer.

Varela, F. (1996). "A science of consciousness as if experience mattered," in *Toward a Science of Consciousness: The first Tucson Discussions and Debates*, eds S. R. Hameroff, A. W. Kaszniak and A. C. Scott (Cambridge, MA: MIT Press), 31–44.

Varela, F. J., Thompson, E., and Rosch, E. (1991). *The Embodied Mind. Cognitive Science and Human Experience.* Cambridge, MA: MIT Press.

Wehner, R. (2003). Desert ant navigation: how miniature brains solve complex tasks. *J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol.* 189, 579–588. doi: 10.1007/s00359-003-0431-1

Is the Motor System Necessary for Processing Action and Abstract Emotion Words? Evidence from Focal Brain Lesions

Felix R. Dreyer[1]*, Dietmar Frey[2], Sophie Arana[3], Sarah von Saldern[1], Thomas Picht[2], Peter Vajkoczy[2] and Friedemann Pulvermüller[1,4]*

[1] Brain Language Laboratory, Department of Philosophy and Humanities, Freie Universität Berlin, Berlin, Germany, [2] Department of Neurosurgery, Charite University Medicine Berlin, Berlin, Germany, [3] Radboud Universiteit, Nijmegen, Netherlands, [4] Berlin School of Mind and Brain, Humboldt Universität zu Berlin, Berlin, Germany

Neuroimaging and neuropsychological experiments suggest that modality-preferential cortices, including motor- and somatosensory areas, contribute to the semantic processing of action related concrete words. Still, a possible role of sensorimotor areas in processing abstract meaning remains under debate. Recent fMRI studies indicate an involvement of the left sensorimotor cortex in the processing of abstract-emotional words (e.g., "love") which resembles activation patterns seen for action words. But are the activated areas indeed necessary for processing action-related and abstract words? The current study now investigates word processing in two patients suffering from focal brain lesion in the left frontocentral motor system. A speeded Lexical Decision Task on meticulously matched word groups showed that the recognition of nouns from different semantic categories – related to food, animals, tools, and abstract-emotional concepts – was differentially affected. Whereas patient HS with a lesion in dorsolateral central sensorimotor systems next to the hand area showed a category-specific deficit in recognizing tool words, patient CA suffering from lesion centered in the left supplementary motor area was primarily impaired in abstract-emotional word processing. These results point to a causal role of the motor cortex in the semantic processing of both action-related object concepts and abstract-emotional concepts and therefore suggest that the motor areas previously found active in action-related and abstract word processing can serve a meaning-specific necessary role in word recognition. The category-specific nature of the observed dissociations is difficult to reconcile with the idea that sensorimotor systems are somehow peripheral or 'epiphenomenal' to meaning and concept processing. Rather, our results are consistent with the claim that cognition is grounded in action and perception and based on distributed action perception circuits reaching into modality-preferential cortex.

Keywords: embodied cognition, category specific impairments, lesion studies, semantic processing

---

Abbreviations: AAT, Aachener Aphasia Test; ACC, Accuracy; fMRI, functional Magnetic Resonance Imaging; LDT, Lexical Decision Task; nTMS, navigated Transcranial Magnetic Stimulation; RSDT, revised standardized difference tests (see Crawford and Garthwaite, 2005 for details); RT, Reaction Time.

INTRODUCTION

A fundamental theoretical debate about the nature of meaning and concepts dominates the cognitive and brain sciences. Classic cognitive psychologists propose that semantic and conceptual processes are carried by a dedicated symbolic semantic system functionally detached from sensory and motor modules and specialized for handling information about meaning and concepts related to signs (e.g., Ellis and Young, 1988). An alternative approach, sometimes referred to by the terms 'embodiment' and 'semantic grounding', states that meaning is intrinsically related to (or grounded in) action and perception information and processed in the brain by distributed action perception circuits that reach into motor and sensory brain areas (Barsalou, 1999, 2008; Pulvermüller, 2005; Glenberg and Gallese, 2012). Some recent attempts to amalgamate both positions into one integrative proposal either maintain that semantic processing is processed by an amodal system, whereas modality-preferential cortices, such as the sensorimotor areas, play an optional, merely "coloring" role (Mahon and Caramazza, 2008; Caramazza et al., 2014), or they postulate semantic integration in a 'semantic hub' (typically placed in temporal cortex) and allow for additional modality-specific semantic centers across the cortex (Patterson et al., 2007; for review, see Binder and Desai, 2011; Kiefer and Pulvermüller, 2012). However, similar to the symbolic systems position, these proposals attribute true semantic processing and related deficits primarily to semantic hub areas. To cite but one relevant statement here: "understanding the word "run" *occurs in* modality-independent neural systems" (Bedny and Caramazza, 2011, p. 92; our own emphasis). Therefore, it is not clear whether this type of 'integrative' position allows for the explanation of category-specific deficits arising from a focal lesion in one modality-preferential cortical system.

Much recent imaging work has accumulated evidence that the motor cortex (Damasio et al., 1996; Martin et al., 1996; Hauk and Pulvermüller, 2004; Hauk et al., 2004; Pulvermüller et al., 2006) and a range of sensory systems (González et al., 2006; Kiefer et al., 2008; Barrós-Loscertales et al., 2012) become active when words and concepts from different semantic categories are being processed. In particular, the motor system instantaneously activates in a somatotopic fashion when subjects hear or read words semantically related to different parts of the body (Pulvermüller et al., 2005b; Shtyrov et al., 2014), thus arguing against the view that the 'grounded' sensorimotor activations may only emerge at a late stage of *post hoc* interpretation and supporting their genuine role in semantic information access. Category-specific semantic activation across the motor system has originally been reported for action-related verbs, but has recently been replicated for nouns semantically related to the mouth and hand (food and tool nouns; Carota et al., 2012). Although some researchers argue some of these effects are difficult to reproduce (Postle et al., 2008; Caramazza et al., 2014), systematic comparison of studies across labs demonstrated good reproducibility (Carota et al., 2012; Kemmerer et al., 2012). Semantically related activation in the motor system has even been reported for abstract words related to emotions (Moseley et al., 2012).

However, although these brain activation studies are consistent with, and confirm predictions of, the grounded-semantic account postulating the relevance of modality preferential areas for semantics, neuroimaging and neurophysiological studies can never proove the functional relevance and necessity of brain areas for cognitive function. To investigate this crucial issue, lesion studies in neurological patients and neurostimulation approaches are necessary.

Here, a range of results has so far been suggestive of a role of sensorimotor systems in semantic processing. For example, Pulvermüller et al. (2005a) applied single TMS pulses to primary hand and foot motor cortex while verbs semantically related to hand or foot actions had to be recognized in a Lexical Decision Task (LDT). As the recognition latencies for hand- and leg-related action verbs was differentially affected by TMS stimulation site (an effect confirmed by a significant interaction of these factors), a causal role of the motor cortex on semantic word type processing was evident. The latter conclusion was also supported by further TMS work in healthy subjects (Willems et al., 2011) and by behavioral experiments in which subjects engaged in motor activity while linguistic-semantic information had to be processed (Glenberg et al., 2008; Witt et al., 2010; Shebani and Pulvermüller, 2013). However, as most of the causal effects of motor activity on semantic processing were manifest in RTs but not accuracies (ACCs), it may still be that the functional role of motor systems for category-specific semantic processing is only relevant for optimizing word processing, but not necessary for it.

Stronger claims about the necessity of modality preferential, including sensorimotor, cortex for semantics can potentially be derived from lesions studies. Important and well-known classic work reported lesion-related category-specific semantic impairments for words related to manipulable objects (Warrington and McCarthy, 1983, 1987), animals and foods (Warrington and Shallice, 1984; for a recent review see also Gainotti, 2010), which were manifest in task ACCs. On closer inspection, the observed patterns of impairments confirm that lesions of regions that include motor areas can lead to selective and pronounced deficits in the processing of action verbs (Damasio and Tranel, 1993; Bak et al., 2001; Neininger and Pulvermüller, 2001, 2003; Arevalo et al., 2012; Kemmerer et al., 2012). Similar results, supportive for theories of embodied cognition, were found after lesions of auditory (Bonner and Grossman, 2012; Trumpp et al., 2013), and visual systems (Gainotti, 2010; Pulvermüller et al., 2010) for the processing of words with auditory or primarily visual semantics. Whereas lesions in modality-preferential sensorimotor cortex brings about deficits in processing action-related words, the granularity of the category-specific deficit is under discussion (see Arevalo et al., 2012; Reilly et al., 2014). At present no evidence exists for a differential involvement of hand- or face-related action words, which can be found amongst verbs ('write' vs. 'chew') but also amongst the nouns (hand-related tool vs. mouth-related food words) (Arevalo et al., 2012).

Unfortunately, several limitations apply to the majority of previous patient studies. First, the patient populations under investigation typically suffered from large lesions typically caused

by stroke or degenerative brain disease. Most of these lesions included motor or sensory cortex but, in addition, other parts of the brain, as in strokes, or even were of diffuse nature, as in motor neuron disease and semantic dementia. Therefore fine grained conclusions about the functional role of specific brain areas in word processing are difficult to derive and it is not entirely clear whether the sensorimotor lesion was indeed the primary cause of the patterns of deficits reported. Second, from a psycholinguistic perspective, the choice of stimulus materials allowed different interpretations of the results. For example, the popular comparison between action-related verbs and object-related nouns frequently led to evidence of a category-deficit, but it is not always clear whether such a deficit is best explained by semantic factors (action- and object-relatedness) or in terms of the lexical (or grammatical) category difference instead (nouns vs. verbs). In addition, relevant psycholinguistic variables such as word length and word frequency were not always matched in previous studies, thus opening further options for alternative explanation of presumed 'category differences'. However, a small number of recent studies looking at rather focal lesions suggest that auditory and action-recognition systems may also be necessary for processing the semantically related words (Neininger and Pulvermüller, 2001; Campanella et al., 2010; Trumpp et al., 2013).

Although some evidence for a causal and possibly even necessary role of modality preferential cortex for category-specific semantic processing exists, no similar data are available for abstract words whose semantic information is somewhat detached from specific sensory and motor modalities. A major claim held by most symbolic systems accounts, and equally the integrative proposals mentioned above, is that abstract semantic processing is removed from, and does not require, sensory and motor systems of the brain. In contrast, proponents of grounded cognition have argued that, in order to learn the meaning of an abstract word, it is necessary to know at least some concrete semantic instantiations and contexts in which it can be used (Barsalou and Wiemer-Hastings, 2005; Borghi and Cimatti, 2009; Pulvermüller, 2013). At the neuromechanistic level, it has therefore been proposed that abstract meanings, similar to concrete ones, are organized as distributed neuronal circuits including neurons in multimodal and sensorimotor systems, although their links into modality-preferential areas may be weaker than those of concrete conceptual representations. This idea is supported by behavioral (Glenberg et al., 2008) and fMRI findings (Moseley et al., 2012), both indicating an involvement of motor processing in comprehension of abstract words. A strong version of a semantic grounding position thus implies that modality preferential sensorimotor cortex also takes a crucial role in abstract word processing (Glenberg et al., 2008; Havas et al., 2010), but to our knowledge this has so far not been shown with neither neurostimulation, nor lesion approaches. If correct, this position predicts that lesions in modality preferential cortex, and in motor areas specifically, can lead to category-specific semantic deficits in processing abstract words. Positive evidence for this statement would certainly falsify symbolical semantic accounts and most integrative proposals still leaning toward

abstract-symbolism, too (Mahon and Caramazza, 2008; Dove, 2009).

On the background of this pre-existing work, the current study addresses the putative necessary role of the modality preferential sensorimotor cortex in the processing of both, action-related and abstract words by examining two patients with focal brain lesions. Although group studies were once claimed necessary for drawing strong conclusions on the brain basis of cognition and language, we would argue that single case reports are indeed suited perfectly well to provide existence proofs for the claimed causality, as they are relevant for the current debate. In addition, some researchers have highlighted the advantages of single case studies, especially if brain localizations of function can strongly differ between individuals – as it is known to be the case for sensorimotor functions (Elbert et al., 1995; Buonomano and Merzenich, 1998) – and the grouping of patients with necessarily non-identical lesions is always debatable (Caramazza, 1986). However, we hasten to add that, whereas case studies can confirm claims about existence ('there is one case for which it applies that...'), they clearly cannot found general ('all') statements.

To overcome the mutual confounding of word semantics and grammatical class, as present in previous patient studies, the current study probed both nouns and verbs separately. This opens the possibility of finding category-semantic deficits that are, in addition, specific to lexical class. With the inclusion of abstract word categories, it becomes possible to investigate whether semantic grounding in modality-preferential cortex applies exclusively to concrete words, or extend to the domain of abstract semantics. To allow conclusions about semantic processing rather than to other stimulus features, semantic categories were matched for a range of psycholinguistic features (see Methods). Word recognition was monitored using a speeded LDT. Performance on this task has previously been shown to be sensitive to aspects of word semantics (Chumbley and Balota, 1984; see also Neininger and Pulvermüller, 2001, 2003). Furthermore, the LDT has important advantages over other tests frequently used in previous studies of semantic category specificity, including, for example, picture naming or categorical classification. These latter tasks require a similar semantic relationship between words and pictures (which, however, differs between concrete and abstract items) and similar perceptual-semantic similarity structure (which differs, for example, between animals and tools), the absence of which limits the scope of their use. In contrast, the LDT offers a straightforward possibility to test performance across word categories differing in their (e.g., abstract vs. concrete) semantics; it has therefore been applied frequently in previous research targeting effects of concreteness and semantic category specificity (e.g., James, 1975; Kroll and Merves, 1986; Jin, 1990; Neininger and Pulvermüller, 2001, 2003; Samson and Pillon, 2004). The rational underlying this research strategy is the following: If semantic processes elicited by one semantic type of words are specifically supported by a given area and if this area is lesioned, the recognition process of the respective word category can be impaired (delayed and/or less accurate). And if a deficit specific for a specific semantic category results from a focal lesion, the lesioned area is a likely key site for processing the affected semantic type. The theoretical

background for this prediction is the theory of distributed semantic circuits, according to which neuronal networks with different cortical distributions underlie the processing of different semantic word types (see Pulvermüller, 2013). A focal lesion in an area belonging to the distributional pattern of one semantic word type, but not other word types, would lead to a reduction of the excitatory feedback in the respective category-specific semantic circuits and therefore to delayed and more errorful word recognition. By testing two patients suffering from focal lesions in their frontocentral sensorimotor cortices, this study aims at fine grained conclusions on the functional involvement and necessity of the focal brain areas for the recognition of words from specific semantic categories. Adding abstract words to the stimulus material thereby allows to test whether such a crucial role of these modal areas just applies to the processing of words related to concrete concepts or even extends to abstract words.

MATERIALS AND METHODS

Patients and Clinical Examination

Patient HS

Patient HS was a 41 years old man, with a singular focal precentral lesion, situated directly inferior to the left hand motor cortex. HS was a native, monolingual German speaker and right handed (LQ = 80), with a total of 18 years of formal education and was serving in the military at the time of testing. Following biopsy, HS' lesion was diagnosed to be the single residual core of an Acute Disseminated Encephalomyelitis (ADEM) of 18 mm in diameter. Fiber tracking on Diffusion-Tensor-Imaging (DTI) data, using hotspots of an nTMS guided motor mapping procedure as seed regions (see Frey et al., 2012 for details) revealed his lesion to be situated in the precentral gyrus, half a centimeter away from the pyramidal tract of the hand motor cortex. A T1 weighted MRI scan of this lesion is shown in **Figure 1A**. At the time of language testing, neurological examination revealed mild paresis of the right arm and leg (grade 4, i.e., movement against external resistance, but less than normal), but no other cognitive or language impairment.

Patient CA

Patient CA was a 52 years old woman with a single lung cancer metastasis (histology: adenocarcinoma as non-small-cell lung carcinoma) in the superior frontal gyrus, affecting the supplementary motor area (SMA) of the left hemisphere, as shown in **Figure 1B**. The patient had been under chemotherapy for three cycles, underwent radiation therapy for two cycles and a first extirpation of the tumor had been performed 6 month prior to testing. All therapeutic measures did not result in control of the solitary cerebral metastasis. Due to growth of the tumor (with an extent of roughly 1.2 cm × 0.9 cm × 2 cm), indication for additional surgical removal was yielded at the time of testing. History revealed hypertension, chronic obstructive pulmonary disease and the regular administration of Pregabalin and Amitriptylin. CA was right handed (LQ = 80 at the time of testing) and a native, monolingual German speaker with 12 years of formal education and had been working as a chef pre-morbidly. CA did not report any sensory, motor, cognitive or language deficits and neurological examination did not reveal any impairments on those dimensions at the time of testing.

Control Participants

A group of 21 participants (five males) without neurological records served as control sample for the LDT paradigm. On average, controls were 40.7 years (SD = 18.7 years) old at the time of testing, with an age range from 18 to 79 years, covering that of the two neurological patients. Likewise, years of formal education were similar to CA and HS, spanning between 11 and 24 years, with an average of 16.5 years (SD = 3.5 years).
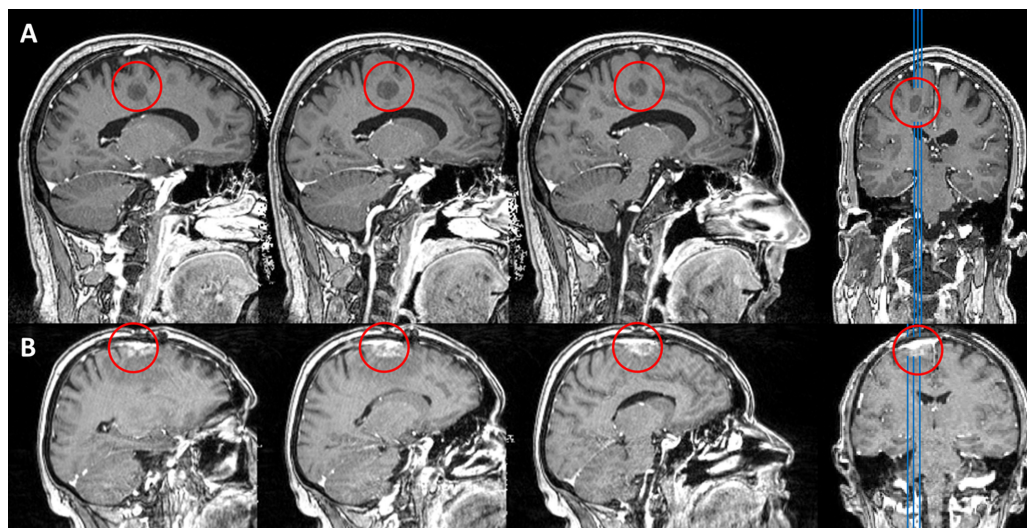

FIGURE 1 | T1 weighted MRI MPRAGE sequences of patients HS (A) and CA (B). Lesion sites are marked with a red circle.

TABLE 1 | Matching on psycholinguistic variables between semantic classes in nouns.

| | Nouns | | | | | | | | |
| | Abstract-emotional | | Animals | | Foods | | Tools | | |
Variables	M	SD	M	SD	M	SD	M	SD	p
Lemma Frequency p. Mio.	8.41	5.56	7.26	5.47	5.95	7.74	6.86	5.97	0.37
Length	5.7	1.4	5.5	1.66	5.78	1.37	5.93	1.47	0.64
Number of Syllables	1.7	0.46	1.7	0.46	1.78	0.42	1.88	0.33	0.21
Character bigram frequency p.Mio.	216786	101041	243304	123503	210825	120275	250937	145228	0.39
Character trigram frequency p.Mio.	120397	67480	148481	68331	124296	78409	125515	88827	0.35
Initial character frequency p.Mio.	12171	5742	13974	5816	14427	6163	14992	7248	0.21
Initial bigram frequency p.Mio.	2346	1926	2349	1901	1956	2000	2599	2321	0.57
Initial trigram frequency p.Mio.	414	926	748	1703	473	1262	913	1882	0.4
Coltheart neighbors frequency p.Mio.	126	488	82	269	28	74	56	117	0.47
Coltheart's N	6.08	6.07	7	6.65	6.01	5.77	7.16	5.7	0.76
Levenshtein neighbors frequency p.Mio.	256.35	1272.33	165.29	547.99	147.26	594.08	61.24	118.91	0.72
Levenshtein N	8.63	7.47	9.9	8.26	8.79	7.32	10.36	6.72	0.67

P-values denote results from one-way ANOVAs on the effect of semantic category.

Both, patients and healthy control participants, provided written informed consent prior to participating in the study and procedures were approved by the ethics committee of the Charité University Hospital, Berlin, Germany.

Paradigm

As critical test, a speeded LDT was carried out, as explained below. To assess clinical language proficiency the Token Test, and the repetition, naming and language comprehension subtests of the "Aachener Aphasie Test", or AAT, a standardized German aphasia test battery (Huber et al., 1983), were applied. Handedness was tested using the Edinburgh Inventory (Oldfield, 1971).

LDT Stimuli

Hundred sixty nouns and 160 verbs were presented, along with 320 matched pseudo-words. Each of the lexical/grammatical categories included 40 stimuli from 4 semantic groups or categories. Among the nouns, there were words used to speak about tools, food items, animals, and abstract-emotional entities. The semantic category groups of the verbs included words typically used to speak about actions performed with parts of the face (e.g., "kauen", *to chew*), hand (e.g., "greifen", *to grap*), or leg (e.g., "rennen", *to run*) and about abstract concepts ("hassen", to *hate;* see Supplementary Data for a complete overview of word stimuli).

Within each lexical category or grammatical word class, all semantic category groups were matched for a range of lexical and sub-lexical psycholinguistic variables, as determined by the dlex corpus (Heister et al., 2011). Matching was achieved for word length, number of syllables, phonological stress, normalized lemma frequency, character bigram frequency, character trigram frequency, initial character-, initial character bigram-, and initial character trigram frequency as well as for number of orthographic neighbors in terms of Coltheart's and Levenshtein's N. *F/t*-tests did not reveal differences between semantic category

groups for any of these psycholinguistic variables (all $p > 0.05$, see **Tables 1** and **2** for details).

In addition, an equal number of pronounceable pseudo-words was generated on the basis of the proper words using the 'Wuggy' software (Keuleers and Brysbaert, 2010). These pseudo-words were chosen to be not homophonous to proper words and to match all proper word categories, both combined and individually, in their sub-lexical psycholinguistic properties of average word length, number of syllables, character bigram frequency, character trigram frequency, initial character frequency, and initial bigram frequency (all $p > 0.05$, see **Table 3** for details). To further mimic appearance of proper words, pseudo-nouns all started with a capital letter and pseudo-verbs all ended in the "-en" suffix, consistent with German noun and verb orthography and morphology.

To empirically evaluate the semantic properties of the word stimuli, semantic ratings were collected from 20 healthy participants (monolingual native speakers of German aged 18–28) before the main experiment. Similar to previous studies (Pulvermüller et al., 2001; Hauk and Pulvermüller, 2004), semantic ratings were expressed on a Likert scales ranging from 1 (no relation) to 7 (strong relation). Each word was rated for its semantic relatedness to hand/arm-, face/mouth-, leg/foot actions, to visual, olfactory, gustatory, and haptic/tactile perceptions, as well as to emotions and mental processes. Ratings of concreteness and word familiarity were also obtained. The concreteness scale was thereby designed with the poles of high abstractness (1) to high concreteness (7). For inclusion into an effector-specific action word category (action verbs and tool/food nouns), words had to achieve an average rating above the neutral mid-point of four for the related question while being rated lower on all other action semantic scales. For animal nouns and abstract words, all action ratings were <4, with abstract items also rating <4 on concreteness and perceptual scales, but >4 on the scale for relation to mental processes. In addition, all abstract-emotional nouns, and also the majority of abstract verbs had

TABLE 2 | Matching on psycholinguistic variables between semantic classes in verbs.

	Verbs								
	Abstract		Face		Leg		Arm		
Variables	M	SD	M	SD	M	SD	M	SD	p
Lemma frequency p. Mio.	25.95	26.71	28.32	57.01	28.97	67.23	26.26	35.18	0.99
Length	6.68	1.29	6.98	1.33	7.08	1.35	6.83	1.24	0.54
Number of syllables	2	0	2	0	2	0	2	0	1
Character bigram frequency p.Mio.	544785	123466	536853	104691	534851	151015	530002	123341	0.96
Character trigram frequency p.Mio.	367231	73446	366884	58928	359082	65621	369878	55976	0.89
Initial character frequency p.Mio.	30979	27868	25033	18439	28458	18747	28579	25280	0.71
Initial bigram frequency p.Mio.	4375	8943	3245	3407	3098	3690	3149	2864	0.67
Initial trigram frequency p.Mio.	1368	2173	1155	2200	1012	1880	1320	2280	0.87
Coltheart neighbors frequency p.Mio.	145	227	61	135	104	347	87	262	0.51
Coltheart's N	5.36	4.08	5.25	4.32	4.38	3.40	4.94	3.5	0.66
Levenshtein neighbors frequency p.Mio.	164.41	233.91	103.64	189.97	114.06	363.51	92.55	262.01	0.65
Levenshtein N	8.47	5.92	7.73	5.14	6.51	4.64	7.33	4.18	0.37

P-values denote results from one-way ANOVAs on the effect of semantic category.

TABLE 3 | Matching on psycholinguistic variables between real and pseudo-words.

	Real words		Pseudo-words		
	M	SD	M	SD	p
Character bigram p.Mio. (Sum)	383542	197428	378278	189726	0.73
Character trigram p.Mio. (Sum)	247720	137424	245700	135682	0.85
Initial character p.Mio	21076	18206	22311	21276	0.43
Initial bigram p.Mio	2890	4051	2476	3274	0.16
Mean bigram p.Mio	51299	22628	51973	23000	0.71
Mean trigram p.Mio.	29493	15449	29945	15739	0.71
Length	6.31	1.5	6.12	1.3	0.08

P values denote results of t-test between both stimulus types.

strong emotional connotations with values >4 on the respective semantic scale. Semantic ratings for all categories are shown in **Figure 2**.

LDT Procedures

Participants were seated approximately 70 cm in front of a computer screen and were instructed to decide whether or not a word flashing on screen resembles a meaningful German word, or a pseudo-word instead. Responses were given via left hand mouse clicks, to assure that responses were not affected by possible motor impairments caused by left hemispheric lesions. Each trial started with a presentation of a central fixation cross. Its presentation time was pseudo-randomly varied between 2250 and 2750 ms (2500 ms on average) and it was followed by an acoustic beep signal of 200 ms length. 800 ms after the offset of this acoustic signal, the fixation cross disappeared and a word was presented tachistoscopically in the center of the screen for

130 ms. After word offset, the screen remained blank until a response was given, or for a maximum of 3000 ms after which the central fixation cross re-appeared. All stimuli were printed in black letters on a light gray background, using monospaced Courier New font with a font size of 13.5 and were spanning a maximum of 2° horizontal and 0.6° vertical visual degree.

Each test session started with 10 practice trials for the LDT, which applied stimuli that were not used in the actual experiment. Those trials were repeated until a task ACC of 80% was achieved, to assure that participants were sufficiently familiarized with task procedures.

The LD experiment was split up into 8 blocks, each including 80 letter strings, five words from each of the eight lexico-semantic categories as well as 20 pseudo-nouns and 20 pseudo-verbs. In addition, two words were presented as filler items at the beginning of each block, which were excluded from analysis. Each block lasted between 6 and 8 min, depending on participants' response speed. Between experimental blocks, participants were offered breaks.

Following the LDT testing, patients conducted the AAT subtests in the following order: Token Test, Verbal Repetition, Naming and Comprehension. To save time, subjects who performed <7 corrected error points on the Token Test (no aphasia diagnosis) were only given the most difficult part of the other subtests and if their performance was flawless, the rest of the subtest was omitted. On average, the whole aphasia test battery could be conducted within 20 min. Each test session was thereafter concluded by the Edinburgh Handedness Inventory and the basic demographics questionnaire.

Data Analysis
Healthy Control Participants

Lexical Decision Task analyses were conducted separately for noun and verb categories. Note again that all noun categories were matched with each other with regard to psycholinguistic variables, and the same applied for verb
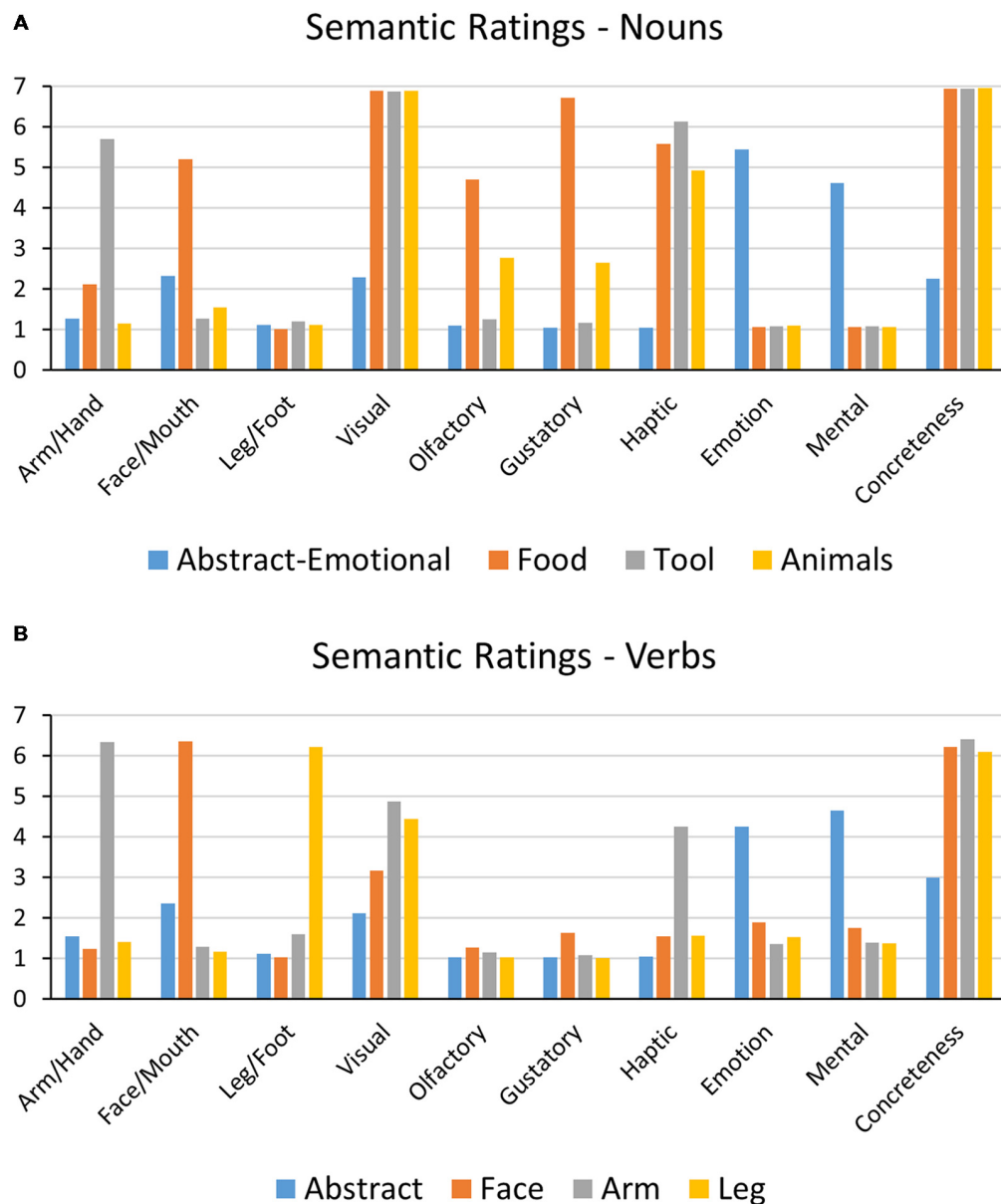
FIGURE 2 | Average semantic ratings for noun (A) and verb (B) categories, given on a scale from 1 (no semantic relation) to 7 (very strong semantic relation).

categories, but it was not possible to match across lexical (grammatical) categories. To allow response bias corrected comparisons with patients, task ACCs for individual lexico-semantic categories were converted into d' scores. To calculate d' values for each lexico-semantic group of nouns (verbs), each category's hit rate and the overall false positive rate of the entire lexical (i.e., either pseudo-noun or pseudo-verb) category was used (see also Pulvermüller et al., 2010). Resultant d' scores were compared between semantic categories using by-subject repeated measures analyses of variance (ANOVAs), by-item ANOVAS, and t-tests with Bonferroni correction for *post hoc* comparisons. Further testing was done to

compare the entire noun and verb groups against each other.

Reaction Times for correct responses were corrected for individual outliers > 2 standard deviations away from the mean. After correction, average RTs for each lexico-semantic category and individual participant were calculated and performance between semantic categories was compared separately for nouns and verbs. By-subject and by-item repeated measures ANOVAs were then used for overall analyses and t-tests for planned comparison testing. An additional analysis step compared the performance between nouns and verbs with repeated measures ANOVAs on d' and RT results.

Patients

Raw AAT scores were calculated, converted into normalized scores and compared to control samples according to the tests' instructions.

Lexical Decision Task ACCs for individual lexico-semantic categories were converted into d' scores as described above, for each patient individually. We tested for general performance differences between semantic groups within each lexical/grammatical category. To this end, ACC (here expressed as number of hits and misses) was compared using χ^2- and, in case of insufficient cell sizes ($n < 5$) Fischer's Exact Tests. In case those tests indicated significant differences, χ^2 tests with Bonferroni correction were conducted once for each semantic category versus the combined other categories within one grammatical word class (four comparisons). In case of significantly different semantic noun categories, a second set of analyses compared each action or abstract category against the reference category of non-action animal nouns (three comparisons). For completeness, all categories were finally pairwise compared against each other (six comparisons). Note again that analyses were done separately for nouns and verbs.

In the analysis of RT of correct responses, individual outliers >2 standard deviations away from the mean were first removed and the corrected single trial RTs were analyzed for effects of semantic word category using by-item ANOVAs and t-tests with Bonferroni correction.

To test whether differences across semantic categories in a specific patient can indeed be considered to be abnormal compared with performance differences between categories seen in the control sample revised standardized difference tests (RSDT; Crawford and Garthwaite, 2005) were conducted as *post hoc* tests. The RSDT resembles a derivate of the t-test, specifically designed to relate performance differences of individual patients directly to results of a group of control participants. To account for the inflated Type II error rate of the RSDT, these additional *post hoc* tests were one-tailed (see Crawford and Garthwaite, 2006 for discussion).

Furthermore, to test for effects of grammatical class, ACC performance was compared between all nouns and verbs using the χ^2 test and with ANOVAs on corrected RTs, respectively, in a separate analysis.

TABLE 4 | Performance of patients in AAT subtests given in *T*-Scores.

	T-Scores	
	CA	**HS**
Token test*	73	73
Verbal repetition*	74	74
Picture naming	80	80
Language comprehension	78	73
Average	76	75

Tests marked with an *were conducted in an abbreviated version.

RESULTS

AAT

No patient exhibited aphasic language impairments, as the AAT scores fell well within the range of healthy control population performance. Individual results for each patient and subtest are listed in **Table 4**.

LDT

Healthy Control Subjects

A repeated measures ANOVA on d' scores did not reveal any significant differences between semantic noun categories [$F(3,60) = 1.59$, $p > 0.1$, $\eta^2 = 0.08$, n.s.] or across verbs subtypes [$F(3,60) = 1.56$, $p < 0.1$, $\eta^2 = 0.08$, n.s.]. However, RTs differed significantly across both semantic categories of nouns [$F(3,60) = 21.4$, $p < 0.001$, $\eta^2 = 0.52$] and verbs [$F(3,60) = 8.3$, $p < 0.001$, $\eta^2 = 0.29$]. This pattern of results was confirmed with additional item-wise ANOVAs on d' [Nouns: $F(3,159) = 0.32$, $p = 0.81$, $\eta^2 = 0.01$; Verbs: $F(3,159) = 0.76$, $p = 0.52$, $\eta^2 = 0.01$] and RT [Nouns: $F(3,159) = 7.95$, $p < 0.001$, $\eta^2 = 0.13$; Verbs: $F(3,159) = 3.76$, $p = 0.01$, $\eta^2 = 0.07$]. For nouns, Bonferroni corrected *post hoc* t-tests revealed that RTs for abstract-emotional ($M = 693$ ms, $SE = 15$ ms) and tool words ($M = 689$ ms, $SE = 15$ ms) were significantly longer than for food ($M = 653$ ms, $SE = 14$ ms) and animal words ($M = 662$ ms, $SE = 16$ ms, all $t(20) > = 5$, $p < 0.001$, Cohen's $d > 1$). *Post hoc* tests conducted on verbs showed RTs for hand verbs ($M = 683$ ms, $SE = 19$ ms) to be significantly shorter than for abstract [$M = 710$ ms, $SE = 17$ ms, $t(20) = 3.5$, $p = 0.01$, Cohen's $d = 0.77$] and face- [$M = 704$ ms, $SE = 18$ ms, $t(20) = 3.3$, $p = 0.02$, Cohen's $d = 0.72$] as well as leg-related action verbs [$M = 720$ ms, $SE = 18$ ms, $t(20) = 3.9$, $p < 0.01$, Cohen's $d = 0.86$]. Overall, ACCs in terms of d' values for nouns and verbs were both high, with a significant advantage of nouns ($M = 3.9$, $SE = 0.09$) over verbs [$M = 3.6$ $SE = 0.13$, $t(20) = 3$, $p < 0.01$, Cohen's $d = 0.65$]. RTs results showed a similar pattern, with RTs for nouns ($M = 674$ ms, $SE = 14$ ms) being significantly shorter than for verbs [$M = 704$ ms, $SE = 18$ ms, $t(20) = 6.4$, $p < 0.001$, Cohen's $d = 1.17$].

Patient HS

Analysis of ACC revealed a significant difference in task performance between the noun categories ($\chi^2 = 10.45$, $df = 3$, $p = 0.01$, Cramer's $V = 0.26$) with performance on tool nouns (ACC = 0.83) being more impaired than that the other three categories combined (ACC = 0.97 on average, $\chi^2 = 9.4$, $df = 1$, $p = 0.02$, Cramer's $V = 0.24$). When comparing tool nouns against the reference category of non-action related animal nouns, a significant difference emerged [$\chi^2 = 7.67$, $df = 1$, $p = 0.036$, Cramer's $V = 0.31$] For verbs no significant differences in ACC between semantic categories was observed ($\chi^2 = 2.11$, $df = 3$ $p = 0.64$, Cramer's $V = 0.14$). The ANOVA on RTs did not show significant differences between categories for either nouns [$F(3,136) = 0.62$, $p > 0.1$, $\eta^2 = 0.01$, n.s.] or verbs [$F(3,137) = 0.62$, $p > 0.1$, $\eta^2 = 0.01$, n.s.]. As the RT distribution for verbs hinted a positive skew (Skewness = 0.84, $SE = 0.2$), the corresponding ANOVA was repeated on log-transformed
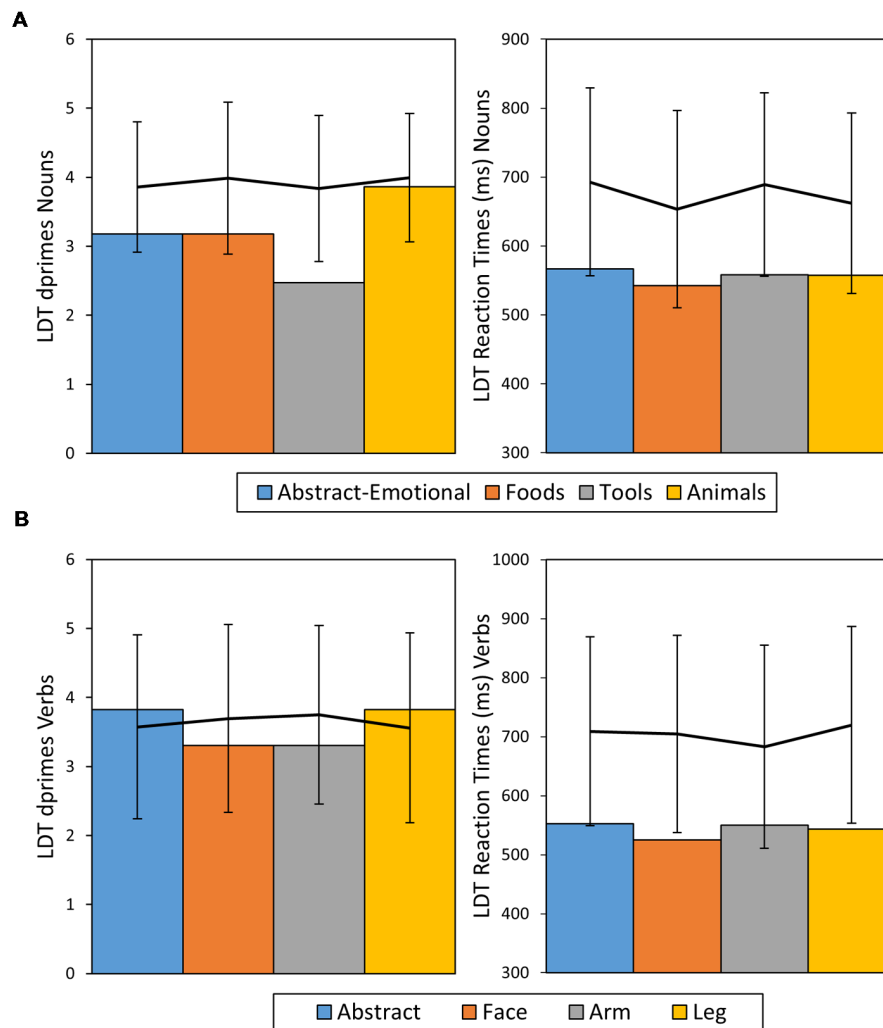
FIGURE 3 | Dprime and RT results of patient HS for nouns (A) and verbs (B) given in bar charts. The line diagrams represent average performance of the control sample with error bars representing ± 2 SDs.

data, but again did not hint significant differences between verbs categories [$F(3,137) = 0.59$, $p = 0.62$, $\eta^2 = 0.01$]. This pattern of results was replicated when comparing HS' performance with that of healthy controls. HS performance fell within the healthy range (mean ± 2 SDs) in terms of ACC and his RTs on verbs even tended to be faster than those of healthy control subjects. In contrast, for tool words, HS's ACCs were more than 2 SDs away from the mean of the control group, thus indicating significant slowing and therefore further confirming a selective impairment for tool nouns. *Post hoc* RSDT results confirmed this observation, as the difference in ACCs between animal and tool nouns was significantly more severe in patient HS than in the control sample [$t(20) = -2.72$, $p = 0.008$, Z-DCC $= -2.77$].

In direct comparison of noun and verb performance, HS exhibited no processing advantage for either word class in terms of ACC [ACC Nouns $= 0.93$, ACC Verbs $= 0.95$, $\chi^2 = 0.5$, $df = 1$, $p = 0.48$, Cramer's $V = 0.04$] or RT [RT Nouns: $M = 556$ ms,

$SE = 6$ ms; RT Verbs: $M = 543$ ms, $SE = 8$ ms, $t(267.8) = 1.19$, $p = 0.23$, Cohen's $d = 0.15$]. A summary of HS' LDT performance in comparison to results of healthy control participants can be found in **Figure 3**.

Patient CA

CA exhibited a strong impairment for the domain of abstract-emotional nouns, with an ACC of 0.43 while the other noun categories showed to be relatively intact in comparison (ACC $= 0.78$ on average). These differences were revealed to be statistically significant, both generally, between noun categories ($\chi^2 = 19.15$, $df = 3$, $p < 0.001$, Cramer's $V = 0.37$) and for the comparison of the abstract-emotional nouns versus the other categories combined ($\chi^2 = 18.13$, $df = 1$, $p < 0.001$, Cramer's $V = 0.34$), whereas *post hoc* comparisons for the other semantic noun categories yielded no significant differences (all $p > 0.2$, n.s.). Furthermore, performance on abstract nouns was also more error-prone than that on the non-action reference category
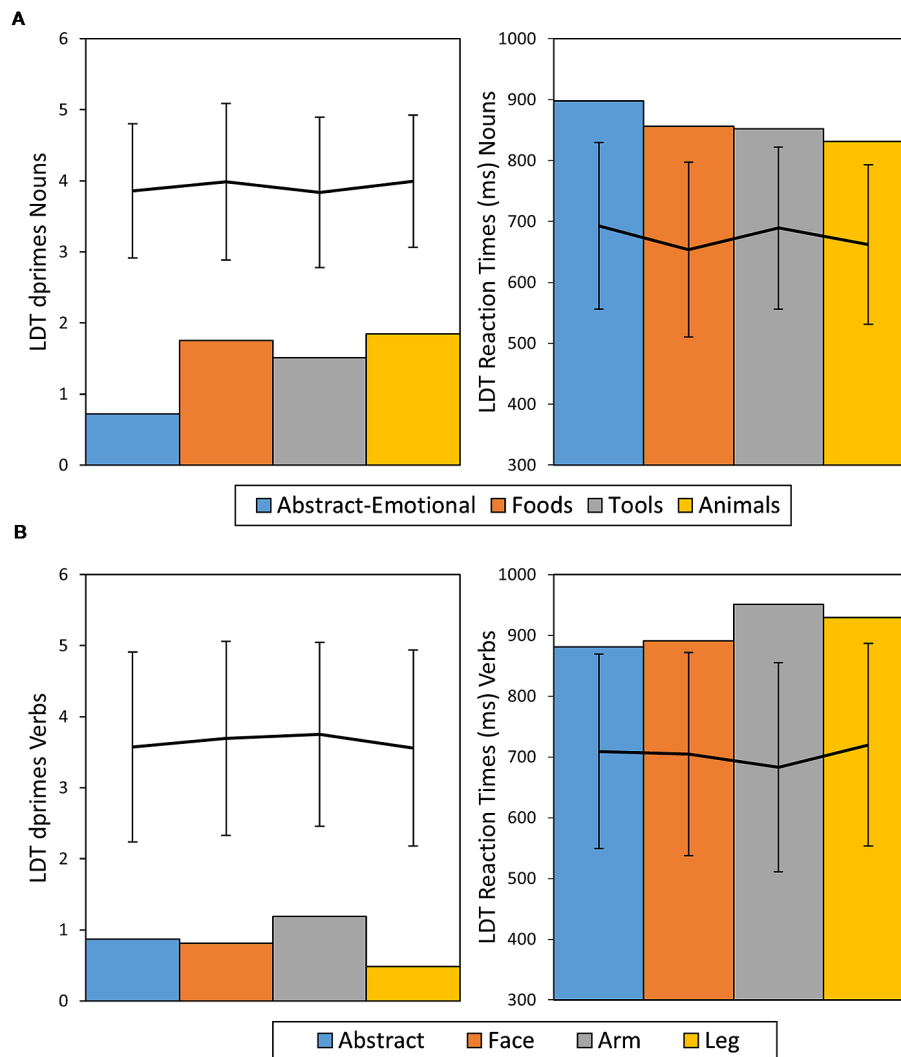
FIGURE 4 | Dprime and RT results of patient CA for nouns (A) and verbs (B) given in bar charts. The line diagrams represent average performance of the control sample with error bars representing ± 2 SDs.

of animal nouns ($\chi^2 = 13.65$, $df = 1$, $p <0.001$, Cramer's $V = 0.41$). *Post hoc* RSDT results confirmed this observation, as the difference in ACC between animal and abstract nouns was significantly more severe in patient CA than in the control sample [$t(20) = -2.05$, $p = 0.027$, Z-DCC $= -2.19$] and likewise the comparison of abstract nouns vs. all other noun categories combined [$t(20) = 3.15$, $p = 0.002$, Z-DCC $= -3.39$]. Finally, even the pairwise χ^2 noun category comparisons showed abstract word ACCs to be lower compared with each of the other noun categories (all $p < 0.05$, Bonferroni corrected), whereas the other noun groups did not significantly differ between each other.

For verbs, overall ACC was poor across categories (0.46 on average) and differences between categories were not significant ($\chi^2 = 6.13$, $df = 3$, $p = 0.11$, Cramer's $V = 0.2$). Analysis of RTs did not show significant effects of semantic word category in either nouns[$F(3,103) = 0.78$, $p > 0.2$, $\eta^2 = 0.02$, n.s.] or

verbs [$F(3,67) = 0.71$, $p = 0.2$, $\eta^2 = 0.03$, n.s.]. Across semantic categories, performance was worse for verbs, than for nouns, as measured by ACC (ACC Nouns $= 0.69$, ACC Verbs $= 0.46$, $\chi^2 = 17.5$, $df = 1$, $p < 0.001$, Cramer's $V = 0.23$) and RT (RTs Nouns $M = 853$ ms, $SE = 14$ ms, Verbs $M = 913$ ms, $SE = 20$ ms, $t(176) = 2.5$, $p = 0.01$, Cohen's $d = 0.38$). Taking the healthy participant sample as a benchmark, RTs and ACCs were considerably impaired across all noun and verb categories in patient CA, with all measures being outside of the range of ± 2 SDs from the mean of the control sample. **Figure 4** provides an overview of CA's LDT results in comparison to performance of healthy control participants.

Post Hoc Matching of Semantic Categories for RTs
Despite the careful matching of word stimuli for psycholinguistic features, RTs in healthy control subjects happened to differ significantly between semantic categories within nouns and verbs

classes. To investigate whether this RT difference may affect the patterns of category specificity seen on ACC data in our patients, an additional *post hoc* stimulus matching was performed, now using average RTs in the healthy control cohort as an additional matching criterion. This was done by removing 20% (i.e., eight) of the items of each semantic noun category, those with the shortest average RT for foods and tools and the 20% slowest items for abstract and animal nouns. The resulting item set did no longer show significant RT differences in the healthy controls [By-subjects: $F(3,60) = 1.19$, $p > 0.2$, $\eta^2 = 0.06$; by-items: $F(3,124) = 1.02$, $p > 0.2$, $\eta^2 = 0.02$] while, the previously reported category-specific patterns in the patients' ACC data could be confirmed for the same item selection on χ^2 and RSDT measures, with patient CA showing the selective deficit for abstract emotional nouns compared to the other categories [$\chi^2 = 5.4$, $df = 1$, $p = 0.02$, Cramer's $V = 0.21$; RSDT: $t(20) = 1.95$, $p = 0.03$, z-DCC $= -2.1$] and patient HS exhibiting a selective impairment for tools compared to the other semantic noun categories [$\chi^2 = 6.4$, $df = 1$, $p = 0.02$, Cramer's $V = 0.22$; RSDT: $t(20) = 2.75$, $p = 0.01$, z-DCC $= -2.96$].

DISCUSSION

Two patients with focal lesions in their dorsal frontocentral primary, premotor, and supplementary motor areas participated in standard aphasia tests as well as a speeded lexical decision paradigm. Albeit general aphasia measures, including comprehension tests, did not indicate neurological language disorders, LDT results of both patients revealed differential impairments of semantic categories of nouns. In patient CA, who suffered from a focal lesion of the left SMA, this impairment was most pronounced for abstract-emotional nouns, whereas patient HS, who suffered from a mild paresis of the right extremities and focal lesion just inferior to the typical hand representation in the primary motor cortex, showed a category-specific deficit in recognizing tool-related nouns. Both observations show that the motor system can be necessary for recognizing and processing of words from specific semantic categories. HS' data confirm a necessary role of motor cortex for action-related tool word processing and CA's results show that motor systems, especially the SMA, can be of relevance for abstract-emotional symbols. These results refute the hypothesis that motor brain areas play merely an epiphenomenal role in processing words with action-related and abstract meaning. It is clear that our present results emerging from the performance patterns of two patients cannot motivate general conclusions on all patients with similar lesions. Single case studies as the ones presented provide the existence proof that category-specific action-semantic deficits can arise from motor system lesions and this observation can be computed against the predictions of established semantic brain theories, as discussed below.

Category-specificity, General Cognitive Deficits, and Lesion Localization

Our proposed conclusions on category-specific semantic deficits imply that the observed performance pattern cannot be a result of general cognitive or linguistic impairments in the patients. Clinical language performance revealed by AAT results showed almost errorless performance and therefore demonstrates absence of aphasia. In particular, the excellent results obtained by both patients on the subtest on Word Comprehension show good general reading skills, which are important for written word and pseudo-word processing required in the LDT. Despite the absence of aphasia, both patients exhibited impairments in the LDT, arguably due to its higher processing load, especially the strict time constraints and emphasis on accuracy, compared with clinical testing with the AAT battery, where speed is not an issue. One might still argue that, possibly, additional general cognitive deficits, e.g., in praxis, attention, memory, or planning might have been present in the patients, but remained undetected and may have affected the results. However, general cognitive deficits can be expected to lead to reduced performance across semantic word categories. In contrast, the processing deficits observed in both patients, which were significantly most pronounced for one semantic word category, argues against an explanation in terms of general cognitive defects and in favor of one emphasizing specific and semantic origins. It is possible that the overall very poor LDT performance across all semantic categories seen in CA was due to the functional role of the affected SMA and adjacent pre-SMA in decision and motor response selection (Hernández et al., 2002; Forstmann et al., 2008; Nachev et al., 2008); however, again, the fact that this deficit was most pronounced for abstract-emotional nouns cannot be explained by such a general cognitive processing impairment.

Given the etiology of the lesions in CA and HS, it may seem that the lesions might not entail the focality needed to draw conclusions about the functional roles of specific brain areas, as proposed in the introduction. As discussed by Karnath and Steinbach (2011), it might be problematic to precisely tell apart functional and non-functional tissue in brain tumor patients. Furthermore, Karnath and Steinbach (2011) also highlight the possibility that gradual functional reorganization may continuously occur during the extended period of tumor growth, thus compensating for the impaired functionality. Both objections would resemble a blurring of the inferences that can be drawn on the functional role of lesioned brain areas. With regard to the former objection, it has to be noted that HS showed a rather circumscribed lesion and patient CA's metastasis (in contrast to other tumors like for example gliomas) did allow for fine grained differentiation between lesioned and non-lesioned tissue. In addition, the disadvantage of a possibly poor spatial resolution of causal inferences on the functional role of brain areas is not unique to tumor patients, but indeed resembles a general problem for all kinds of lesion studies (Shallice and Skrap, 2011). This has been investigated in detail, for example in the context of reperfusion of the 'penumbra' of stroke-related lesions (Hillis et al., 2006). Similarly, the argument of better functional restitution in tumor patients does not apply to the current cases, as a category-specific deficits were in fact manifest and detectable using psycholinguistic methods in both patients, whereas functional reorganization would have predicted absence of such specificity. Even if

functional reorganization occurred in any of the patients, it can be assumed to be insufficient to recover normal function so that a functional role of the lesioned brain areas can still be soundly derived (Duffau, 2011). However, we should remark that for many other patients, the argument is still valid and significant category-deficits may not arise from motor systems lesions. Functional reorganization provides one important reason why category differences may be frequently absent after focal lesions.

With regard to the lesion in patient HS, it has to be noted that although ADEMs are most often diagnosed with multiple lesion foci (Karussis, 2014; Koudriavtseva et al., 2015), cases with monofocal lesions have been reported on multiple occasions (Kesselring et al., 1990; Miller et al., 1993; Murthy et al., 1999), allowing to assume a focal etiology. In patient CA, who suffered from a circumscribed metastasis, areas adjacent to the lesioned SMA, including pre-SMA and primary and pre-motor cortex, may have been affected in their function. We should, however, draw attention to the fact that the patient's tumor had been subject to intensive therapy previous to testing, including partial extirpation, so that it appears unlikely that pressure was exerted on adjacent areas. Still, the possibility that partial lesion of pre-SMA played some role in causing the deficit in abstract noun processing cannot be ruled out with certainty based on our present data.

Clinical observations are consistent with the claim that patient CA, but not HS, was suffering from depressive symptoms at the time of testing, although this could not be objectified using psychological tests. Intuitively, this depressive mood could be seen as a (non-neurological) reason for the processing deficit for abstract-emotional words. However, previous studies either indicated that LDT performance was not affected by depression (Clark et al., 1983; Challis and Krane, 1988) or even led to a facilitation of LDT performance on emotionally congruent word stimuli (Olafson and Ferraro, 2001). Note that most of the abstract emotion words used in the present study were negative in valence and therefore congruent with the negative emotional state of depression. The observed performance reduction for abstract emotion words seen in patient CA contrasts with these earlier observations, rendering it unlikely that the observed category-specific semantic word processing deficit was based on emotional state of the patient at the time of language testing.

Apart from the neurological and clinical factors mentioned above, one could try to argue that the specific impairments found in the two patients might in fact not be due to compromised processing of word semantics, but rather to impairments of basic visual or linguistic processing. It is well-known that, in order to solve the LDT, it is not necessary to engage semantic processing, because words, but not pseudo-words, are familiar entities stored as whole lexical entries in the brain-internal 'mental lexicon'. Nevertheless, the LDT paradigm has previously been shown to be sensitive to manipulation of semantic content (James, 1975; Chumbley and Balota, 1984; Kroll and Merves, 1986; Jin, 1990; Samson and Pillon, 2004) and a range of pre-existing neuropsychological studies demonstrated category-specificity in processing semantic word categories after focal

brain lesions (Gainotti, 2010). In the present study, the examined semantic word categories, within each greater lexical category, were meticulously matched for a range of psycholinguistic features, including word length, lemma frequency, character, bi- and trigram frequencies and their word-initial counterparts, as well as number and word frequency of orthographic neighbors. Therefore, the observed category-effects can soundly be attributed to differences in word semantics and not to sublexical, morphological or other psycholinguistic properties, some of which have previously been shown to modulate the activity of motor areas during language processing, independent of semantics (Pulvermüller et al., 2006; de Zubicaray et al., 2013). In addition, the close matching of words and pseudo-words with regard to character, bi- and trigram frequencies as well as word initial character and bigram frequencies, argues against the possibility that sublexical strategies played a role in the present LDT, instead of actual semantic processing of target stimuli.

Category-effects Across Participants, Measures, and Lexical Classes

In contrast to the category-specific patterns shown by both patients with focal lesions in the motor system, d' and ACC data showed that the healthy control population performed similarly on all semantic noun categories and the same applied for the matched verb categories too. However, semantic category differences may be suggested by the control subjects' response time data, which yielded significant differences due to slightly slower responses to abstract and hand-action related nouns. These were the two categories, respectively, affected in our patients. To examine the theoretical possibility that the processing difference suggested by controls' RT data may explain the category specific patterns in our patients, analyses were repeated with a subset of the word stimuli matched for response times in healthy controls. The RT-matched semantic word category sets did not yield any significant performance difference in our healthy subjects, neither in ACCs nor in RTs, but the category-differences for semantic noun categories in both patients' ACC values were reconfirmed. These results rule out the possibility that, whatever might have caused the RT differences in our control population could explain the category differences seen in the patients.

In both patients, the category specific impairments were only found for nouns, but not for verbs. This observation might appear surprising, as the majority of previous studies on motor semantics highlighted the role of the sensory-motor systems for the processing of action verbs. Considering the stimuli selected for the present LDT though, one cannot conclude from this result that the functional role of motor areas applies exclusively to the processing of nouns. The experimental setup was designed to compare processing of semantic categories separately within semantic subtypes of nouns and, again, for subtypes of verbs. Because psycholinguistic matching was not performed across noun and verb categories, a direct comparison between the lexical classes is not straightforward. For example, verbs had higher lemma frequencies than nouns

and therefore were more familiar. This implies that the LDT was generally easier for verbs compared with nouns. At the same time, pseudo-verbs consistently differed in only one syllable from proper verbs (because of the shared suffix '-en'), whereas nouns differed between each other in both of their syllables, thus making it necessary to process more information for making lexical decisions on nouns than on verbs. In addition, it is well known that verbs carry more syntactic information and are generally more strongly action-related semantically but are, on the other hand, less imageable than nouns (Pulvermüller et al., 1999; Bird et al., 2000). Some of these differences between the lexical categories (e.g., the greater imageability of nouns) may underlie the observed processing advantage of nouns over verbs, as found in the healthy controls' d' and RT results and in CA's reduced performance on all verb categories. These general psycholinguistic differences between nouns and verbs may also in part account for the fact that category differences could only be documented for one of the lexical categories, because a difference on one of the psycholinguistic dimensions may have moved one of the categories away from a ceiling or floor so that performance differences could become selectively manifest.

While patient HS' overall performance for verbs on the LDT was comparable to that of healthy controls, results for CA revealed a strong impairment across all verb categories, which was only paralleled by the severely affected abstract word category of nouns. Being aware of the mentioned psycholinguistic differences between our lexical class stimuli, we should still mention the possibility that the latter observation could, in theory, originate from the relatively higher relevance of action knowledge for the semantics of verbs. From an Embodied Cognition perspective, the observed impairment for all verb categories with action dominant semantics seem to fit to CA's lesion site in the left SMA, an area known to be involved in motor planning independent of motor effector and body part (Roland et al., 1980; Fried et al., 1991). Nevertheless, given that potential differences in task difficulty cannot be ruled out when comparing nouns and verbs, this interpretation has to be treated with caution before less ambiguous experimental evidence is available. In the case of patient HS, the fact that no semantic category effects were seen for verbs could be seen as a side effect of the high performance close to ceiling for verbs, whereas average performance on nouns was relatively reduced. Our data did not show significant differences in processing different semantic sub-categories of verbs, thus confirming the corresponding observation by Arevalo et al. (2012). To disentangle the possible factors influencing verb and noun performance, future studies should aim to match semantic categories between those grammatical word classes in terms of semantic features, psycholinguistic characteristics as well as general task difficulty. However, we once again remind the reader that such matching is not trivial and might be not possible on all dimensions (for discussion, see Bird et al., 2000; Neininger and Pulvermüller, 2003).

Relationship of the Present Results to Known Neuropsychological Dissociations

The reported selective impairment for tool nouns in patient HS adds to previous findings on impairments in neurological patients, specifically for words with action related semantics (Bak et al., 2001, 2006; Neininger and Pulvermüller, 2001, 2003; Pulvermüller et al., 2010; Arevalo et al., 2012; Kemmerer et al., 2012). In contrast to these earlier works, the current study shows that those selective impairments can be induced by rather small focal lesions (of 18 mm diameter in the case of HS) in the motor areas and confirms that the corresponding category-specific semantic deficits are not restricted to action-related verbs but can also arise for nouns used to speak about objects that afford actions, as for example tool words. HS' results on tool nouns also fit well with the results of earlier neuro-stimulation experiments, which pointed out the functional relevance of motor areas for action verb processing, using facilitatory (Pulvermüller et al., 2005a) or virtual lesion approaches (Willems et al., 2011), although in those studies effects were found solely on RTs. As substantial numbers of errors were here documented to arise from motor system lesion for nouns with action-affording referents, the present results show a necessary role of motor and premotor cortex in one single neurological case. Over and above previous research, we show a rather narrow level of category-specificity, in so far as it applied only to nouns used to speak about objects affording actions typically performed with the hand. This specificity is consistent with semantic somatotopy in the motor system (Pulvermüller, 2005; Pulvermüller and Fadiga, 2010).

Observations on the performance of patient CA on the other hand revealed a functional involvement of supplementary motor systems also for the processing of abstract-emotional nouns, which lack the transparent sensory-motor components of their concrete counterparts. This can be seen as first evidence that activity in motor areas during the processing of abstract-emotional nouns, as revealed by earlier fMRI results (Moseley et al., 2012), does in fact not resemble an epiphenomenon, but an integral part of word comprehension instead, which is necessary for optimal word processing. This result appears consistent with semantic grounding theories postulating involvement of motor circuits in abstract semantic processing, thus suggesting that the 'embodiment' does not necessarily need to limit its scope to the processing of words referring to concrete entities. At the theoretical level, there is indeed motivation to see an intrinsic connection between abstract-emotional meaning and the bodily actions with which such meanings are expressed (for discussion, see Barsalou and Wiemer-Hastings, 2005; Moseley et al., 2012; Pulvermüller, 2013). Whether this holds exclusively for abstract-emotional words, or renders an effect that is valid also for non-emotional abstract symbols and concepts, has to be determined by future studies, for example by investigating stimuli across different subcategories of abstract words.

Distributed Semantic Circuit Account of the Current Results

In order to explain the category-preferential semantic deficit in processing hand-action-affording and abstract-emotional nouns, one may claim that our results are consistent with theories that view the motor system as the main carrier of meaning processing for these specific semantic types. Although such strong statements – that motor cortices but no other areas integrate concepts and word meanings – have hardly been made, some arguments against semantic grounding (e.g., in Mahon and Caramazza, 2008) seem to focus on this hypothetical position. Indeed, some authors have stated "that the modalities of action and perception are integrated at the level of the sensorimotor system itself and not via higher association areas" (Gallese and Lakoff, 2005, p. 459), and such statements may have laid the ground for the idea that motor systems, but not association or convergence zones such as the prefrontal or anterior-temporal cortex, might carry meaning. Although even such a strong postulate about semantic integration in motor but no other multimodal brain systems could indeed be strengthened by the present data, it is not the only position that explains the present results. Considering a wider spectrum of data, which also show semantic activation of and semantic deficits after lesion in multimodal areas (e.g., Patterson et al., 2007; Binder et al., 2009; Vigliocco et al., 2014), the more appropriate explanation of the present data needs to be phrased in terms of distributed semantic circuits in which neurons in motor areas play a functional, causal and necessary role.

In this perspective, the sensorimotor parts of the distributed semantic circuits would carry aspects of word meaning and contribute to a process of immediate 'simulation' of semantic information (in the sense of Jeannerod, 2006) when symbols are perceived, even if subjects do not actively attend to them (Pulvermüller et al., 2005b; Shtyrov et al., 2014). Therefore, the observations made in both patients on nouns seem to fit especially well into theoretical frameworks that assume distributed cell assemblies with different cortical distributions to be the basis of semantic processing of words (Pulvermüller, 1999). Those cell assemblies are assumed to be the result of correlational learning mechanisms driven by Hebbian learning principles (Hebb, 1949). If a word often co-occurs with specific sensory and or motor experiences, or likewise with specific sensory or motor imagery, that word's semantic circuit would gradually be represented by a distributed cell assembly reaching into the sensory or motor areas where relevant activations had been present. A word like "hammer" co-occurring with performance, perception or imagery of specific motor movements afforded by the tool, would co-activate the perception action circuit for the word form and the action-related neuronal circuit, thus yielding a higher-order distributed semantic circuit in which neurons in motor areas take a causal and necessary functional role. This proposal does not postulate a unique role of the motor system (or 'modality specific cortices') as a seat of semantics, but a semantic role of cortical circuits distributed over perisylvian, sensorimotor and multimodal convergence areas. Specificity in cortical function arises from the fact that, for different meaning types, these semantic circuits have different cortical distributions – with some (action related) semantic circuits, but not others (non-action related ones), reaching into the motor system. Importantly, in this view, the word 'hammer' is not exhaustively semantically processed in multimodal areas, as postulated by disembodiment (or weak 'integrative') approaches to semantics, and there is no preferential status of the motor system for semantics either. Semantic circuits for abstract-emotional words would include neurons in the limbic system – because emotional-affective 'inner states' are essential for at least some abstract words (Meteyard et al., 2012) – and in the motor system – because the learning of at least some abstract-emotional words requires the grounding of word forms in emotions expressed in overt body movements (Moseley et al., 2012). This integrative action perception model appears to us to be consistent with known lesion results on brain-lesion-elicited semantic impairments (Kiefer and Pulvermüller, 2012; Pulvermüller, 2013) and to do best justice to the present data.

CONCLUSION

Category-specific semantic deficits in a LDT seen in two patients with focal lesions in their left hemispheres reveal the functional necessity of primary/pre- and supplementary motor areas for the processing of concrete hand-action affording as well as for abstract-emotional nouns. Processing of concrete tool nouns was selectively impaired after lesions of hand motor cortex, while a lesion in the left SMA resulted in impaired processing of abstract-emotional nouns.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpsyg.2015.01661

REFERENCES

Arevalo, A. L., Baldo, J. V., and Dronkers, N. F. (2012). What do brain lesions tell us about theories of embodied semantics and the human mirror neuron system? *Cortex* 48, 242–254. doi: 10.1016/j.cortex.2010.06.001

Bak, T. H., O'Donovan, D. G., Xuereb, J. H., Boniface, S., and Hodges, J. R. (2001). Selective impairment of verb processing associated with pathological changes in Brodmann areas 44 and 45 in the motor neurone disease–dementia–aphasia syndrome. *Brain* 124, 103–120. doi: 10.1093/brain/124.1.103

Bak, T. H., Yancopoulou, D., Nestor, P. J., Xuereb, J. H., Spillantini, M. G., Pulvermüller, F., et al. (2006). Clinical, imaging and pathological correlates of a hereditary deficit in verb and action processing. *Brain* 129, 321–332. doi: 10.1093/brain/awh701

Barrós-Loscertales, A., González, J., Pulvermüller, F., Ventura-Campos, N., Bustamante, J. C., Costumero, V., et al. (2012). Reading salt activates gustatory brain regions: fMRI evidence for semantic grounding in a novel sensory modality. *Cereb. Cortex* 22, 2554–2563. doi: 10.1093/cercor/bhr324

Barsalou, L. W. (1999). Perceptions of perceptual symbols. *Behav. Brain Sci.* 22, 637–660. doi: 10.1017/S0140525X99532147

Barsalou, L. W. (2008). Grounded cognition. *Annu. Rev. Psychol.* 59, 617–645. doi: 10.1146/annurev.psych.59.103006.093639

Barsalou, L. W., and Wiemer-Hastings, K. (2005). "Situating abstract concepts," in *Grounding Cognition: The Role of Perception and Action in Memory, Language, and Thought*, eds D. Pecher and R. Zwaan (Oxford: Oxford University Press), 129–163.

Bedny, M., and Caramazza, A. (2011). Perception, action, and word meanings in the human brain: the case from action verbs. *Ann. N. Y. Acad. Sci.* 1224, 81–95. doi: 10.1111/j.1749-6632.2011.06013.x

Binder, J. R., and Desai, R. H. (2011). The neurobiology of semantic memory. *Trends Cogn. Sci.* 15, 527–536. doi: 10.1016/j.tics.2011.10.001

Binder, J. R., Desai, R. H., Graves, W. W., and Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cereb. Cortex* 19, 2767–2796. doi: 10.1093/cercor/bhp055

Bird, H., Ralph, M. A. L., Patterson, K., and Hodges, J. R. (2000). The rise and fall of fquency and imageability: noun and verb production in semantic dementia. *Brain Lang.* 73, 17–49. doi: 10.1006/brln.2000.2293

Bonner, M. F., and Grossman, M. (2012). Gray matter density of auditory association cortex relates to knowledge of sound concepts in primary progressive aphasia. *J. Neurosci.* 32, 7986–7991. doi: 10.1523/JNEUROSCI.6241-11.2012

Borghi, A. M., and Cimatti, F. (2009). "Words as tools and the problem of abstract words meanings," in *Proceedings of the 31st Annual Conference of the Cognitive Science Society*, eds N. Taatgen and H. van Rijn (Amsterdam: Cognitive Science Society), 2304–2309.

Buonomano, D. V., and Merzenich, M. M. (1998). Cortical plasticity: from synapses to maps. *Annu. Rev. Neurosci.* 21, 149–186. doi: 10.1146/annurev.neuro.21.1.149

Campanella, F., D'Agostini, S., Skrap, M., and Shallice, T. (2010). Naming manipulable objects: anatomy of a category specific effect in left temporal tumours. *Neuropsychologia* 48, 1583–1597. doi: 10.1016/j.neuropsychologia.2010.02.002

Caramazza, A. (1986). On drawing inferences about the structure of normal cognitive systems from the analysis of patterns of impaired performance: the case for single-patient studies. *Brain Cogn.* 5, 41–66. doi: 10.1016/0278-2626(86)90061-8

Caramazza, A., Anzellotti, S., Strnad, L., and Lingnau, A. (2014). Embodied cognition and mirror neurons: a critical assessment. *Annu. Rev. Neurosci.* 37, 1–15. doi: 10.1146/annurev-neuro-071013-013950

Carota, F., Moseley, R., and Pulvermüller, F. (2012). Body-part-specific representations of semantic noun categories. *J. Cogn. Neurosci.* 24, 1492–1509. doi: 10.1162/jocn_a_00219

Challis, B. H., and Krane, R. V. (1988). Mood induction and the priming of semantic memory in a lexical decision task: asymmetric effects of elation and depression. *Bull. Psychon. Soc.* 26, 309–312. doi: 10.3758/BF03337666

Chumbley, J. I., and Balota, D. A. (1984). A word's meaning affects the decision in lexical decision. *Mem. Cogn.* 12, 590–606. doi: 10.3758/BF03213348

Clark, D. M., Teasdale, J. D., Broadbent, D. E., and Martin, M. (1983). Effect of mood on lexical decisions. *Bull. Psychon. Soc.* 21, 175–178. doi: 10.3758/BF03334679

Crawford, J. R., and Garthwaite, P. H. (2005). Testing for suspected impairments and dissociations in single-case studies in neuropsychology: evaluation of alternatives using monte carlo simulations and revised tests for dissociations. *Neuropsychology* 19, 318–331. doi: 10.1037/0894-4105.19.3.318

Crawford, J. R., and Garthwaite, P. H. (2006). Methods of testing for a deficit in single-case studies: evaluation of statistical power by Monte Carlo simulation. *Cogn. Neuropsychol.* 23, 877–904. doi: 10.1080/02643290500538372

Damasio, A. R., and Tranel, D. (1993). Nouns and verbs are retrieved with differently distributed neural systems. *Proc. Natl. Acad. Sci. U.S.A.* 90, 4957–4960. doi: 10.1073/pnas.90.11.4957

Damasio, H., Grabowski, T. J., Tranel, D., Hichwa, R. D., and Damasio, A. R. (1996). A neural basis for lexical retrieval. *Nature* 380, 499–505. doi: 10.1038/380499a0

de Zubicaray, G., Arciuli, J., and McMahon, K. (2013). Putting an "end" to the motor cortex representations of action words. *J. Cogn. Neurosci.* 25, 1957–1974. doi: 10.1162/jocn_a_00437

Dove, G. (2009). Beyond perceptual symbols: a call for representational pluralism. *Cognition* 110, 412–431. doi: 10.1016/j.cognition.2008.11.016

Duffau, H. (2011). Do brain tumours allow valid conclusions on the localization of human brain functions? *Cortex* 47, 1016–1017. doi: 10.1016/j.cortex.2010.11.010

Elbert, T., Pantev, C., Wienbruch, C., Rockstroh, B., and Taub, E. (1995). Increased cortical representation of the fingers of the left hand in string players. *Science* 270, 305–307. doi: 10.1126/science.270.5234.305

Ellis, A. W., and Young, A. W. (1988). *Human Cognitive Neuropsychology*. Hove: Lawrence Erlbaum Associates Ltd.

Forstmann, B. U., Dutilh, G., Brown, S., Neumann, J., Von Cramon, D. Y., Ridderinkhof, K. R., et al. (2008). Striatum and pre-SMA facilitate decision-making under time pressure. *Proc. Natl. Acad. Sci. U.S.A.* 105, 17538–17542. doi: 10.1073/pnas.0805903105

Frey, D., Strack, V., Wiener, E., Jussen, D., Vajkoczy, P., and Picht, T. (2012). A new approach for corticospinal tract reconstruction based on navigated transcranial stimulation and standardized fractional anisotropy values. *Neuroimage* 62, 1600–1609. doi: 10.1016/j.neuroimage.2012.05.059

Fried, I., Katz, A., McCarthy, G., Sass, K. J., Williamson, P., Spencer, S. S., et al. (1991). Functional organization of human supplementary motor cortex studied by electrical stimulation. *J. Neurosci.* 11, 3656–3666.

Gainotti, G. (2010). The influence of anatomical locus of lesion and of gender-related familiarity factors in category-specific semantic disorders for animals, fruits and vegetables: a review of single-case studies. *Cortex* 46, 1072–1087. doi: 10.1016/j.cortex.2010.04.002

Gallese, V., and Lakoff, G. (2005). The brain's concepts: The role of the sensory-motor system in conceptual knowledge. *Cogn. Neuropsychol.* 22, 455–479. doi: 10.1080/02643290442000310

Glenberg, A. M., and Gallese, V. (2012). Action-based language: a theory of language acquisition, comprehension, and production. *Cortex* 48, 905–922. doi: 10.1016/j.cortex.2011.04.010

Glenberg, A. M., Sato, M., and Cattaneo, L. (2008). Use-induced motor plasticity affects the processing of abstract and concrete language. *Curr. Biol.* 18, R290–R291. doi: 10.1016/j.cub.2008.02.036

González, J., Barros-Loscertales, A., Pulvermüller, F., Meseguer, V., Sanjuán, A., Belloch, V., et al. (2006). Reading cinnamon activates olfactory brain regions. *Neuroimage* 32, 906–912. doi: 10.1016/j.neuroimage.2006.03.037

Hauk, O., Johnsrude, I., and Pulvermüller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron* 41, 301–307. doi: 10.1016/S0896-6273(03)00838-9

Hauk, O., and Pulvermüller, F. (2004). Neurophysiological distinction of action words in the fronto-central cortex. *Hum. Brain Mapp.* 21, 191–201. doi: 10.1002/hbm.10157

Havas, D. A., Glenberg, A. M., Gutowski, K. A., Lucarelli, M. J., and Davidson, R. J. (2010). Cosmetic use of botulinum toxin-a affects processing of

emotional language. *Psychol. Sci.* 21, 895–900. doi: 10.1177/0956797610374742

Hebb, D. O. (1949). *The Organization of Behavior; A Neuropsychological Theory.* Hove: Psychology Press.

Heister, J., Würzner, K. M., Bubenzer, J., Pohl, E., Hanneforth, T., Geyken, A., et al. (2011). dlexDB–eine lexikalische Datenbank für die psychologische und linguistische Forschung. *Psychol. Rundsch.* 62, 10–20. doi: 10.1026/0033-3042/a000029

Hernández, A., Zainos, A., and Romo, R. (2002). Temporal evolution of a decision-making process in medial premotor cortex. *Neuron* 33, 959–972. doi: 10.1016/S0896-6273(02)00613-X

Hillis, A. E., Kleinman, J. T., Newhart, M., Heidler-Gary, J., Gottesman, R., Barker, P. B., et al. (2006). Restoring cerebral blood flow reveals neural regions critical for naming. *J. Neurosci.* 26, 8069–8073. doi: 10.1523/JNEUROSCI.2088-06.2006

Huber, W., Poeck, K., Weniger, D., and Willmes, K. (1983). *Aachener Aphasie Test (AAT).* Göttingen: Hogrefe.

James, C. T. (1975). The role of semantic information in lexical decisions. *J. Exp. Psychol.* 1, 130.

Jeannerod, M. (2006). *Motor Cognition: What Actions Tell the Self*, Vol. 42. Oxford: Oxford University Press.

Jin, Y. S. (1990). Effects of concreteness on cross-language priming in lexical decisions. *Percept. Mot. Skills* 70, 1139–1154. doi: 10.2466/pms.1990.70.3c.1139

Karnath, H. O., and Steinbach, J. P. (2011). Do brain tumours allow valid conclusions on the localisation of human brain functions?–Objections. *Cortex* 47, 1004–1006. doi: 10.1016/j.cortex.2010.08.006

Karussis, D. (2014). The diagnosis of multiple sclerosis and the various related demyelinating syndromes: a critical review. *J. Autoimmun.* 48, 134–142. doi: 10.1016/j.jaut.2014.01.022

Kemmerer, D., Rudrauf, D., Manzel, K., and Tranel, D. (2012). Behavioral patterns and lesion sites associated with impaired processing of lexical and conceptual knowledge of actions. *Cortex* 48, 826–848. doi: 10.1016/j.cortex.2010.11.001

Kesselring, J., Miller, D. H., Robb, S. A., Kendall, B. E., Moseley, I. F., Kingsley, D., et al. (1990). Acute disseminated encephalomyelitis. *Brain* 113, 291–302. doi: 10.1093/brain/113.2.291

Keuleers, E., and Brysbaert, M. (2010). Wuggy: a multilingual pseudoword generator. *Behav. Res. Methods* 42, 627–633. doi: 10.3758/BRM.42.3.627

Kiefer, M., and Pulvermüller, F. (2012). Conceptual representations in mind and brain: theoretical developments, current evidence and future directions. *Cortex* 48, 805–825. doi: 10.1016/j.cortex.2011.04.006

Kiefer, M., Sim, E. J., Herrnberger, B., Grothe, J., and Hoenig, K. (2008). The sound of concepts: four markers for a link between auditory and conceptual brain systems. *J. Neurosci.* 28, 12224–12230. doi: 10.1523/JNEUROSCI.3579-08.2008

Koudriavtseva, T., Renna, R., Plantone, D., and Mainero, C. (2015). Demyelinating and thrombotic diseases of the central nervous system: common pathogenic and triggering factors. *Front. Neurol.* 6:63. doi: 10.3389/fneur.2015.00063

Kroll, J. F., and Merves, J. S. (1986). Lexical access for concrete and abstract words. *J. Exp. Psychol.* 12, 92.

Mahon, B. Z., and Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *J. Physiol. Paris* 102, 59–70. doi: 10.1016/j.jphysparis.2008.03.004

Martin, A., Wiggs, C. L., Ungerleider, L. G., and Haxby, J. V. (1996). Neural correlates of category-specific knowledge. *Nature* 379, 649–652. doi: 10.1038/379649a0

Meteyard, L., Cuadrado, S. R., Bahrami, B., and Vigliocco, G. (2012). Coming of age: a review of embodiment and the neuroscience of semantics. *Cortex* 48, 788–804. doi: 10.1016/j.cortex.2010.11.002

Miller, D. H., Scaravilli, F., Thomas, D. C., Harvey, P., and Hirsch, N. P. (1993). Acute disseminated encephalomyelitis presenting as a solitary brainstem mass. *J. Neurol. Neurosurg. Psychiatry* 56, 920–922. doi: 10.1136/jnnp.56.8.920

Moseley, R., Carota, F., Hauk, O., Mohr, B., and Pulvermüller, F. (2012). A role for the motor system in binding abstract emotional meaning. *Cerebral. Cortex* 22, 1634–1647. doi: 10.1093/cercor/bhr238

Murthy, J. M. K., Yangala, R., Meena, A. K., and Reddy, J. J. (1999). Acute disseminated encephalomyelitis: clinical and MRI study from South India. *J. Neurol. Sci.* 165, 133–138. doi: 10.1016/S0022-510X(99)00094-5

Nachev, P., Kennard, C., and Husain, M. (2008). Functional role of the supplementary and pre-supplementary motor areas. *Nat. Rev. Neurosci.* 9, 856–869. doi: 10.1038/nrn2478

Neininger, B., and Pulvermüller, F. (2001). The right hemisphere's role in action word processing: a double case study. *Neurocase* 7, 303–317. doi: 10.1093/neucas/7.4.303

Neininger, B., and Pulvermüller, F. (2003). Word-category specific deficits after lesions in the right hemisphere. *Neuropsychologia* 41, 53–70. doi: 10.1016/S0028-3932(02)00126-4

Olafson, K. M., and Ferraro, F. R. (2001). Effects of emotional state on lexical decision performance. *Brain Cogn.* 45, 15–20. doi: 10.1006/brcg.2000.1248

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113. doi: 10.1016/0028-3932(71)90067-4

Patterson, K., Nestor, P. J., and Rogers, T. T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nat. Rev. Neurosci.* 8, 976–987.

Postle, N., McMahon, K. L., Ashton, R., Meredith, M., and de Zubicaray, G. I. (2008). Action word meaning representations in cytoarchitectonically defined primary and premotor cortices. *Neuroimage* 43, 634–644. doi: 10.1016/j.neuroimage.2008.08.006

Pulvermüller, F. (1999). Words in the brain's language. *Behav. Brain Sci.* 22, 253–279. doi: 10.1017/S0140525X9900182X

Pulvermüller, F. (2005). Brain mechanisms linking language and action. *Nat. Rev. Neurosci.* 6, 576–582. doi: 10.1038/nrn1706

Pulvermüller, F. (2013). How neurons make meaning: brain mechanisms for embodied and abstract-symbolic semantics. *Trends Cogn. Sci.* 17, 458–470. doi: 10.1016/j.tics.2013.06.004

Pulvermüller, F., Cooper-Pye, E., Dine, C., Hauk, O., Nestor, P. J., and Patterson, K. (2010). The word processing deficit in semantic dementia: all categories are equal, but some categories are more equal than others. *J. Cogn. Neurosci.* 22, 2027–2041. doi: 10.1162/jocn.2009.21339

Pulvermüller, F., and Fadiga, L. (2010). Active perception: sensorimotor circuits as a cortical basis for language. *Nat. Rev. Neurosci.* 11, 351–360. doi: 10.1038/nrn2811

Pulvermüller, F., Härle, M., and Hummel, F. (2001). Walking or talking? Behavioral and neurophysiological correlates of action verb processing. *Brain Lang.* 78, 143–168. doi: 10.1006/brln.2000.2390

Pulvermüller, F., Hauk, O., Nikulin, V. V., and Ilmoniemi, R. J. (2005a). Functional links between motor and language systems. *Eur. J. Neurosci.* 21, 793–797. doi: 10.1111/j.1460-9568.2005.03900.x

Pulvermüller, F., Shtyrov, Y., and Ilmoniemi, R. (2005b). Brain signatures of meaning access in action word recognition. *J. Cogn. Neurosci.* 17, 884–892. doi: 10.1162/0898929054021111

Pulvermüller, F., Huss, M., Kherif, F., del Prado Martin, F. M., Hauk, O., and Shtyrov, Y. (2006). Motor cortex maps articulatory features of speech sounds. *Proc. Natl. Acad. Sci. U.S.A.* 103, 7865–7870. doi: 10.1073/pnas.0509989103

Pulvermüller, F., Lutzenberger, W., and Preissl, H. (1999). Nouns and verbs in the intact brain: evidence from event-related potentials and high-frequency cortical responses. *Cereb. Cortex* 9, 497–506. doi: 10.1093/cercor/9.5.497

Reilly, J., Harnish, S., Garcia, A., Hung, J., Rodriguez, A. D., and Crosson, B. (2014). Lesion symptom mapping of manipulable object naming in nonfluent aphasia: can a brain be both embodied and disembodied? *Cogn. Neuropsychol.* 31, 287–312. doi: 10.1080/02643294.2014.914022

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.

Samson, D., and Pillon, A. (2004). Orthographic neighborhood and concreteness effects in the lexical decision task. *Brain Lang.* 91, 252–264. doi: 10.1016/j.bandl.2004.03.001

Shallice, T., and Skrap, M. (2011). Localisation through operation for brain tumour: a reply to Karnath and Steinbach. *Cortex* 47, 1007–1009. doi: 10.1016/j.cortex.2010.12.006

Shebani, Z., and Pulvermüller, F. (2013). Moving the hands and feet specifically impairs working memory for arm-and leg-related action words. *Cortex* 49, 222–231. doi: 10.1016/j.cortex.2011.10.005

Shtyrov, Y., Butorina, A., Nikolaeva, A., and Stroganova, T. (2014). Automatic ultrarapid activation and inhibition of cortical motor systems in spoken word comprehension. *Proc. Natl. Acad. Sci. U.S.A.* 111, E1918–E1923. doi: 10.1073/pnas.1323158111

Trumpp, N. M., Kliese, D., Hoenig, K., Haarmeier, T., and Kiefer, M. (2013). Losing the sound of concepts: damage to auditory association cortex impairs the processing of sound-related concepts. *Cortex* 49, 474–486. doi: 10.1016/j.cortex.2012.02.002

Vigliocco, G., Kousta, S. T., Della Rosa, P. A., Vinson, D. P., Tettamanti, M., Devlin, J. T., et al. (2014). The neural representation of abstract words: the role of emotion. *Cereb. Cortex* 24, 1767–1777. doi: 10.1093/cercor/bht025

Warrington, E. K., and McCarthy, R. A. (1983). Category specific access dysphasia. *Brain* 106, 859–878. doi: 10.1093/brain/106.4.859

Warrington, E. K., and McCarthy, R. A. (1987). Categories of knowledge: further fractionations and an attempted integration. *Brain* 110, 1273–1296. doi: 10.1093/brain/110.5.1273

Warrington, E. K., and Shallice, T. (1984). Category specific semantic impairments. *Brain* 107, 829–854. doi: 10.1093/brain/107.3.829

Willems, R. M., Labruna, L., D'Esposito, M., Ivry, R., and Casasanto, D. (2011). A functional role for the motor system in language understanding evidence from theta-burst transcranial magnetic stimulation. *Psychol. Sci.* 22, 849–854. doi: 10.1177/0956797611412387

Witt, J. K., Kemmerer, D., Linkenauger, S. A., and Culham, J. (2010). A functional role for motor simulation in identifying tools. *Psychol. Sci.* 21, 1215–1219. doi: 10.1177/0956797610378307

Dopaminergic modulation of positive expectations for goal-directed action: evidence from Parkinson's disease

Noham Wolpe[1,2]*, Cristina Nombela[1,2] and James B. Rowe[1,2,3]

[1] Department of Clinical Neurosciences, University of Cambridge, Cambridge, UK, [2] Medical Research Council Cognition and Brain Sciences Unit, Cambridge, UK, [3] Behavioural and Clinical Neuroscience Institute, University of Cambridge, Cambridge, UK

Parkinson's disease (PD) impairs the control of movement and cognition, including the planning of action and its consequences. This provides the opportunity to study the dopaminergic influences on the perception and awareness of action. Here we examined the perception of the outcome of a goal-directed action made by medicated patients with PD. A visuomotor task probed the integration of sensorimotor signals with the positive expectations of outcomes (Self priors), which in healthy adults bias perception toward success in proportion to trait optimism. We tested the hypotheses that (i) the priors on the perception of the consequences of one's own actions differ between patients and age- and sex-matched controls, and (ii) that these priors are modulated by the levodopa dose equivalent (LDEs) in patients. There was no overall difference between patients and controls in the perceptual priors used. However, the precision of patient priors was inversely related to their LDE. Patients with high LDE showed more accurate priors, representing predictions that were closer to the true distribution of performance. Such accuracy has previously been demonstrated when observing the actions of others, suggesting abnormal awareness of action in these patients. These results confirm a link between dopamine and the positive expectation of the outcome of one's own actions, and may have implications for the management of PD.

Keywords: positive expectation, voluntary action, agency, Parkinson's disease, dopamine, Bayesian, placebo, inverted U-shaped function

Introduction

Parkinson's disease (PD) is a common neurodegenerative disease, associated with the loss of dopaminergic projections from the substantia nigra and ventral tegmentum to the striatum and frontal cortex respectively (reviewed in Cools, 2006). PD causes a disorder of movement with tremor, rigidity, and bradykinesia (Hughes et al., 1992). However, it also affects motor cognition, including executive function such as planning, sequencing and initiating movements (Owen et al., 1992; Williams-Gray et al., 2007a; Hughes et al., 2010, 2013). Dopamine replacement therapies alleviate some of these cognitive functions and their underlying neural circuits, but impair others (Gotham et al., 1988; Kehagia et al., 2010; Rowe et al., 2010).

Parkinson's disease can also change the perception and awareness of voluntary action. For example, the perception of the position and motion of one's own body parts, known as

'kinaesthesia,' is impaired by PD (Klockgether et al., 1995), possibly due to an abnormal processing of sensory feedback (Konczak et al., 2012). Levodopa alleviates kinaesthetic deficits in some studies (Li et al., 2010), but not in others (O'Suilleabhain et al., 2001).

The effect of PD on the perception of self-generated actions was also assessed by the intentional binding paradigm. Intentional binding refers to the perceived temporal attraction between a voluntary action and its sensory effect in instrumental behavior (Haggard et al., 2002). It has been used as an objective measure for the awareness of action and sense of agency (Wolpe and Rowe, 2014). This temporal attraction is unchanged in PD patients, but is enhanced by levodopa (Moore et al., 2010).

Together, these inconsistent results demonstrate a complex effect of PD and levodopa on the awareness of action, which might reflect a change in the expectations of outcomes rather than changes in sensation *per se*. We therefore studied the effect of PD and levodopa on the positive expectations of the outcome of goal-directed actions.

We recently reported that people's perception of the outcome of their own goal-directed action is normally biased toward goal success (Wolpe et al., 2014). The bias can be explained by optimistic Bayesian Self priors – that is, the exaggerated reliability of expectations of success set by the goal of the action, and integrated with the sensorimotor signals for perception. The priors are optimistic in the sense that they have a narrower distribution around the intended goal, relative to the actual distribution of performance. In contrast, for observed actions people use priors that more accurately represents the observed performance (Wolpe et al., 2014).

Optimistic expectations are normally integrated with sensorimotor signals that contribute to the awareness of action (Frith et al., 2000). For example, a cognitive process can 'exaggerate' the reliability of low-level sensorimotor prediction signals, such as the efference copy of the motor command (Wolpert and Ghahramani, 2000). The integration leads to narrow Self priors that support the correct attribution of actions by facilitating the difference between perception of one's own and others' actions (Wolpe et al., 2014).

Dopamine is a key neuromodulator that signals reward expectation in the striatum during operant action-reward conditioning (de la Fuente-Fernández et al., 2001). Striatal dopamine loss might therefore alter the positive expectation of outcomes of one's actions. Moreover, the representation of expected reward in medial frontal cortex is also dependent on dopamine (Rowe et al., 2008), and liable to hyper-dopaminergic effects in early PD (Rakshi et al., 1999) as well as dopamine loss in later stages.

The current study sought to examine the effect of dopamine on the perception of the outcome of goal-directed action. We tested the specific hypotheses that PD and levodopa modulate the perception of action by changing the exaggeration of expected performance (optimism). Patients in mild to moderate stages of PD participated, on their usual medication. We hypothesized that (i) PD would diminish the positive expectations for goal-directed actions, represented in reduced reliability of Self priors; and that (ii) the levodopa dose equivalents (LDE) of dopaminergic

medication would determine the reliability of these priors and alter the perception of one's goal-directed actions.

Materials and Methods

Participants

Twenty patients (13 men; aged 48–81 years mean: 68; SD: 10) were recruited from the John van Geest Centre for Brain Repair, PD research clinic. Patients met clinical diagnostic criteria of idiopathic PD, according to the UK PD brain bank criteria (Hughes et al., 1992), and were mild to moderate stages of disease, [Hoehn and Yahr stages 1–3] (Hoehn and Yahr, 1967). In addition, 20 age- and sex-matched, neurologically healthy controls, (12 men; aged 55–76 years, mean: 68; SD: 6) were included in the study, and were compensated with £12 for their participation. All subjects were right-handed, and gave written informed consent before the experiment. The study was approved by the Cambridge Research Ethics Committee.

Assessment of motor and cognitive symptoms in patients was performed at the beginning of the testing session. The severity of motor features was assessed with the Unified Parkinson's Disease Rating Scale (UPDRS) motor subscale III (Fahn and Elton, 1987). Cognitive abilities in both patients and controls were assessed through the Mini-Mental State Examination (MMSE; Folstein et al., 1975) and Addenbrooke's Cognitive Examination Revised (ACE-R; Mioshi et al., 2006). Cognitive impairment was an exclusion criterion: as ACE-R has better sensitivity and specificity for cognitive impairment in PD than MMSE (Reyes et al., 2009), we used the ACE-R cut-off score of 83/100 (Reyes et al., 2009).

All patients were on dopamine replacement therapy at the time of the experiment: patients were tested in the morning after taking their morning medication as normal. The time interval between levodopa self administration and testing varied between 1 and 3 h, such that patients were in a relative 'on' state. Our patients had mild to moderate PD (see **Table 1**) and were not affected by clinically significant on-off phenomena or freezing. LDE was computed according to Tomlinson et al. (2010).

Self Prior Task and Procedure

Subjects performed a modified version of the Self Stop task described in Wolpe et al. (2014) (**Figure 1A**). Subjects were seated 0.5 m from a 17′ Protouch LCD touch screen with 1024 × 768 resolution (26 pixels/cm) that refreshed at 60 Hz. All stimuli were displayed with Matlab Psychophysics toolbox (Brainard, 1997). In brief, subjects were asked to stop a blue ball (15 pixel radius), which repeatedly swept the screen horizontally in a rightward motion. When it was vertically aligned with a red circle target (15 pixel radius), subjects pressed a key with their right index finger to stop the ball. The ball vanished following the key press, and 250 ms later the target disappeared. The ball's starting position was randomized across trials (drawn from a uniform distribution covering the horizontal extent of the screen). The target was horizontally centered, and was displayed close to the top of the screen, just above the sweeping blue ball. The number of ball sweeps was limited to three, in order to control

TABLE 1 | Demographic details of Parkinson's disease (PD) patients participating in the study.

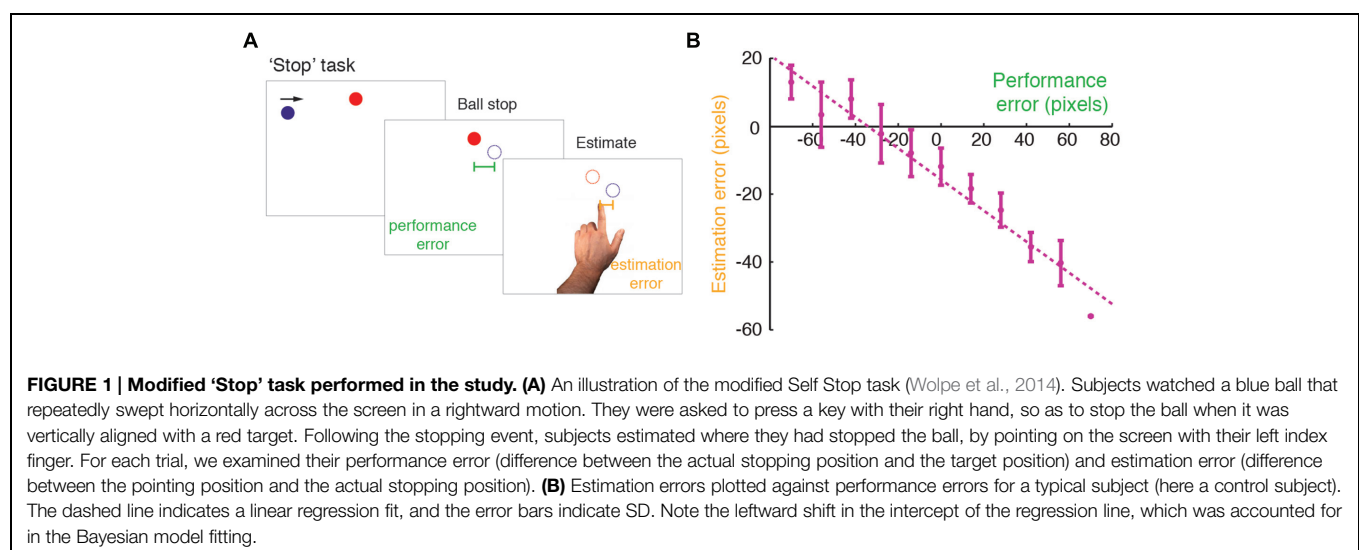| No.	Gender (M/F)	Age	Disease duration (years)	Disease stage* (0-5)	UPDRS-III motor subscale (0-56)	MMSE (0-30)	ACE-R (0-100)	LDE **
1	M	74	16	2	20	30	98	620
2	F	76	10	2	26	29	91	560
3	M	81	15	2.5	13	29	94	410
4	M	54	15	1	13	29	89	1817
4	F	73	13	2.5	16	28	95	565
6	F	76	9	3	37	30	91	855
7	F	81	13	2	15	29	97	1355
8	M	77	13	2	24	30	97	1722
9	M	64	6	1.5	21	29	94	1175
10	M	48	6	2.5	16	29	88	536
11	M	72	26	1	21	28	85	460
12	M	69	10	1	21	23	84	276
13	M	57	14	1	31	28	94	1180
14	F	66	17	2	19	27	95	1740
15	M	76	12	1	18	27	87	1500
16	F	64	11	3	16	28	96	1000
17	M	77	9	3	19	27	88	700
18	F	55	8	2	20	29	98	200
19	M	56	11	3	14	29	86	1210
20	M	63	11	1.5	16	29	94	1315
Mean		68	12	2	20	28	92	960

UPDRS, Unified Parkinson's Disease Rating Scale; MMSE, Mini-Mental State Examination; ACE-R, Addenbrooke's Cognitive Examination Revised; LDE, Levodopa Dose Equivalent; *According to Hoehn and Yahr (1967); **Calculated according to Tomlinson et al. (2010).

for stopping times across groups, thereby minimizing potential group differences in task demands. When failing to respond after three sweeps, the trial was stopped, and before it was restarted subjects received a message on the screen prompting them to respond faster.

Subjects were asked to point with their left index finger on the screen where the ball had stopped (i.e., final position before vanishing). Screen pointing was used after a pilot experiment showed that it was preferable to using a mouse for many patients. The left hand was chosen to facilitate fast responses, as the right

hand remained resting on the key in preparation for the next trial. The spatial shift in pointing due to hand use was corrected in the analysis by including an additional spatial shift parameter (see *Analyses*).

Subjects were also given the option to skip the trial during the estimation of the stopping point, by pointing at the word 'Skip' displayed at the bottom of the screen. Any trial that was skipped was excluded, and an additional trial was added instead, so that all subjects' dataset had an equal number of trials (on average four trials skipped per subject). The experiment was performed

FIGURE 1 | Modified 'Stop' task performed in the study. (A) An illustration of the modified Self Stop task (Wolpe et al., 2014). Subjects watched a blue ball that repeatedly swept horizontally across the screen in a rightward motion. They were asked to press a key with their right hand, so as to stop the ball when it was vertically aligned with a red target. Following the stopping event, subjects estimated where they had stopped the ball, by pointing on the screen with their left index finger. For each trial, we examined their performance error (difference between the actual stopping position and the target position) and estimation error (difference between the pointing position and the actual stopping position). **(B)** Estimation errors plotted against performance errors for a typical subject (here a control subject). The dashed line indicates a linear regression fit, and the error bars indicate SD. Note the leftward shift in the intercept of the regression line, which was accounted for in the Bayesian model fitting.

in blocks of trials. The experiment started with a short practice block of 12 trials to familiarize subjects with the task. Following practice, two blocks of 52 trials each were performed.

Analyses

As in Wolpe et al. (2014), for each trial we calculated the estimation error (distance between estimated stopping position and true stopping position) and the performance error (distance between true stopping position and target). For each subject, we first fitted a linear regression of estimation errors against performance errors to examine any bias in estimation. In order to have a consistency in the estimation procedure and to minimize a possible interference from different memory processes, we excluded outlier trials with longer estimation times in the following manner: trials with estimation times greater than 2 SDs from the mean were excluded (on average one trial for each control and four trials for each patient). One additional patient was excluded from the study cohort, as his mean estimation time was larger than 3 SDs from the patient group mean.

We inferred the subjects' priors by fitting Bayesian models through maximum likelihood estimation – that is, maximizing the probability of each subject's dataset. We used a constrained model, in which the mean of priors was centered on the target, and the mean of sensory evidence was centered on the true stopping position. However, as subjects used their left hand to estimate the stopping position, we expected there would be a consistent spatial (left) shift in estimates that should be accounted for in the model.

Such a spatial shift might be incorporated in the prior mean, in the visual evidence mean or in both. In the non-linear model fitting of the maximum likelihood estimation process, a model with a shift in the prior mean; a model with a shift in sensory evidence mean; and a model with a shift in both prior and sensory evidence, are all mathematically equivalent and will converge on the same parameters. However, a model that includes a shift in both prior and sensory evidence has an additional free parameter compared to only a shift in either. Moreover, we expected the spatial shift to arise due to the screen pointing process and independent of the prior. We therefore used a model with a shift in evidence mean (which in this experiment includes both sensory noise and noise added during the screen pointing) in addition to the constrained model above. We used the following equation derived in Wolpe et al. (2014):

$$x_{\text{estimate}} = w * \overline{x}_{\text{prior}} + (1 - w) * (\overline{x}_{\text{evidence}} + \text{shift}_{\text{evidence}})$$

with $\overline{x}_{\text{prior}}$ centered on the target, and the weighting w given by:

$$w = \frac{\sigma^2_{\text{evidence}}}{\sigma^2_{\text{prior}} + \sigma^2_{\text{evidence}}} \text{ (Ghahramani et al., 1997).}$$

To summarize, we fitted the following models:

1) Constrained model, with only prior SD and visual evidence SD as free parameters (2 parameters).
2) Prior SD and visual evidence SD as well as a spatial shift in visual evidence (3 parameters).

The model that best explained the data was selected using the Bayesian Information Criterion (BIC) with a threshold difference of 6 for 'strong' evidence (Schwarz, 1978; Raftery, 1995). Only parameters of the winning models were presented. For correlations with the clinical data and age, non-parametric Spearman's correlations were performed.

Results

The clinical details of patients are summarized in **Table 1**. Patients had a mixed severity of motor features and duration of illness, but they were all in mild to moderate stages of disease (Hoehn and Yahr, 1967).

Patients and controls were all able to perform the modified 'Stop' task (**Figure 1A**). Our first objective was to examine the difference in the perception of the consequences of one's own goal-directed actions between PD patients and controls. To this end, we compared the perceptual bias toward the target measured in the regression slopes, followed by a comparison of the Bayesian model parameters.

Before examining the perceptual bias and priors, we first compared the time it took subjects to complete the task as a possible confound. Patients and controls did not differ in the time required for stopping the ball ($t_{38} = -0.69$, $p = 0.49$). In contrast, there was a group difference in the time to estimate the ball stopping position ($t_{38} = 2.09$, $p = 0.043$). Mean Estimation time in patients was 1.07 s, but only 0.88 s in controls. We therefore included mean estimation times as a nuisance covariate in both the between-group and within-group analyses of the Bayesian priors below.

Perceptual Biases Toward the Target Across Groups

We examined subjects' bias toward the target, by fitting a linear regression of estimation errors against performance errors for each subject (**Figure 1B**). Mean intercept (i.e., ball stopping position) in patients was just six pixels left of the target, and 15 pixels left of the target in controls (see below). Estimation errors were always biased in a graded manner, as measured by a negative regression slope, for both patients (slope smaller than zero in patients: $t_{19} = -9.83$, $p < 0.001$) and controls ($t_{19} = -8.16$, $p < 0.001$). The extent of the bias, in terms of the regression slope, was not different between patients and controls ($t_{38} = -1.277$, $p = 0.419$; Bonferroni corrected).

Comparison of Bayesian Model Parameters Across Groups

We next fitted the data with two Bayesian models: (1) a model with SDs of sensory evidence and prior, constrained with the prior centered on the target and the evidence centered on the true stopping position; (2) a model with SDs of evidence and prior, and a spatial shift on evidence, likely to arise due to the estimation procedure of using the left hand (see **Figure 1B**). On a group-level, average difference in BIC was only 1.4 in favor of model 2. However, on an individual-subject level, model 2 was more likely than model 1 in 29/40 subjects (BIC difference greater than 6 in

12/20 patients and 17/20 controls), which was significant by a sign test (two-tailed; $p = 0.006$). We therefore report the parameters of model 2.

Examining the model parameter of shift from stopping position in the distribution of sensory evidence, subjects had a consistent leftward shift on their sensory evidence ($t_{38} = -4.8$, $p < 0.001$). This shift is likely to have resulted from using the left hand during the screen pointing procedure. Importantly, the shift did not differ across groups (mean shift in patients: 18 pixels; mean shift in controls: 23 pixels; $t_{38} = 0.95$, $p = 0.35$; uncorrected).

Examining the distribution of sensory evidence, there was a significant increase in the SDs of sensory evidence in patients compared to controls ($t_{38} = 3.14$, $p = 0.003$). This increase in patients could reflect an elevated sensory noise, or more noise attributed to the estimation procedure. Our main focus in this study, however, was on the differences in priors between patients and controls and within the patient group with relation to levodopa.

The SDs of priors correlated with performance error SDs for both patients ($r = 0.706$, $p < 0.001$) and controls ($r = 0.582$, $p = 0.007$) (**Figure 2A**), as in Wolpe et al. (2014). Based on this strong correlation, and as patient performance error SD was greater than that in controls ($t_{38} = 2.33$, $p = 0.025$), we examined priors with their values normalized to performance distribution (**Figure 2B**). These normalized priors reflect the degree of exaggeration of positive expectations for each subject. SDs of priors for both patients and controls were significantly smaller than SDs of performance distribution (normalized values smaller than 0.5 for patients: $t_{19} = -5.92$, $p < 0.001$; and controls: $t_{19} = -4.482$, $p < 0.001$). These results replicate the exaggerated Self priors found in Wolpe et al. (2014).

As patients and controls showed a significant difference in estimation time (see above), we included estimation time as a nuisance covariate in the comparison across groups in an analysis

of covariance. Normalized priors did not differ across groups (**Figure 2B**), with very small effect sizes: there was neither a main effect of group [$F_{(1,36)} = 0.063$, $p = 0.803$, $\eta^2 = 0.002$], or estimation time [$F_{(1,36)} = 0.081$, $p = 0.777$, $\eta^2 = 0.002$], nor a group $\times$ estimation time interaction [$F_{(1,36)} = 0.12$, $p = 0.731$, $\eta^2 = 0.003$]. An additional exploratory analysis showed no correlation between prior width and disease severity as measured by UPDRS motor subscale (Spearman's $\rho = -0.01$, ns). Together, these results suggest that PD patients demonstrated a normal group-average prior despite the likelihood of individual differences in disease and treatment. We next explored the source of within-patient variability, by examining the relation between priors and LDE.

Relation between Priors and Levodopa

The second objective of our study was to examine the relationship between the degree of exaggeration in priors and its relationship to levodopa doses in PD patients. Specifically, we examined the relation between individual differences in Self priors and LDEs (**Figure 3**). We found a strong correlation between normalized prior SDs and LDEs (Spearman's $\rho = 0.723$, $p = 0.002$), arising from a significant correlation between prior SDs and LDEs across patients (Spearman's $\rho = 0.585$, $p = 0.039$). The correlation was positive, indicating that patients with higher LDE had wider priors.

The correlation between normalized priors and LDEs remained significant even after accounting for disease duration (partial correlation; Spearman's $\rho = 0.723$, $p = 0.002$); severity of motor features measured using the UPDRS motor subscale III (partial correlation; Spearman's $\rho = 0.76$, $p < 0.001$); or subject mean estimation time (partial correlation; Spearman's $\rho = 0.746$, $p = 0.001$; with Bonferroni correction). For completeness, we note that there was no consistent relation between the SDs of sensory evidence and LDEs (Spearman's $\rho = 0.32$, $p = 0.17$). Taken together, these results suggest that the extent
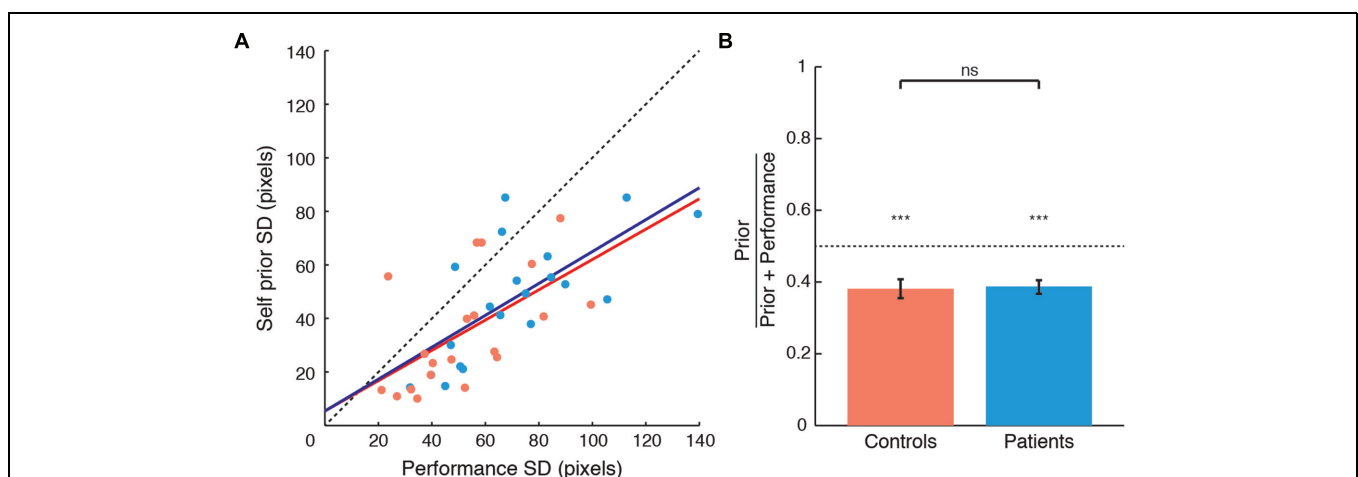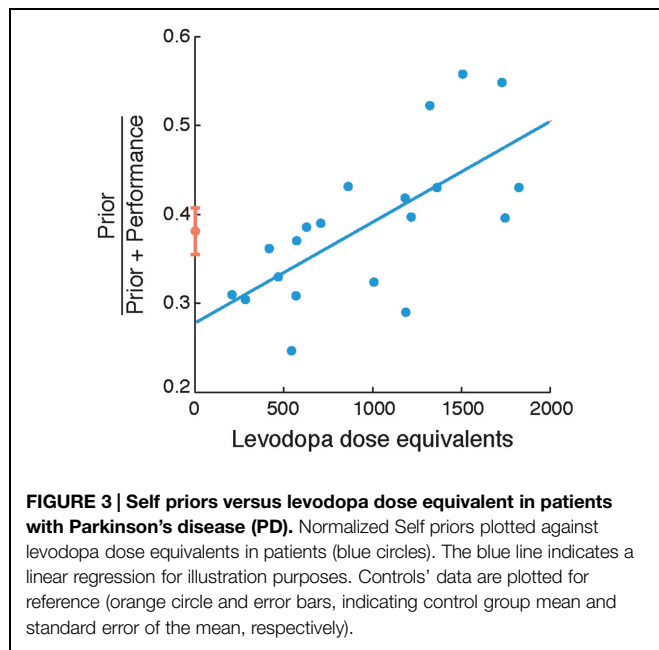

FIGURE 2 | Comparison between Self priors across groups. (A) SDs of Self priors plotted against SDs of performance errors for patients (blue circles and regression line) and controls (orange circles and regression line). Data points lie below the line of equality (black dashed line) in 34 of the 40 subjects. **(B)** Prior SDs normalized to performance error SDs, plotted for patients (blue) and controls (orange). For both groups, the SDs of priors were smaller than the SDs of performance errors. No difference was found across groups. Significance level indicated by ***$p < 0.001$.

FIGURE 3 | Self priors versus levodopa dose equivalent in patients with Parkinson's disease (PD). Normalized Self priors plotted against levodopa dose equivalents in patients (blue circles). The blue line indicates a linear regression for illustration purposes. Controls' data are plotted for reference (orange circle and error bars, indicating control group mean and standard error of the mean, respectively).

of exaggeration of the reliability of Self priors in patients was strongly related to levodopa doses.

The correlation between the extent of exaggeration of Self priors and levodopa in patients suggests that patients on higher levodopa had priors that were closer to the true performance distribution. However, the 'accuracy' of priors is not linearly related to the values of normalized priors, which can vary between 0 and 1, but would be 0.5 in the case when priors and performance error SDs are equal. To formally test how the accuracy of priors is related to levodopa, we therefore correlated the absolute difference between normalized priors and 0.5 with levodopa doses. A strongly significant negative correlation emerged (Spearman's $\rho = -0.717$, $p = 0.001$; Bonferroni corrected). This negative relation suggests that patients with high levodopa doses used priors that deviated less from the distribution of performance, and were thus more accurate, similar to the perception of observed actions (Wolpe et al., 2014).

Discussion

The principal results of this study are that (i) the group-mean of the positive distortion of perceived outcomes of one's own goal-directed action was similar in medicated PD patients and healthy individuals; but (ii) there is a strong correlation between dopaminergic dose and the degree of positive exaggeration of priors in patients. This correlation indicates that patients on higher levodopa doses more accurately perceive the outcome of their own actions, in a way that healthy people perceive the actions of others but not of themselves.

Group Comparison of Self Priors

Although Self priors were related to dopaminergic medication, there was no overall group difference. We suggest that this

reflects the dynamic effects of disease on basal ganglia function. For example, in early stages of PD, the dorsal regions of the basal ganglia that are implicated in habitual actions are most affected, while ventral regions that are most associated with goal-directed control are largely preserved (Redgrave et al., 2010). As a result, goal-directed behaviors gradually replace or compensate for the impairment of automated movements based on stimulus-response associations (Redgrave et al., 2010; Torres et al., 2011). As patients included in our study had mild to moderate disease (median Hoehn and Yahr stage two), we suggest that goal-directed actions and their perception remained relatively intact.

Another feature of abnormal motor control in PD is the increased reliance of patients on external cues. The significance of this feature is underscored in the common clinical observation of improved gait and postural stability in PD patients following sensory cueing (Azulay et al., 2002, 2006). Several studies have demonstrated the strong influence of visual feedback on patient movements (Flash et al., 1992; Klockgether and Dichgans, 1994; Schettino et al., 2006), which may partially compensate for kinaesthetic impairments (Azulay et al., 2006). When visual feedback is occluded, patients tend to be slower and less accurate than when it is available (Klockgether and Dichgans, 1994; Schettino et al., 2006).

Our visuomotor task could reinforce the tendency of patients to over-rely on visual cues. As it disappears just before the estimation phase of the task, the target potentially provides patients with a visual cue to help them complete the estimation task more quickly and accurately. Bias toward the target as a visual cue for the estimation procedure may offset the diminution of priors. This behavioral feature of PD is unlikely to account for the relationship between levodopa and the width of priors, as levodopa does not significantly affect the reliance on external cues for movement (Burleigh-Jacobs et al., 1997). In addition to the altered integration of visual signals in PD, impaired oculomotor function (e.g., Shibasaki et al., 1979; Perneczky et al., 2011) could influence performance on our task. However, the absence of a group difference in our task and the minimal effect of levodopa on oculomotor function (Corin et al., 1972) suggest that the effect of oculomotor dysfunction on our principal results is minimal.

Although we have focussed on the degree of exaggeration of sensorimotor prediction for the perception of action, it is worth noting that previous studies have suggested preservation of other sensorimotor prediction signals in PD. For example, intentional binding, the perceived temporal attraction between a voluntary action and its sensory effect, is dependent on an intact sensorimotor prediction (Waszak et al., 2012; Wolpe et al., 2013; Wolpe and Rowe, 2014), but is unaffected in PD (Moore et al., 2010). Moreover, kinaesthetic deficits in PD have been found to be driven by abnormal low-level sensory processing, rather than changes in prediction processes (Konczak et al., 2012). Similarly, we found increased noise on sensory evidence in patients, which can be attributed to abnormal processing of sensory input or to a greater noise in the motor output during the estimation process.

A second potential contributor to the increased sensory noise shown by patients is the time it took patients to complete the procedure. On average, patients were 200 ms slower than

controls in estimating the sensory consequences of their actions, although they were not slower in their response time for stopping the ball. This result is perhaps not surprising, considering the bradykinesia that is characteristic of PD, but simple motor response times are not necessarily increased in PD and the additional estimation time may indicate an impairment in the cognitive decision processes. The estimation of time or temporal intervals is also affected by PD (Pastor et al., 1992; Buhusi and Meck, 2005). This might have been expected to influence the perception of velocity of the moving ball, or the horizontal distance traveled during, although neither the simple regression models nor Bayesian models to estimate the Self priors suggested a group-wise difference even when accounting for the different estimation times. However, PD is a heterogeneous disorder and in the next section we consider the causes of individual differences in priors.

A Relation between Self Priors and Levodopa Dose

The marked individual differences in our patients (see **Figure 3**) is typical of studies of heterogeneous neurodegenerative disorders like PD. This variation is in part due to individual differences in the extent and severity of striatal dopaminergic denervation and cortical involvement (e.g., Cools, 2006; Rowe et al., 2008). Much of the variability in our patients' priors was explained by LDE (Spearman's ρ 0.723), which positively correlated with the degree of exaggeration in priors. Patients with higher levodopa doses demonstrated more accurate Self priors that are more similar to the priors used for observing others' actions (Wolpe et al., 2014).

We suggest that exogenous dopamine in PD (indexed by the LDE) is related to positive expectations: patients under higher levodopa doses do not show the normal exaggeration of Self priors. However, the effect of dopamine on fronto-striatal functions in PD is often described as an inverted U-shaped curve, such that level of baseline dopaminergic function determines whether dopaminergic medication enhances or impairs cognitive functions (Rowe et al., 2008; Wu et al., 2012; Hughes et al., 2013). These baseline levels are in part determined by the differences in the integrity of dopaminergic innervation in the parallel corticostriatal circuits for motor, oculomotor, limbic and cognitive function. There are also individual differences due to polymorphism in the genotype of Catechol O-Methyltransferase enzyme (Williams-Gray et al., 2007b; Nombela et al., 2014). In our study, the correlation between the width of Self priors and LDE was centered on the mean of controls (no group difference), confirming that low levodopa doses have the opposite effect to high doses in PD patients, a result of inverted U-shape dose-response relationships (Cools and D'Esposito, 2011).

This U-shaped relation suggests that low doses of levodopa might preserve, or even further enhance the precision of priors, leading to increased optimistic expectations. It has been shown that a single administration of levodopa in young healthy adults can enhance hedonic expectations for future events (Sharot et al., 2009). Subjects gave higher ratings of prospective 'happiness' to different vacation destinations when given levodopa compared

to placebo. The levodopa dose used by Sharot et al. (2009) was 100 mg, which according to our data might lead to such narrower priors (see Figure 5.3A) and the related enhancement in optimism bias. Taken together, these results confirm the dose-dependent effect of dopaminergic medication on goal priors and the resulting optimism bias. It also means that levodopa administration in young healthy adults cannot be equated to older neurodegenerative patients in which the dopaminergic systems are severely perturbed at baseline.

As the narrow Self priors might support the normal attribution of action (Wolpe et al., 2014), these results imply that PD patients on high levodopa doses have an impaired sense of agency. Interestingly, levodopa treatment in PD has indirectly been shown to alter the sense of agency, increasing overall intentional binding of action and its effect (Moore et al., 2010). Together, these findings not only suggest an impaired awareness of action in medicated PD patients, but also support the application of Self priors as an objective measure of agency processes (Wolpe and Rowe, 2014).

The Mechanism of Action of Dopamine on Self Priors

Both cognitive and sensorimotor processes are required for generating predictions for the perception of one's goal-directed action (Wolpe et al., 2014). For example, there is a prediction of the ensuing sensory effect from the efference copy of the motor command using a forward model; a process which is optimized by learning the relation between an action and its sensory effect (Wolpert and Ghahramani, 2000). These predictions might be adjusted according to prior knowledge about the world, for example the distribution of one's motor performance (Körding and Wolpert, 2004). The reliability of this low-level sensorimotor prediction can be 'exaggerated' by top–down cognitive mechanisms, such as conscious expectations (Sterzer et al., 2008), motivational states (Balcetis and Dunning, 2006) and the illusions of superiority (Wolpe et al., 2014).

The role of striatal dopamine in such prediction processes has been established in health (Pessiglione et al., 2006). In medicated PD patients, the relation between the degree of exaggeration measured in Self priors and dopamine could stem from either the aberrant endogenous dopamine release, or the influx of exogenous dopamine from the levodopa treatment. We next discuss these alternative mechanisms.

Dopamine release in the nigrostriatal system in PD has been linked to the strength of placebo effect – that is, the clinical improvement following a non-active treatment due to the positive expectation of benefit (de la Fuente-Fernández et al., 2001). Increased nigrostriatal damage may lead to diminished positive expectations, thereby reducing the placebo effect in PD (de la Fuente-Fernández et al., 2001). This link can also explain the relation between Self priors and dopamine in our study: as levodopa doses are tailored for each patient's motor impairments, the computed LDEs may, at least in part, be regarded as a proxy for nigrostriatal integrity. Patients on higher levodopa doses, suggesting greater nigrostriatal damage, would thus show impairment in the normal positive expectations reflected in

the narrow Self priors, resulting in a more 'accurate' and less optimistic perception.

However, levodopa therapies can have detrimental effects on patient cognition (Cools, 2006), due to the uneven distribution of striatal and cortical dopaminergic pathology in PD. Dorsal striatum is most affected early, while ventral striatum and mesocortex are relatively intact in early disease stages (reviewed in Kish et al., 1988; Kehagia et al., 2010). As patients' levodopa doses are usually adjusted according to the motor deficits (closely related to the dorsal striatum), they can cause a dopaminergic overdose of the ventral striatum system (Cools, 2006). Areas connected to the ventral striatum and the mescortical system, including anterior cingulate and ventromedial cortex, are hyperdopaminergic in early medicated PD (Rakshi et al., 1999; Wu et al., 2012). The anterior cingulate and ventromedial cortex have been implicated in the illusions of superiority (Beer and Hughes, 2010) and optimism bias (Sharot et al., 2007), which can be explained in terms of exaggerated Self priors (Wolpe et al., 2014). In our study, increasing doses of levodopa could have thus impaired the exaggeration of reliability of sensorimotor prediction by overdosing the mesocortical pathway.

Implications for PD

In healthy adults, the degree of exaggeration in Self priors has been shown to be associated with trait optimism bias: people who show wider priors tend to be less optimistic (Wolpe et al., 2014). The exaggerated priors and the related 'positive' cognitive illusions could thereby facilitate motivation and adaptive behavior (Taylor and Brown, 1988). In PD patients, however, increasing damage to nigrostriatal circuits and/or long-term dopaminergic overdose of the mesocortex impair the exaggeration of the reliability of priors, leading to a more accurate but pessimistic perception of one's actions. The persistence of this pessimistic and accurate perception might lead to a reduced motivation as a result of a 'depressive realism' (Alloy and Abramson, 1988). The abnormally accurate perception of the consequences of ones actions does not imply accurate choices in the selection of actions. Indeed, dopaminergic dysregulation in PD alters value based decision making (Voon et al., 2010), which predisposes to impulsive and addictive behaviors (Napier et al., 2014).

A more accurate perception of the outcomes of one's actions might put patients at risk for depression and poor motivation, which are indeed often observed in PD (e.g., Gotham et al., 1986); and which contribute to the reduced quality of life in patients (Schrag et al., 2000). Similarly, broadening of Self priors might reflect diminished positive expectations, including reduced expectation of benefit from treatment (de la Fuente-Fernández et al., 2001), impairing the alleviation of symptoms.

The administration of dopamine in PD can *elevate* mood and reduce anxiety (Maricle et al., 1995; Czernecki et al., 2002). Critically, the reported improvement in depressive symptoms in such studies was the result of a short-term levodopa administration following overnight withdrawal and were measured using self-reports and questionnaires (Maricle et al., 1995; Czernecki et al., 2002). However, the effect of dopamine on patients withdrawn from chronic dopaminergic treatment could be due to the alleviation of acute withdrawal effects, rather than indicating a long-term role in behavior (Nestler, 2001). In contrast, our study measured sensorimotor and perceptual processes while the patients were on their usual medication.

The current study has several limitations. Firstly, we do not separate our analysis according to the laterality of dominant motor symptoms. This varied across patients, and could affect the motor performance on the visuomotor task, and may add unexplained variability to the correlation analyses. However, in the absence of *a priori* hypotheses of lateralised processes, small numbers and the presence of a systemic therapy, we suggest that collapsing across the factor of laterality is preferable. In addition to altered control of the hands which may affect performance in the task, their impaired oculomotor function (Corin et al., 1972; Shibasaki et al., 1979) may have also compromised patient performance of the visuomotor task.

In light of the demands of the visuomotor task and to facilitate compliance, patients were only tested whilst on their usual dopaminergic medication. We did not measure their perception of action 'off' medication. Moreover, within each patient we did not test across drug phases of their on–off cycle, i.e., in different times relative to levodopa administration. We could not therefore dissociate the possible mechanisms through which dopamine modulates Self priors in PD; and could not establish a causal effect of dopamine on Self priors as our results are correlational.

Conclusion

For the perception of one's own goal-directed actions, the use of exaggerated priors persists in PD. However, the dopaminergic dose was negatively correlated with the degree of exaggeration of priors, such that patients on more dopaminergic medication had more accurate priors that closely represented the true distribution of performance. This veridical perception is normally only present for the observation of others' actions. Our results suggest that positive expectations can be increasingly impaired in PD, especially with a relative dopaminergic overdose of the ventral striatum and mesocortical systems. The resulting changes in perception of action outcomes have implications for understanding the normative mechanisms, and the risks of affective and behavioral symptoms in PD.

Acknowledgments

References

Alloy, L. B., and Abramson, L. Y. (1988). "Depressive realism: four theoretical perspectives," in *Cognitive Processes in Depression*, ed. L. B. Alloy (New York, NY: Guilford Press), 223–265.

Azulay, J.-P., Mesure, S., Amblard, B., and Pouget, J. (2002). Increased visual dependence in Parkinson's disease. *Percept. Mot. Skills* 95, 1106–1114. doi: 10.2466/pms.2002.95.3f.1106

Azulay, J.-P., Mesure, S., and Blin, O. (2006). Influence of visual cues on gait in Parkinson's disease: contribution to attention or sensory dependence? *J. Neurol. Sci.* 248, 192–195. doi: 10.1016/j.jns.2006.05.008

Balcetis, E., and Dunning, D. (2006). See what you want to see: motivational influences on visual perception. *J. Pers. Soc. Psychol.* 91, 612–625. doi: 10.1037/0022-3514.91.4.612

Beer, J. S., and Hughes, B. L. (2010). Neural systems of social comparison and the "above-average" effect. *Neuroimage* 49, 2671–2679. doi: 10.1016/j.neuroimage.2009.10.075

Brainard, D. H. (1997). The psychophysics toolbox. *Spat. Vis.* 10, 433–436. doi: 10.1163/156856897X00357

Buhusi, C. V., and Meck, W. H. (2005). What makes us tick? Functional and neural mechanisms of interval timing. *Nat. Rev. Neurosci.* 6, 755–765. doi: 10.1038/nrn1764

Burleigh-Jacobs, A., Horak, F. B., Nutt, J. G., and Obeso, J. A. (1997). Step initiation in Parkinson's disease: influence of levodopa and external sensory triggers. *Mov. Disord.* 12, 206–215. doi: 10.1002/mds.870120211

Cools, R. (2006). Dopaminergic modulation of cognitive function-implications for L-DOPA treatment in Parkinson's disease. *Neurosci. Biobehav. Rev.* 30, 1–23. doi: 10.1016/j.neubiorev.2005.03.024

Cools, R., and D'Esposito, M. (2011). Inverted-U–shaped dopamine actions on human working memory and cognitive control. *Biol. Psychiatry* 69, e113–e125. doi: 10.1016/j.biopsych.2011.03.028

Corin, M. S., Elizan, T. S., and Bender, M. B. (1972). Oculomotor function in patients with Parkinson's disease. *J. Neurol. Sci.* 15, 251–265. doi: 10.1016/0022-510X(72)90068-8

Czernecki, V., Pillon, B., Houeto, J. L., Pochon, J. B., Levy, R., and Dubois, B. (2002). Motivation, reward, and Parkinson's disease: influence of dopatherapy. *Neuropsychologia* 40, 2257–2267. doi: 10.1016/S0028-3932(02)00108-2

de la Fuente-Fernández, R., Ruth, T. J., Sossi, V., Schulzer, M., Calne, D. B., and Stoessl, A. J. (2001). Expectation and dopamine release: mechanism of the placebo effect in Parkinson's disease. *Science* 293, 1164–1166. doi: 10.1126/science.1060937

Fahn, S., and Elton, R. (1987). "UPDRS program members. Unified Parkinsons disease rating scale," in *Recent Developments in Parkinsons Disease*, eds S. Fahn, C. Marsden, M. Goldstein, and D. Calne (Florham Park, NJ: Macmillan Healthcare Information), 153–163.

Flash, T., Inzelberg, R., Schechtman, E., and Korczyn, A. D. (1992). Kinematic analysis of upper limb trajectories in Parkinson's disease. *Exp. Neurol.* 118, 215–226. doi: 10.1016/0014-4886(92)90038-R

Folstein, M. F., Folstein, S. E., and McHugh, P. R. (1975). "Mini-mental state." A practical method for grading the cognitive state of patients for the clinician. *J. Psychiatr. Res.* 12, 189–198.

Frith, C. D., Blakemore, S. J., and Wolpert, D. M. (2000). Abnormalities in the awareness and control of action. *Philos. Trans. R. Soc. B Biol. Sci.* 355, 1771–1788. doi: 10.1098/rstb.2000.0734

Ghahramani, Z., Wolpert, D. M., and Jordan, M. I. (1997). "Computational models of sensorimotor integration," in *Advances in Psychology,* eds P. G. Morasso and V. Sanguineti (Amsterdam: Elsevier), 117–147.

Gotham, A. M., Brown, R. G., and Marsden, C. D. (1986). Depression in Parkinson's disease: a quantitative and qualitative analysis. *J. Neurol. Neurosurg. Psychiatry* 49, 381–389. doi: 10.1136/jnnp.49.4.381

Gotham, A. M., Brown, R. G., and Marsden, C. D. (1988). "Frontal" cognitive function in patients with Parkinson's disease "on" and "off" levodopa. *Brain* 111, 299–321. doi: 10.1093/brain/111.2.299

Haggard, P., Clark, S., and Kalogeras, J. (2002). Voluntary action and conscious awareness. *Nat. Neurosci.* 5, 382–385. doi: 10.1038/nn827

Hoehn, M. M., and Yahr, M. D. (1967). Parkinsonism: onset, progression and mortality. *Neurology* 17, 427–442. doi: 10.1212/WNL.17.5.427

Hughes, A. J., Daniel, S. E., Kilford, L., and Lees, A. J. (1992). Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases. *J. Neurol. Neurosurg. Psychiatry* 55, 181–184. doi: 10.1136/jnnp.55.3.181

Hughes, L. E., Altena, E., Barker, R. A., and Rowe, J. B. (2013). Perseveration and choice in Parkinson's disease: the impact of progressive frontostriatal dysfunction on action decisions. *Cereb. Cortex* 23, 1572–1581. doi: 10.1093/cercor/bhs144

Hughes, L. E., Barker, R. A., Owen, A. M., and Rowe, J. B. (2010). Parkinson's disease and healthy aging: independent and interacting effects on action selection. *Hum. Brain Mapp.* 31, 1886–1899. doi: 10.1002/hbm.20979

Kehagia, A. A., Barker, R. A., and Robbins, T. W. (2010). Neuropsychological and clinical heterogeneity of cognitive impairment and dementia in patients with Parkinson's disease. *Lancet Neurol.* 9, 1200–1213. doi: 10.1016/S1474-4422(10)70212-X

Kish, S. J., Shannak, K., and Hornykiewicz, O. (1988). Uneven pattern of dopamine loss in the striatum of patients with idiopathic Parkinson's disease. Pathophysiologic and clinical implications. *N. Engl. J. Med.* 318, 876–880. doi: 10.1056/NEJM198804073181402

Klockgether, T., Borutta, M., Rapp, H., Spieker, S., and Dichgans, J. (1995). A defect of kinesthesia in Parkinson's disease. *Mov. Disord.* 10, 460–465. doi: 10.1002/mds.870100410

Klockgether, T., and Dichgans, J. (1994). Visual control of arm movement in Parkinson's disease. *Mov. Disord.* 9, 48–56. doi: 10.1002/mds.870090108

Konczak, J., Sciutti, A., Avanzino, L., Squeri, V., Gori, M., Masia, L., et al. (2012). Parkinson's disease accelerates age-related decline in haptic perception by altering somatosensory integration. *Brain* 135, 3371–3379. doi: 10.1093/brain/aws265

Körding, K. P., and Wolpert, D. M. (2004). Bayesian integration in sensorimotor learning. *Nature* 427, 244–247. doi: 10.1038/nature02169

Li, K., Pickett, K., Nestrasil, I., Tuite, P., and Konczak, J. (2010). The effect of dopamine replacement therapy on haptic sensitivity in Parkinson's disease. *J. Neurol.* 257, 1992–1998. doi: 10.1007/s00415-010-5646-9

Maricle, R. A., Nutt, J. G., Valentine, R. J., and Carter, J. H. (1995). Dose-response relationship of levodopa with mood and anxiety in fluctuating Parkinson's disease: a double-blind, placebo-controlled study. *Neurology* 45, 1757–1760. doi: 10.1212/WNL.45.9.1757

Mioshi, E., Dawson, K., Mitchell, J., Arnold, R., and Hodges, J. R. (2006). The addenbrooke's cognitive examination revised (ACE-R): a brief cognitive test battery for dementia screening. *Int. J. Geriatr. Psychiatry* 21, 1078–1085. doi: 10.1002/gps.1610

Moore, J. W., Schneider, S. A., Schwingenschuh, P., Moretto, G., Bhatia, K. P., and Haggard, P. (2010). Dopaminergic medication boosts action–effect binding in Parkinson's disease. *Neuropsychologia* 48, 1125–1132. doi: 10.1016/j.neuropsychologia.2009.12.014

Napier, T. C., Corvol, J.-C., Grace, A. A., Roitman, J. D., Rowe, J., Voon, V., et al. (2014). Linking neuroscience with modern concepts of impulse control disorders in Parkinson's disease. *Mov. Disord.* 30, 1–9. doi: 10.1002/mds.26068

Nestler, E. J. (2001). Molecular basis of long-term plasticity underlying addiction. *Nat. Rev. Neurosci.* 2, 119–128. doi: 10.1038/35053570

Nombela, C., Rowe, J. B., Winder-Rhodes, S. E., Hampshire, A., Owen, A. M., Breen, D. P., et al. (2014). Genetic impact on cognition and brain function in newly diagnosed Parkinson's disease: ICICLE-PD study. *Brain* 137, 2743–2758. doi: 10.1093/brain/awu201

O'Suilleabhain, P., Bullard, J., and Dewey, R. B. (2001). Proprioception in Parkinson's disease is acutely depressed by dopaminergic medications. *J. Neurol. Neurosurg. Psychiatry* 71, 607–610. doi: 10.1136/jnnp.71.5.607

Owen, A. M., James, M., Leigh, P. N., Summers, B. A., Marsden, C. D., Quinn, N. P., et al. (1992). Fronto-striatal cognitive deficits at different stage of Parkinson's disease. *Brain* 115, 1727–1751. doi: 10.1093/brain/115.6.1727

Pastor, M. A., Artieda, J., Jahanshahi, M., and Obeso, J. A. (1992). Time estimation and reproduction is abnormal in Parkinson's disease. *Brain* 115(Pt 1), 211–225. doi: 10.1093/brain/115.1.211

Perneczky, R., Ghosh, B. C. P., Hughes, L., Carpenter, R. H. S., Barker, R. A., and Rowe, J. B. (2011). Saccadic latency in Parkinson's disease correlates with executive function and brain atrophy, but not motor severity. *Neurobiol. Dis.* 43, 79–85. doi: 10.1016/j.nbd.2011.01.032

Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J., and Frith, C. D. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature* 442, 1042–1045. doi: 10.1038/nature05051

Raftery, A. E. (1995). Bayesian model selection in social research. *Sociol. Methodol.* 25, 111–163. doi: 10.2307/271063

Rakshi, J. S., Uema, T., Ito, K., Bailey, D. L., Morrish, P. K., Ashburner, J., et al. (1999). Frontal, midbrain and striatal dopaminergic function in early and advanced Parkinson's disease A 3D [(18)F]dopa-PET study. *Brain* 122, 1637–1650. doi: 10.1093/brain/122.9.1637

Redgrave, P., Rodriguez, M., Smith, Y., Rodriguez-Oroz, M. C., Lehericy, S., Bergman, H., et al. (2010). Goal-directed and habitual control in the basal ganglia: implications for Parkinson's disease. *Nat. Rev. Neurosci.* 11, 760–772. doi: 10.1038/nrn2915

Reyes, M. A., Lloret, S. P., Gerscovich, E. R., Martin, M. E., Leiguarda, R., and Merello, M. (2009). Addenbrooke's cognitive examination validation in Parkinson's disease. *Eur. J. Neurol.* 16, 142–147. doi: 10.1111/j.1468-1331.2008.02384.x

Rowe, J. B., Hughes, L. E., Barker, R. A., and Owen, A. M. (2010). Dynamic causal modelling of effective connectivity from fMRI: are results reproducible and sensitive to Parkinson's disease and its treatment? *Neuroimage* 52, 1015–1026. doi: 10.1016/j.neuroimage.2009.12.080

Rowe, J. B., Hughes, L., Ghosh, B. C. P., Eckstein, D., Williams-Gray, C. H., Fallon, S., et al. (2008). Parkinson's disease and dopaminergic therapy-differential effects on movement, reward and cognition. *Brain* 131, 2094–2105. doi: 10.1093/brain/awn112

Schettino, L. F., Adamovich, S. V., Hening, W., Tunik, E., Sage, J., and Poizner, H. (2006). Hand preshaping in Parkinson's disease: effects of visual feedback and medication state. *Exp. Brain Res.* 168, 186–202. doi: 10.1007/s00221-005-0080-4

Schrag, A., Jahanshahi, M., and Quinn, N. (2000). What contributes to quality of life in patients with Parkinson's disease? *J. Neurol. Neurosurg. Psychiatry* 69, 308–312. doi: 10.1136/jnnp.69.3.308

Schwarz, G. (1978). Estimating the dimension of a model. *Ann. Stat.* 6, 461–464. doi: 10.2307/2958889

Sharot, T., Riccardi, A. M., Raio, C. M., and Phelps, E. A. (2007). Neural mechanisms mediating optimism bias. *Nature* 450, 102–105. doi: 10.1038/nature06280

Sharot, T., Shiner, T., Brown, A. C., Fan, J., and Dolan, R. J. (2009). Dopamine enhances expectation of pleasure in humans. *Curr. Biol.* 19, 2077–2080. doi: 10.1016/j.cub.2009.10.025

Shibasaki, H., Tsuji, S., and Kuroiwa, Y. (1979). Oculomotor abnormalities in Parkinson's disease. *Arch. Neurol.* 36, 360–364. doi: 10.1001/archneur.1979.00500420070009

Sterzer, P., Frith, C., and Petrovic, P. (2008). Believing is seeing: expectations alter visual awareness. *Curr. Biol.* 18, R697–R698. doi: 10.1016/j.cub.2008.06.021

Taylor, S. E., and Brown, J. D. (1988). Illusion and well-being: a social psychological perspective on mental health. *Psychol. Bull.* 103, 193–210. doi: 10.1037/0033-2909.103.2.193

Tomlinson, C. L., Stowe, R., Patel, S., Rick, C., Gray, R., and Clarke, C. E. (2010). Systematic review of levodopa dose equivalency reporting in Parkinson's disease. *Mov. Disord.* 25, 2649–2653. doi: 10.1002/mds.23429

Torres, E. B., Heilman, K. M., and Poizner, H. (2011). Impaired endogenously evoked automated reaching in Parkinson's disease. *J. Neurosci.* 31, 17848–17863. doi: 10.1523/JNEUROSCI.1150-11.2011

Voon, V., Reynolds, B., Brezing, C., Gallea, C., Skaljic, M., Ekanayake, V., et al. (2010). Impulsive choice and response in dopamine agonist-related impulse control behaviors. *Psychopharmacology (Berl.)* 207, 645–659. doi: 10.1007/s00213-009-1697-y

Waszak, F., Cardoso-Leite, P., and Hughes, G. (2012). Action effect anticipation: neurophysiological basis and functional consequences. *Neurosci. Biobehav. Rev.* 36, 943–959. doi: 10.1016/j.neubiorev.2011.11.004

Williams-Gray, C. H., Foltynie, T., Brayne, C. E. G., Robbins, T. W., and Barker, R. A. (2007a). Evolution of cognitive dysfunction in an incident Parkinson's disease cohort. *Brain* 130, 1787–1798. doi: 10.1093/brain/awm111

Williams-Gray, C. H., Hampshire, A., Robbins, T. W., Owen, A. M., and Barker, R. A. (2007b). Catechol O-methyltransferase Val158Met genotype influences frontoparietal activity during planning in patients with Parkinson's disease. *J. Neurosci.* 27, 4832–4838. doi: 10.1523/JNEUROSCI.0774-07.2007

Wolpe, N., Haggard, P., Siebner, H. R., and Rowe, J. B. (2013). Cue integration and the perception of action in intentional binding. *Exp. Brain Res.* 229, 467–474. doi: 10.1007/s00221-013-3419-2

Wolpe, N., and Rowe, J. B. (2014). Beyond the "urge to move": objective measures for the study of agency in the post-Libet era. *Front. Hum. Neurosci.* 8:450. doi: 10.3389/fnhum.2014.00450

Wolpe, N., Wolpert, D. M., and Rowe, J. B. (2014). Seeing what you want to see: priors for one's own actions represent exaggerated expectations of success. *Front. Behav. Neurosci.* 8:232. doi: 10.3389/fnbeh.2014.00232

Wolpert, D. M., and Ghahramani, Z. (2000). Computational principles of movement neuroscience. *Nat. Neurosci.* 3, 1212–1217. doi: 10.1038/81497

Wu, K., O'Keeffe, D., Politis, M., O'Keeffe, G. C., Robbins, T. W., Bose, S. K., et al. (2012). The catechol-O-methyltransferase Val(158)Met polymorphism modulates fronto-cortical dopamine turnover in early Parkinson's disease: a PET study. *Brain* 135, 2449–2457. doi: 10.1093/brain/aws157

Imitation as a mechanism in cognitive development: a cross-cultural investigation of 4-year-old children's rule learning

Zhidan Wang[1]*, Rebecca A. Williamson[1] and Andrew N. Meltzoff[2]

[1] Department of Psychology, Georgia State University, Atlanta, GA, USA, [2] Institute for Learning and Brain Sciences, University of Washington, Seattle, WA, USA

Children learn about the social and physical world by observing other people's acts. This experiment tests both Chinese and American children's learning of a rule. For theoretical reasons we chose the rule of categorizing objects by the weight. Children, age 4 years, saw an adult heft four visually-identical objects and sort them into two bins based on an invisible property—the object's weight. Children who saw this categorization behavior were more likely to sort those objects by weight than were children who saw control actions using the same objects and the same bins. Crucially, children also generalized to a novel set of objects with no further demonstration, suggesting rule learning. We also report that high-fidelity imitation of the adult's "hefting" acts may give children crucial experience with the objects' weights, which could then be used to infer the more abstract rule. The connection of perception, action, and cognition was found in children from both cultures, which leads to broad implications for how the imitation of adults' acts functions as a lever in cognitive development.

Keywords: imitation, rule learning, weight, categorization, cross-culture, social learning

Introduction

The ability to learn from others' actions sets our species apart. Human infants and toddlers have a proclivity, rare in the animal kingdom, for imitating a broad range of acts (Meltzoff et al., 2009; Whiten et al., 2009). This includes reproducing not only the overall outcome or endstates that others achieve with objects, but also the precise means used to attain them. For example, after witnessing the novel act of an adult touching a light panel with his head to illuminate it, 18-month-olds are likely to perform this novel act even after a 1-week delay (Meltzoff, 1988). The neural basis for infant and childhood imitation is being uncovered using electroencephalography (EEG; Marshall and Meltzoff, 2014).

Imitation has several advantages for cognitive development. Reproducing others' precise actions accelerates and supports cultural learning of instrumental actions and arbitrary rituals (Tomasello, 1999; Boyd and Richerson, 2005; Meltzoff et al., 2009; Herrmann et al., 2013). Instrumental innovations and social routines can spread through communities through imitation, thereby leading these behaviors to be maintained across generations and providing more opportunities for cumulative progress.

A particular benefit of high-fidelity imitation is that it increases learning opportunities (Williamson and Markman, 2006). Even if acts are not fully understood, children who are able to imitate them in precise detail gain opportunities to discover a deeper meaning and cognitive

understanding of the acts, which are first grasped only in a more superficial manner. In this paper, we hypothesize that action imitation can spark cognitive change and test this idea with a novel procedure using the categorization of objects by their weight. We conducted these tests in two cultures, China and the USA.

What children learn from others' actions is not limited to specific observable movements. Children also infer and reproduce the goals others strive to achieve and cognitive rules that guide others' behaviors (for review, see Meltzoff and Williamson, 2013). For example, children imitate an adult's intended goal (e.g., Meltzoff, 1995), causal relations (Horner and Whiten, 2005; Schulz et al., 2008; Buchsbaum et al., 2011; Waismeyer et al., 2015), the organization guiding others' acts (Whiten et al., 2006; Flynn and Whiten, 2008; Loucks and Meltzoff, 2013), and abstract rules (Subiaul et al., 2007a,b, 2014; Williamson et al., 2010; Wang et al., 2015).

Evidence for what has been dubbed "abstract imitation" comes from Williamson et al. (2010), which is the basis for the current experiment. Children in that study saw an adult sort four objects into two bins according to either a visual property, color (Experiment 1), or the sounds the objects produced when shaken (Experiment 2). When given a chance to manipulate the objects, children in the experimental groups were more likely to categorize the objects by these respective properties than were controls. The children were then presented with a generalization task—a different set of objects that differed from the originals in kind as well as in their color or the sound they produced. Although the adult never manipulated this second set, the children in the experimental group sorted these objects by the key object property (color or sound), suggesting that children learned an abstract rule that could be generalized across stimuli.

Here we extended this idea of "abstract imitation" to children's learning about an interesting domain in physics—object weight. Categorizing by weight is a cognitively demanding task for preschool children. Results from Wang et al. (2015) show that 36-month-old children, the same age that readily learned to sort objects by their colors and sounds, were unable to learn the weight-sorting rule through observation and imitation. This finding is in line with previous research establishing that preschool-aged children struggle on tasks that require considering weights independently of object appearances (Smith et al., 1985; Schrauf et al., 2011).

Cross-cultural methods have been used to assess which aspects of social learning are culturally universal and which vary. Overall, these studies have shown substantial similarity in children's early imitation, despite considerable differences in cultural milieu (Callaghan et al., 2011; Wang et al., 2012). For example, highly similar reactions have been demonstrated in children from an industrialized Australian city and children from remote Bushman and Aborigine communities (Nielsen and Tomaselli, 2010; Nielsen et al., 2014).

It is possible that the imitation of cognitive rules is susceptible to cultural experience, and Chinese culture presents an interesting theoretical test. China and other Asian countries have been dubbed "collectivist" cultures (Markus and Kitayama, 1991; Oyserman et al., 2002). Because of language and culture, people raised in China are thought to place relatively more emphasis on harmonious relationships than those raised in the USA and other Western cultures, dubbed "individualist" cultures. Chinese parenting practices highlight the value of groups, social cohesion, and conformity in behavior (Jose et al., 2000). Chinese society also emphasizes allowing others to save "Mian Zi" or "Face," which commonly leads to implicit and conservative expressions of one's opinion (Redding and Ng, 1982). Chinese child-rearing practices may provide a fertile training ground for highlighting the invisible rules and motivations that explain visible behaviors.

The current experiment tests Chinese children's abstract imitation of rules and compares it to Wang et al.'s (2015) existing data set from American children. All children were presented with four visually-identical objects, two heavy and two light. In the Experimental group, children saw an adult heft each object and sort them (by weight) into two bins. Two control groups were used to determine what elements of the demonstration were needed to promote weight sorting. Specifically, we tested whether seeing an intentional sorting demonstration (Experimental treatment) was more effective for eliciting weight sorting than was seeing the hefting acts alone (a control for "stimulus enhancement") or the hefting acts + the sorted endstate (a control for "emulation," or duplicating the endstate).

One question was whether the focus on group cohesion and conformity in China may emphasize the underlying meaning of others' behavior, which would give Chinese children an advantage in learning a non-obvious cognitive rule such as categorizing by the invisible property of weight. However, the abstract imitation of rules may be available during the early years in all cultures—a cultural universal that propels further cognitive development.

Equally important to the cross-cultural aspect, we sought to illuminate how imitation can inform theories about the relation between perception, action, and cognitive development. Past research has suggested that reproducing specific actions may prompt children to learn the underlying purpose of an act (e.g., Williamson and Markman, 2006). If this is the case, children's imitation of the adult's specific weighing and "hefting actions" (lifting up and down) may help them isolate and infer that underlying weight differences are the basis for categorizing the visually-identical objects. If so, it would illuminate how action imitation could foster the development of cognitive rules (see Discussion for further elaboration).

Materials and Methods

Participants

The participants were ninety-six 4-year-old children. Half were Chinese ($N = 48$; $M = 53.06$ months, SD = 3.77 months; 24 males) and half were American ($N = 48$, $M = 48.92$ months, SD = 1.66 months; 24 males). Chinese participants were recruited from a kindergarten affiliated with a university in China, which primarily enrolls children of Han ethnicity. American participants were recruited from a large metropolitan area (the sample was 78% White, 16% Black/African American, 3% other, with 2% being of Hispanic ethnicity, and 1% not reporting).

American children were tested individually in the laboratory, and their behaviors were videotaped for subsequent scoring.
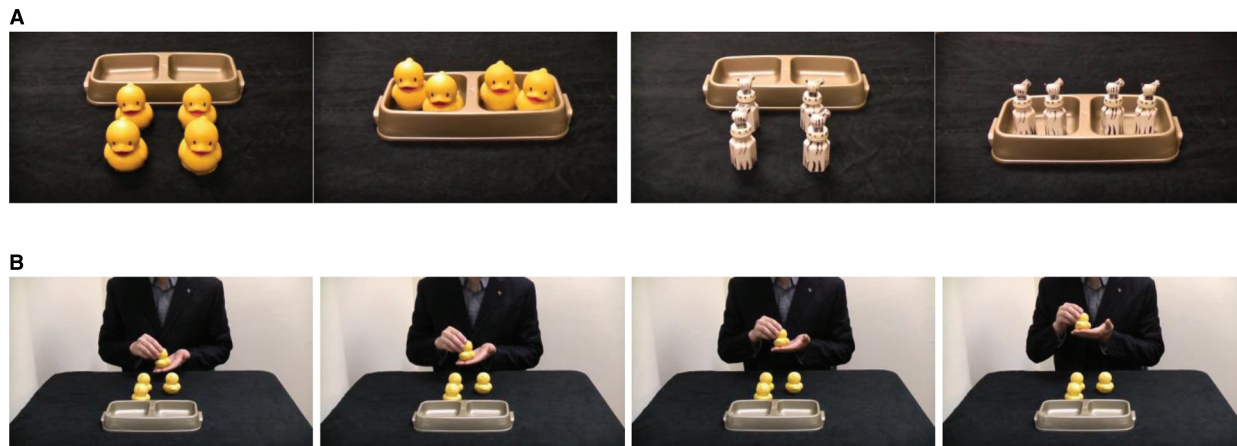
FIGURE 1 | Photographs of the experimental materials (A), which consist of sets of four visually-identical ducks, four visually-identical zebras, and bins to sort the objects into. Within each set, two objects are heavy and two are light. Row **(B)** shows an adult making an upward "hefting" motion, which consisted of a full cycle of raising and lowering the object on a flat palm as if weighing the object.

Chinese children were tested individually in a quiet room at their school. Georgia State University's institutional review board (IRB) provided oversight of the project.

Materials

Four sets of four objects were used as stimuli (**Figure 1A**). Two sets consisted of four yellow rubber ducks (5.5 cm × 4.5 cm × 5 cm) each. The other two sets consisted of four plastic zebras (5 cm × 5 cm × 4 cm). In each set, the four objects were visually identical, but unbeknownst to the child, differed in the invisible property of weight. For each duck set, two ducks weighed 87.5 g ("heavy"), and two weighed 21.7 g ("light"). For each zebra set, two zebras weighed 41.5 g ("heavy"), and two weighed 11.6 g ("light"). Pilot work suggested that the two weights used in each set were readily discriminable by untrained adults. The objects could not be discriminated by vision or audition (none of the objects made sound when manipulated, because the interior chambers were either filled or empty). The objects were spatially sorted into a two-bowled tray (23.5 cm × 5 cm × 4.5 cm), hereafter referred to as "bins."

Procedure

Each child was randomly assigned to one of three independent experimental groups. In all groups, the procedure consisted of a demonstration and a response period. The following three factors were counterbalanced within and between the experimental groups: (a) child's gender, (b) the order in which the stimuli were presented (ducks or zebras as the first set), and (c) the side on which the heavy objects were placed during the demonstration (left vs. right). Each group had 16 Chinese and 16 American children.

Demonstration Period

Experimental group: hefting + sorting

The experimenter placed one set of objects (e.g., the ducks) on the table in a square arrangement (approximately 12 cm ×

12 cm). The two objects of one weight were located on the right of the square, and the two objects of the other weight were on the left. The weight difference was not visible and thus unknown to the child. From the experimenter's viewpoint, the bins were placed on the table behind the objects (**Figure 1B**). Then the experimenter drew the child's attention (e.g., "It's my turn first").

In this group, children saw the experimenter intentionally sort the objects by weight. The experimenter picked up the object that was closest to the child (and on the child's right), put it on his palm, and "hefted" it six times, as if to test the object's weight by bobbing it up and down on a flat palm in a weighing motion (see **Figure 1B**). The object was then placed into the bin on the child's right. Next, the experimenter picked up the second object from the child's right side, hefted it in the same way, and placed it in the same bin. The experimenter then hefted each of the two remaining objects in the same way, and placed each of them into the other bin. The experimenter had a neutral, pleasant facial expression throughout this demonstration. The hefting motion was identical for all objects, because the experimenter practiced doing it in the same way for each object, and the difference in weight was so minimal that the kinematics of the lift could be done in the same manner.

Control-group 1: hefting + no sorting

In this control group, the experimenter handled each object, but did not sort them. This group was used to control for "stimulus enhancement" that may occur when the adult handles the test objects. The experimenter placed one set of objects on the table in the square arrangement, and drew the children's attention to the objects ("it's my turn"). Then, the experimenter picked up each object and hefted it, exactly as in the Experimental group, but instead of sorting the objects, each one was placed back on the table in its original location after it was hefted. Thus, in this control group, the children saw only the weighing process, but not the sorting behavior.

Control-group 2: hefting + presorted

In this control, children saw the experimenter handle each object and also saw the endstate of the objects sorted in the bins. The crucial difference was that the experimenter never sorted the objects into the bins. Instead, the four objects were brought on the table already pre–sorted into the bins. This group controls for, "emulation," or duplication of the endstate array. The experimenter drew the child's attention ("it's my turn"), picked up each of the objects in turn, hefted them, and returned each to its location in the bins. Thus, for this group the children saw the weighing behavior and also the perceptual endstate that was shown in the Experimental group, but the participant never saw the adult sort the objects.

Response Period

The response period was identical for all groups. The experimenter placed the four objects in front of the child, with the bins behind the objects (from the child's viewpoint, see **Figure 1A**). The objects were placed in a square configuration, but the two objects with the same weight were now switched (unbeknownst to the children) and placed in the horizontal rows. The spatial positioning of the objects was changed from the demonstration period so that the children had to use the object weights, and not simply the experimenter's picking and placing movements, in order to correctly sort the objects. If children only copied the literal movements of the experimenter, they would not succeed in sorting by weight, because the array was transformed between the demonstration and response period as described. (Furthermore, the location of the heavy and light objects in the front vs. back rows was alternated for the response periods in each of the four trials. Thus if the two heavy objects were in the row closest to the child in the response period in trial 1, then they were in the row farthest from the child in trial 2, etc.)

The children were given a prompt to act, but there was no linguistic description about the content of the act. The experimenter simply made the neutral comment, "Now it's your turn." Children were allowed to manipulate the objects until they placed all four into the bins. If needed, the children were prompted with the question, "Can you put them inside?" After the children placed the four objects into the bins, the experimenter removed the bins for later scoring. For trial 2, the children were given an identical group of objects to sort. No demonstration was given for this set. This second set of materials was necessary because it was not always possible to score from the video with 100% certainty what the child did with the heavy/light objects, because they all were visually identical, and sometimes the child's arm blocked a camera view; thus we retained the bins for subsequent scoring.

After these two trials, a visually novel set of four objects was introduced. If the duck set was used in the demonstration, the zebra set was used as the generalization set and *vice versa*. Crucially, these objects also differed in their absolute weights from the original (see Materials), and the experimenter did not perform any sorting demonstration with these objects. These trials were designed to assess whether children would *generalize* the weight-sorting rule to the novel stimuli. The experimenter placed the four objects of the generalization set on the table in a square arrangement (with the heavy vs. light objects in horizontal rows, see

counterbalancing above) and children were given two response periods as described above.

Dependent Measures and Scoring

Sorting score

The primary dependent measure is the number of trials in which the participants sorted the four objects by weight. To be credited with a "correct sort," children had to group the two objects of one weight in one bin and the two objects of the other weight in the other bin. Each correct sort was scored as a 1, which yields a sorting score ranging from 0 to 4 across the four trials.

Hefting score

Another dependent measure was also scored—children's imitation of the hefting action that the adult had used (**Figure 1B**). There were three components: (a) holding the object from underneath with a flat palm, (b) hefting the object by raising the hand and letting it fall, and (c) stabilizing the object with the second hand. If children reproduced all the three components at least once in a trial, they received a score of 1 for that trial. Otherwise, the score for the trial was 0. A child's hefting score ranged from 0 to 4 (1 possible point for each of the four test trials).

Scoring agreement

The primary scorer was a research assistant who remained uninformed of the participant's group assignment and the study hypotheses. A second scorer, also unaware of group assignment, coded a randomly selected 25% of the participants. Intercoder agreement was assessed using the Intraclass correlation coefficient (ICC = 0.98). Due to IRB restrictions, videos are not available for the Chinese children. Only the American children's hefting was scored. (In the American sample, three video records were unavailable resulting in a final $N = 45$ for the hefting analysis.)

Results

Preliminary analyses showed no significant effects of participant sex, the side on which the weights were placed, object type (ducks vs. zebras), or presentation order (ducks vs. zebras first). We collapsed across these factors in all subsequent analyses.

Object Categorization

Our first analyses test for differences in whether children sorted the sets of objects by weight as a function of experimental group. Children's sorting scores were analyzed using a 2(Culture: Chinese vs. American) × 3(Test group: Experimental, Control-1, Control-2) × 2(Object set: Demonstration set vs. Generalization set) repeated-measures ANOVA. **Figure 2** shows the sorting scores as a function of Culture and Test group. This analysis revealed a significant main effect of Test group, $F(2,96) = 9.03$, $p < 0.001$, $\eta_p^2 = 0.17$. Follow-up pairwise comparisons (Student–Newman–Keuls) indicated that children in the Experimental group ($M = 2.50$, SD = 0.95) had significantly higher sorting scores than did children in either the Control-1 ($M = 1.41$, SD = 1.18; $p < 0.001$) or Control-2 ($M = 1.53$, SD = 1.19; $p = 0.002$) groups, with no significant difference between the two controls ($p = 0.87$).

This analysis also revealed several notable non-significant comparisons. Culture showed no significant main effect, $F(1,96) = 1.91$, $p = 0.17$, $\eta_p^2 = 0.02$, or interaction with Test group, $F(2,95) = 0.48$, $p = 0.62$, $\eta_p^2 = 0.01$. There was also no significant main effect of Object set, $F(1,90) = 0.11$, $p = 0.74$, $\eta_p^2 = 0.001$, or Test group × Object set interaction, $F(2,90) = 0.33$, $p = 0.72$, $\eta_p^2 = 0.007$.

There was evidence of generalization. Children's sorting scores on the Demonstration and Generalization objects were respectively: Experimental group: $M = 1.19$, $SD = 0.64$, $M = 1.31$, $SD = 0.69$; Control-1: $M = 0.75$, $SD = 0.84$, $M = 0.69$, $SD = 0.59$; Control-2: $M = 0.75$, $SD = 0.76$, $M = 0.78$, $SD = 0.83$. No significant difference was found between children's performance on Demonstration and Generalization objects, $t(31) = -0.75$, $p = 0.46$, $d = 0.18$, indicating that children in the Experimental group did just as well on sorting the novel objects by weight as they did in sorting the ones that the adult originally used in the demonstration–generalization. Further evidence of generalization is that 50% (16/32) of the children in the Experimental group sorted objects in three or four trials versus 20.3% (13/64) in the controls, $\chi^2 (4,92) = 14.70$, $p = 0.005$, Cramer's $V = 0.28$.

We also conducted a more over-arching test of children's performance. Children's sorting scores were compared to chance. To calculate the chance value, we assumed that two objects were placed into each bin (children did this on 93.9% of trials). There are 24 possible arrangements of the four objects in the two bins. By chance combinations alone, in 8 of these 24 combinations the heavy objects would be grouped together in one bin and the light objects in the other bin. Thus, the chance probability that the final array will consist of two objects of the same weight placed in each bin is 0.33. Considering that there are four trials, chance performance is a sorting score of 1.33 (4 trials × 0.33). A one-sample t-test revealed that children in the Experimental group categorized the objects by weight significantly more often than is expected by chance, $t(31) = 6.96$, $p < 0.001$, $d = 2.50$. In contrast, children's performance in the Control-1 ($p = 0.72$) and Control-2 ($p = 0.35$) groups was not significantly different from chance. This same effect was also obtained for the Chinese and American cultures tested individually.

Hefting Behavior

This analysis assesses whether children imitated the specific "hefting" act and how this interacted with their learning the cognitive rule of categorizing the objects by weight. This question is of interest because one way that children could learn about weight is by imitating the motor acts of hefting (bobbing the object up and down in the hand while supporting it), even if they did not fully understand why the adult was doing this act. In this way, imitation of the motor act might potentially engender learning about the property of the object. For this analysis, we classified children across test groups into one of three sorting types based on their sorting scores. Children who correctly sorted the objects on three or four trials were considered to have a high sorting score (high sorters, $n = 13$). Children who sorted the objects on two trials were considered medium sorters ($n = 7$). Children with a sorting score of 0 or 1 were categorized as low sorters ($n = 25$).
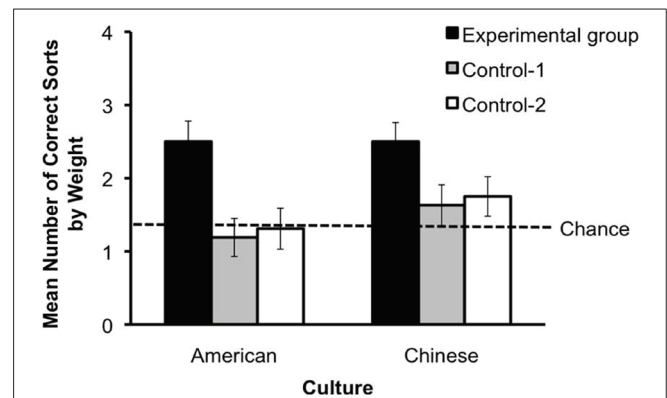

FIGURE 2 | Mean number of weight sorts ($\pm$SE) as a function of test group and culture.
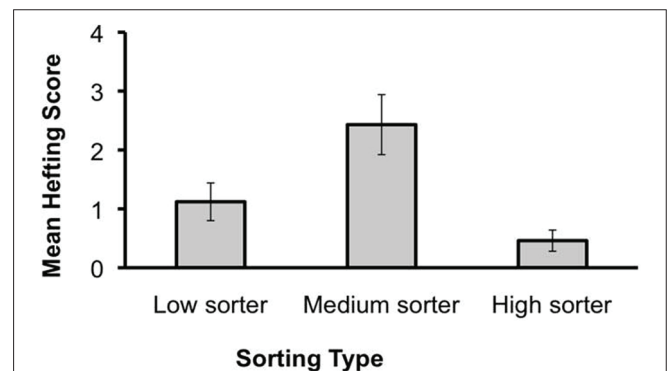

FIGURE 3 | Mean hefting score ($\pm$SE) as a function of sorting type.

A one-way ANOVA using sorting type (three levels) as the between-subject factor was conducted on children's hefting scores. Children's imitation of the adult's hefting act was related to their sorting performance, $F(2,42) = 4.04$, $p = 0.03$, $\eta_p^2 = 0.16$ (**Figure 3**). A follow-up pairwise test (Student–Newman–Keuls) indicated that the medium sorters had significantly higher hefting scores than did the high sorters ($p = 0.01$), with intermediate performance by the low sorters (see Discussion for further consideration).

Discussion

Based on the adult demonstration, both American and Chinese children abstracted the categorization rule of sorting objects by weight. The low levels of sorting by the children in Control-1 (hefting + no sorting) establishes that merely seeing the adult's weighing actions alone is not enough to induce children to categorize the objects. Control-2 (hefting + presort) establishes that seeing both the adult's hefting gestures and the final sorted endstate is also not enough. This latter result is particularly striking and important because the behaviors used during the demonstration period of Control-2 closely trace those used in the Experimental group. In the Control-2 group, the experimenter picked up the presorted objects from the bins and returned them to the same

position; in the Experimental group, the experimenter picked the objects from the table and sorted them into the bins. Neither the hefting nor the final endstate was sufficient to promote weight sorting. *We therefore suggest that the rule learning was based on the perception and imitation of the adult's goal-directed sorting behavior.*

Action Observation and Cognitive Rule Learning

Children who saw the experimental demonstration of categorizing visually-identical objects by the invisible property of weight showed higher rates of sorting the objects by weight than would be expected by chance. Several elements of the experimental design indicate that the children had to go beyond copying the adult's specific motor actions alone to succeed. The spatial positioning of the heavy and light objects was switched between the adult's demonstration and the response period. This means that if the children duplicated the literal picking up and placing movements of the adult, the objects would not have been grouped by weight. Further, the objects in each set looked identical—there were no visual cues and no auditory cues for categorizing the objects. The finding of weight sorting is in line with arguments that children's categorization is not limited to considering only visual perceptual features, but can include the consideration of invisible and internal properties of objects (e.g., Gelman and Wellman, 1991; Gelman, 2003).

Extensive previous research has established that understanding object weight is a challenging cognitive task for children of the age tested here and even older (Piaget, 1951; Smith et al., 1985; Schrauf and Call, 2009; Schrauf et al., 2011; Povinelli, 2012). During the preschool years, in particular, children struggle to consider this internal and invisible property in the absence of correlated visible cues (Smith et al., 1985). Profound difficulties with weight have also been reported in comparative work (Vonk and Povinelli, 2006; Schrauf and Call, 2009; Povinelli, 2012). The key suggestion made in this paper is that social learning and imitation can prompt children's attention and cognitive inferences about the invisible property of weight.

We come, then, to the crux of the problem: what exactly did children learn about weight from observing the adult's sorting actions? One possibility is that they learned that the objects had different weights. The hefting movements used by the adult may be one cue to this invisible property. Seeing the hefting act coupled with intentional sorting behavior by the adult may have prompted children to seek an explanation for this complex behavioral stream. A good candidate explanation may be an internal, invisible property such as weight (for related discussions, see Legare et al., 2010; Legare and Lombrozo, 2014; Meltzoff and Gopnik, 2013). An additional possibility, not mutually exclusive, is that children might have already had an inkling about object weight and gained information about the adult's goals or how to behave in this contextual situation—people sort by weight.

An important characteristic of children's weight sorting in this experiment is that it was generalizable. The adult manipulated only the first set of objects, but the children in the Experimental group were equally likely to sort on the generalization trials. This finding highlights that rules, once abstracted, can be applied to new objects and across situations. Thus, if a child learns to

consider weight when picking melons, she could also consider this invisible property in relation to other types of objects. Overall, these current findings indicate that observing the act of categorization promoted children to make use of weight with novel objects on new trials.

Action Imitation and Cognition

Some children were more likely to imitate the hefting acts that the adult demonstrated. Children with medium sorting scores hefted the objects on significantly more trials than did children who had high sorting scores.

One function of imitating others' hefting actions with high fidelity is that it may afford children the opportunity to discover the significance of behaviors that are not understood (Williamson and Markman, 2006). Whether or not children actually understand the deeper purpose of the hefting acts, children gain first-hand experience with the weight of the objects when they imitate the hefting behaviors. This experience may have been less important for children who readily inferred the sorting rule (the high sorters)—indeed they may have realized that imitating the hefting acts was unnecessary for completing the goal of categorizing the objects by weight. However, it is possible that imitating those specific acts with high fidelity helped the intermediate sorters to attend to or recognize the weight difference and its significance, and then to use this property to categorize the objects. Although the data are too limited to draw strong conclusions, they raise intriguing links between action imitation and cognitive development—with action observation sparking action production, which may direct attention, experience, and cognitive change.

Cross-Cultural Universals in Imitation

In China, there is generally a greater emphasis on conformity and the implicit expression of ideas than in the individualistic American culture (Markus and Kitayama, 1991; Oyserman et al., 2002). However, despite the differences in parenting practices and cultural norms, we found no difference in children's imitation of the rules tested here. It is possible that culture exerts an influence on rule imitation that was not detected in this experiment with this specific physical-based rule (vs. a more psychological attribution). It should also be recognized that the children in both the USA and China were recruited from middle- to upper-middle class families, and with increasing globalization, it is possible that any cultural differences due to traditional child-rearing practices are not as pronounced in people of closely matched socio-economic backgrounds. Additionally, preschool children may not have had sufficient cultural experience to show differences that may emerge later; or there may be a different developmental time course for social rules and customs than for those based on physical properties such as weight. One recent example showed a different time course in the acquisition of cultural stereotypes about math in children raised in Asian vs. North American culture (Cvencek et al., 2014).

The findings of the current study are consistent with a growing body of research showing similarities in children's imitation across a variety of cultures (e.g., Callaghan et al., 2011; Nielsen et al., 2014). Past studies have generally targeted the reproduction of specific actions on objects while the current study targeted

the reproduction of an abstract cognitive rule underlying such behaviors. In early development especially, the observation and use of others' actions may draw primarily on cultural universals. Children around the world may use imitation in similar ways to learn new, generalizable information from other social agents.

Conclusion

The current study provides three contributions. First, it shows that children's imitation goes beyond replicating specific motor movements. Children also imitate abstract rules or strategies that guide behavior, such as rules for categorization. Such "abstract imitation" (Williamson et al., 2010) is important for children's acquisition of both instrumental skills and cultural practices. Second, this

research also suggests that a different type of imitation, specifically children's high-fidelity imitation of motor acts, may serve as a lever in their acquisition of abstract cognitive rules. Children who did not understand the weight-sorting demonstration may have benefitted from reproducing the adult's exact "hefting" acts. This use of imitation of literal behavior as a mechanism for rule learning deserves more research. Third, this research extends previous findings of cross-cultural similarity in social learning to an area beyond the imitation of particular acts to the imitation of more generalizable rules (categorization rules). Overall, these findings and others support the view that action representation and imitation may be key mechanisms for the rapid acquisition and spread of generalizable skills, knowledge, and customs in human cultures.

References

Boyd, R., and Richerson, P. J. (2005). *The Origin and Evolution of Cultures*. Oxford: Oxford University Press.

Buchsbaum, D., Gopnik, A., Griffiths, T. L., and Shafto, P. (2011). Children's imitation of causal action sequences is influenced by statistical and pedagogical evidence. *Cognition* 120, 331–340. doi: 10.1016/j.cognition.2010.12.001

Callaghan, T., Moll, H., Rakoczy, H., Warneken, F., Liszkowski, U., Behne, T., et al. (2011). Early social cognition in three cultural contexts. *Monogr. Soc. Res. Child Dev.* 76, 1–142. doi: 10.1111/j.1540-5834.2011.00603.x

Cvencek, D., Meltzoff, A. N., and Kapur, M. (2014). Cognitive consistency and math–gender stereotypes in Singaporean children. *J. Exp. Child Psychol.* 117, 73–91. doi: 10.1016/j.jecp.2013.07.018

Flynn, E., and Whiten, A. (2008). Cultural transmission of tool use in young children: a diffusion chain study. *Soc. Dev.* 17, 699–718. doi: 10.1111/J.1467-9507.2007.00453.X

Gelman, S. A. (2003). *The Essential Child: Origins of Essentialism in Everyday Thought*. New York, NY: Oxford University Press.

Gelman, S. A., and Wellman, H. M. (1991). Insides and essences: early understandings of the non-obvious. *Cognition* 38, 213–244. doi: 10.1016/0010-0277(91)90007-Q

Herrmann, P. A., Legare, C. H., Harris, P. L., and Whitehouse, H. (2013). Stick to the script: the effect of witnessing multiple actors on children's imitation. *Cognition* 129, 536–543. doi: 10.1016/j.cognition.2013.08.010

Horner, V., and Whiten, A. (2005). Causal knowledge and imitation/emulation switching in chimpanzees (*Pan troglodytes*) and children (*Homo sapiens*). *Anim. Cogn.* 8, 164–181. doi: 10.1007/S10071-004-0239-6

Jose, P. E., Huntsinger, C. S., Huntsinger, P. R., and Liaw, F. R. (2000). Parental values and practices relevant to young children's social development in Taiwan and the United States. *J. Cross Cult. Psychol.* 31, 677–702. doi: 10.1177/0022022100031006002

Legare, C. H., Gelman, S. A., and Wellman, H. M. (2010). Inconsistency with prior knowledge triggers children's causal explanatory reasoning. *Child Dev.* 81, 929–944. doi: 10.1111/j.1467-8624.2010.01443.x

Legare, C. H., and Lombrozo, T. (2014). Selective effects of explanation on learning during early childhood. *J. Exp. Child Psychol.* 126, 198–212. doi: 10.1016/j.jecp.2014.03.001

Loucks, J., and Meltzoff, A. N. (2013). Goals influence memory and imitation for dynamic human action in 36-month-old children. *Scand. J. Psychol.* 54, 41–50. doi: 10.1111/sjop.12004

Markus, H. R., and Kitayama, S. (1991). Culture and the self: implications for cognition, emotion, and motivation. *Psychol. Rev.* 98, 224–253. doi: 10.1037/0033-295X.98.2.224

Marshall, P. J., and Meltzoff, A. N. (2014). Neural mirroring mechanisms and imitation in human infants. *Philos. Trans. R. Soc. Biol. Sci.* 369, 20130620. doi: 10.1098/rstb.2013.0620

Meltzoff, A. N. (1988). Infant imitation after a 1-week delay: long-term memory for novel acts and multiple stimuli. *Dev. Psychol.* 24, 470–476. doi: 10.1037/0012-1649.24.4.470

Meltzoff, A. N. (1995). Understanding the intentions of others: re-enactment of intended acts by 18-month-old children. *Dev. Psychol.* 31, 838–850. doi: 10.1037/0012-1649.31.5.838

Meltzoff, A. N., and Gopnik, A. (2013). "Learning about the mind from evidence: children's development of intuitive theories of perception and personality," in *Understanding Other Minds*, 3rd Edn, eds S. Baron-Cohen, H. Tager-Flusberg, and M. Lombardo (Oxford: Oxford University Press), 19–34.

Meltzoff, A. N., Kuhl, P. K., Movellan, J., and Sejnowski, T. J. (2009). Foundations for a new science of learning. *Science* 325, 284–288. doi: 10.1126/science.1175626

Meltzoff, A. N., and Williamson, R. A. (2013). "Imitation: social, cognitive, and theoretical perspectives," in *Oxford Handbook of Developmental Psychology*, Vol. 1, ed. P. R. Zelazo (New York, NY: Oxford University Press), 651–682.

Nielsen, M., and Tomaselli, K. (2010). Over-imitation in Kalahari Bushman children and the origins of human cultural cognition. *Psychol. Sci.* 21, 729–736. doi: 10.1177/0956797610368808

Nielsen, M., Tomaselli, K., Mushin, I., and Whiten, A. (2014). Where culture takes hold: "overimitation" and its flexible deployment in Western, Aboriginal, and Bushmen children. *Child Dev.* 85, 2169–2184. doi: 10.1111/cdev.12265

Oyserman, D., Coon, H. M., and Kemmelmeier, M. (2002). Rethinking individualism and collectivism: evaluation of theoretical assumptions and meta-analyses. *Psychol. Bull.* 128, 3–72. doi: 10.1037/0033-2909.128.1.3

Piaget, J. (1951). *The Psychology of Intelligence*. London: Routledge and Kegan Paul Ltd.

Povinelli, D. J. (2012). *World without Weight: Perspectives on an Alien Mind*. Oxford: Oxford University Press.

Redding, S. G., and Ng, M. (1982). The role of "face" in the organizational perceptions of Chinese managers. *Organ. Stud.* 3, 201–219.

Schrauf, C., and Call, J. (2009). Great apes' performance in discriminating weight and achromatic color. *Anim. Cogn.* 12, 567–574. doi: 10.1007/s10071-009-0216-1

Schrauf, C., Call, J., and Pauen, S. (2011). The effect of plausible versus implausible balance scale feedback on the expectancies of 3- to 4-year-old children. *J. Cogn. Dev.* 12, 518–536. doi: 10.1080/15248372.2011.571647

Schulz, L. E., Hooppell, C., and Jenkins, A. C. (2008). Judicious imitation: children differentially imitate deterministically and probabilistically effective actions. *Child Dev.* 79, 395–410. doi: 10.1111/j.1467-8624.2007.01132.x

Smith, C., Carey, S., and Wiser, M. (1985). On differentiation: a case study of the development of the concepts of size, weight, and density. *Cognition* 21, 177–237. doi: 10.1016/0010-0277(85)90025-3

Subiaul, F., Lurie, H., Romansky, K., Klein, T., Holmes, D., and Terrace, H. (2007a). Cognitive imitation in typically-developing 3- and 4-year olds and individuals with autism. *Cogn. Dev.* 22, 230–243. doi: 10.1016/J.Cogdev.2006.10.003

Subiaul, F., Romansky, K., Cantlon, J. F., Klein, T., and Terrace, H. (2007b). Cognitive imitation in 2-year old children (*Homo sapiens*): a comparison with rhesus monkeys (*Macaca mulatta*). *Anim. Cogn.* 10, 369–375. doi: 10.1016/J.Cogdev.2006.10.003

Subiaul, F., Patterson, E. M., Schilder, B., Renner, E., and Barr, R. (2014). Becoming a high-fidelity—super—imitator: what are the contributions of social and individual learning? *Dev. Sci.* doi: 10.1111/desc.12276 [Epub ahead of print].

Tomasello, M. (1999). *The Cultural Origins of Human Cognition.* Cambridge, MA: Harvard University Press.

Vonk, J., and Povinelli, D. J. (2006). "Similarity and difference in the conceptual systems of primates: the unobservability hypothesis," in *Comparative Cognition: Experimental Explorations of Animal Intelligence*, eds T. Zentall and E. A. Wasserman (Oxford: Oxford University Press), 363–387.

Waismeyer, A., Meltzoff, A. N., and Gopnik, A. (2015). Causal learning from probabilistic events in 24-month-olds: an action measure. *Dev. Sci.* 18, 175–182. doi: 10.1111/desc.12208

Wang, L., Fu, X., Zimmer, H. D., and Aschersleben, G. (2012). Familiarity and complexity modulate the way children imitate tool-use actions: a cross-cultural study. *J. Cogn. Psychol.* 24, 221–228. doi: 10.1080/20445911.2011.617300

Wang, Z., Meltzoff, A. N., and Williamson, R. A. (2015). Social learning promotes understanding of the physical world: preschool children's imitation of weight sorting. *J. Exp. Child Psychol.* doi: 10.1016/j.jecp.2015.02.010 [Epub ahead of print].

Whiten, A., Flynn, E., Brown, K., and Lee, T. (2006). Imitation of hierarchical action structure by young children. *Dev. Sci.* 9, 574–582. doi: 10.1111/j.1467-7687.2006.00535.x

Whiten, A., McGuigan, N., Marshall-Pescini, S., and Hopper, L. M. (2009). Emulation, imitation, over-imitation and the scope of culture for child and chimpanzee. *Philos. Trans. R. Soc. B Biol. Sci.* 364, 2417–2428. doi: 10.1098/rstb.2009.0069

Williamson, R. A., Jaswal, V. K., and Meltzoff, A. N. (2010). Learning the rules: observation and imitation of a sorting strategy by 36-month-old children. *Dev. Psychol.* 46, 57–65. doi: 10.1037/a0017473

Williamson, R. A., and Markman, E. M. (2006). Precision of imitation as a function of preschoolers' understanding of the goal of the demonstration. *Dev. Psychol.* 42, 723–731. doi: 10.1037/0012-1649.42.4.723

Walking indoors, walking outdoors: an fMRI study

Riccardo Dalla Volta[1], Fabrizio Fasano[2], Antonio Cerasa[3], Graziella Mangone[3], Aldo Quattrone[1,3] and Giovanni Buccino[1,4]*

[1] Dipartimento di Scienze Mediche e Chirurgiche, Università Magna Graecia, Catanzaro, Italy, [2] Dipartimento di Neuroscienze, Università di Parma, Parma, Italy, [3] IBFM Istituto di Bioimmagini e Fisiologia Molecolare, Consiglio Nazionale delle Ricerche, Germaneto, Italy, [4] IRCCS Neuromed, Pozzilli, Italy

An observation/execution matching system for walking has not been assessed yet. The present fMRI study was aimed at assessing whether, as for object-directed actions, an observation/execution matching system is active for walking and whether the spatial context of walking (open or narrow space) recruits different neural correlates. Two experimental conditions were employed. In the execution condition, while being scanned, participants performed walking on a rolling cylinder located just outside the scanner. The same action was performed also while observing a video presenting either an open space (a country field) or a narrow space (a corridor). In the observation condition, participants observed a video presenting an individual walking on the same cylinder on which the actual action was executed, the open space video and the narrow space video, respectively. Results showed common bilateral activations in the dorsal premotor/supplementary motor areas and in the posterior parietal lobe for both execution and observation of walking, thus supporting a matching system for this action. Moreover, specific sectors of the occipital–temporal cortex and the middle temporal gyrus were consistently active when processing a narrow space versus an open one, thus suggesting their involvement in the visuo-motor transformation required when walking in a narrow space. We forward that the present findings may have implications for rehabilitation of gait and sport training.

Keywords: walking, fMRI, mirror neuron system, space coding, rehabilitation

Introduction

Action observation and recognition are fundamental tasks on which social interactions are based. From these abilities also derives the capacity to quickly and accurately recognize intentions and feelings of other individuals based on their non-verbal behavior. There is increasing evidence that these cognitive tasks may be explained to some extent by a mechanism matching the observed action and motor behavior with an internal motor representation of that action or motor behavior in the brain of the observer. Cortical areas endowed with this observation/execution matching mechanism (*mirror mechanism*) are known as the mirror neuron system (MNS) (Fabbri-Destro and Rizzolatti, 2008). This system has been involved in a number of cognitive functions (Hari and Kujala, 2009) including social cognition (Gallese and Goldman, 1998). In addition, this system appears to be impaired in autism spectrum disorder where social cognition is markedly affected (Iacoboni and Dapretto, 2006). It has been forwarded that the MNS embodies a unifying mechanism active whenever motor representations are recalled as during action observation,

motor imagery, dreams with a motor content and so on, even in the absence of overt action (Jeannerod, 2001).

Walking is a complex motor behavior with a special relevance in social interactions (for review, see Pavlova, 2012). By observing walking, people can extract a considerable amount of information including emotional states and intentions of the agent, even from sketchy descriptions of the body segments as it occurs with point-light biological motion stimuli (for review, see Johansson et al., 1980). That said, the cortical representation of human walking is still poorly understood. Most studies in the field assessed brain activation specifically related to execution, imagery and observation of walking, taken separately. Some studies investigated brain areas common to imagery and execution or, alternatively, imagery and observation of walking. To the best of our knowledge, however, the presence of an observation/execution matching system for this action remains to be assessed. For sake of completeness, in the following paragraphs we will shortly review current literature in the field considering first studies where execution, imagery and observation of walking were taken separately. Then, we will report about studies combining two of these tasks. Finally, we will take into account studies where the motor representation of foot actions, but not specifically walking, was investigated.

During execution of walking, brain imaging studies showed activations of several cortical (medial part of primary sensory-motor cortex, pSM, supplementary motor area, SMA, and premotor cortex, PM) and subcortical (basal ganglia and cerebellar vermis) structures (Greenstein et al., 1995; Ishii et al., 1995; Fukuyama et al., 1997; Tashiro et al., 2001). In some cases, also the recruitment of occipital and associative temporo-parietal cortices was found. A brain activation pattern similar to that of walking execution was also found during pure motor imagery of walking (Malouin et al., 2003; Sacco et al., 2006; Bakker et al., 2008; Jahn et al., 2008; Wang et al., 2008; la Fougere et al., 2010; van der Meulen et al., 2012). In very recent studies, a set of parietal, frontal and temporo-occipital areas was also found during observation of walking (Abdollahi et al., 2013; Maffei et al., 2015).

Other studies combining walking execution and walking imagery (Miyai et al., 2001; la Fougere et al., 2010) showed a pattern of activation largely shared by both tasks. Differential activation between the two tasks was present in the primary motor cortex, which was typically engaged only during the actual execution of action. Motor imagery of walking has been compared also to walking observation (Iseki et al., 2008). Among common cortical structures subserving both tasks there were dorsal PM area bilaterally, left SMA and right SPL. In summary, results indicate a recruitment of the cortical sensory-motor system, with a significant convergence between execution and imagery on the one hand and imagery and observation on the other. Indeed, the notion that motor imagery and motor execution share common neural substrates is well established also for hand actions (Stephan et al., 1995; Decety, 1996; Porro et al., 1996; Gerardin et al., 2000; Solodkin et al., 2004).

It is worth stressing that for technical reasons, in all studies mentioned above where PET was employed, actual walking was an offline task executed before undergoing scanning. As for fMRI

studies, since walking is a motor task rather hard to be performed in a scanner, participants were asked to perform a motor imagery since imagery and actual walking execution partially share the same neural substrates.

Some studies have assessed the neural structures involved in foot and leg actions, but not specifically walking. The activation of a dorsal sector of the PM cortex and the parietal lobe has been shown during mere observation of foot actions (Buccino et al., 2001; Wheaton et al., 2004; Sakreida et al., 2005). During motor imagery of foot plantar- and/or dorsiflexion (Cramer et al., 2005, 2007; Gustin et al., 2010), activations were found in the pSM cortex and SMA but also in the cerebellum and subcortical structures (basal ganglia and thalamus). Similar activations were also found in studies where participants were asked to perform actual execution of foot actions in combination with motor imagery or observation of the same actions (Lafleur et al., 2002; Alkadhi et al., 2005; Enzinger et al., 2008; Hotz-Boendermaker et al., 2008; Orr et al., 2008; Rocca and Filippi, 2010; Yuan et al., 2010). Christensen et al. (2000) showed a similar pattern of activation during execution and imagination of bicycling.

The present fMRI study was aimed at investigating the common neural structures recruited during the execution and observation of walking. In order to study the correlates of active walking inside the MR scanner, we employed a rolling cylinder that allowed participants to move lower limbs as if they were really walking. Moreover, we were interested in assessing whether walking in different environments, namely an open space (country field) or a narrow space (corridor) may recruit different and specific neural substrates. Because of its motor relevance, in fact, there is evidence that the space near the body is differently coded from the space far from the body (Fogassi and Luppino, 2005; Brozzoli et al., 2012). At least for the hand, neurons were discovered in the monkey that preferred actions performed either near or far from the animal (Caggiano et al., 2009). Recent results by Costantini et al. (2010), suggest that perceiving affordances of an object recruits a motor act only when the object is presented within the near space of participants where interactions with objects are possible. As far as walking is concerned, behavioral and clinical evidence suggest that different space features in which walking occurs may be differently coded (Hashimoto, 2006; Salsone et al., 2009; Schicke et al., 2009).

Materials and Methods

Subjects

We studied 18 healthy Italian subjects (7 females, mean age 24, range 19–28 years) with no previous history of neurological or psychiatric disorders. All participants gave written informed consent, according to the Helsinki Declaration. The present study was approved by the Ethical Committee of the University "Magna Graecia" of Catanzaro.

Stimuli

The stimuli presented in the experiment consisted of video clips depicting either a walking action or different spatial contexts

(**Figure 1**). The *Walking* video clip showed the lower limbs of an individual lying supine and performing a walking action on a rolling cylinder. The *Open Space* video clip showed a countryside view while the *Narrow Space* video clip showed a narrow corridor. In both the videos depicting a space, the scene was filmed while the cameraman was actually walking in the countryside or in the corridor. In this way, the observation of these videos gave participants the feeling of walking into the observed space. Still images taken from the above mentioned video clips served as controls (*Still Walking, Still Open Space* and *Still Narrow Space*). Each video clip and the corresponding still image lasted 21 s and were preceded by the presentation (3 s) of written words at the center of the screen in order to cue participants.

Experimental Procedure

Each participant comfortably lay in the scanner with a forehead restraining strip and foam pads to ensure head fixation and minimize motion during scanning. Moreover, an adhesive band was fixed on the participant's jaw to help control movement. Visual stimuli were projected on an acrylic screen inside the MRI room. A mirror was placed on the head coil at 45° to the screen and the participant's line of sight. Functional MRI timing parameters and triggering of the visual stimulation were performed by an in-house software developed in LabView (National Instruments, Austin, TX, USA). Before scanning, all participants completed a 10-min practice session (which included stimuli different from those presented in the scanner). A cylinder rolling around a pivot was positioned in correspondence of the feet of participants that had their legs supported by a semi-rigid wedge (**Figure 2**).

The study consisted of two experimental conditions: (1) execution and (2) observation. During execution, subjects performed three different tasks: (a) walking on a rolling cylinder while looking at a gray screen (*Walking*, WW), (b) walking on a rolling cylinder while looking at the *Open Space* video (*Open Space Walking,* WO) and (c) walking on a rolling cylinder while looking at the *Narrow Space* video (*Narrow Space Walking,* WN). For each of these tasks, we used the following as controls: (a) gently pressing the rolling cylinder with the feet while looking at a gray screen (*Control for Walking*, CWW), (b) gently pressing the rolling cylinder with the feet while looking at the *Still Open Space* (*Control for Open Space Walking*, CWO),

and (c) gently pressing the rolling cylinder with the feet while looking at *Still Narrow Space* (*Control for Narrow Space Walking*, CWN). Before scanning, participants were trained to walk on the rolling cylinder and minimize head and trunk movements. During observation, participants performed the following tasks: (a) observing the *Walking* video (*Walking Observation,* OW), (b) observing the *Open Space* video (*Open Space Observation,* OO) and (c) observing the *Narrow Space* video (*Narrow Space Observation,* ON). For each of these tasks, we used the following as controls (a) observing the *Still Walking* (*Control for Walking Observation*, COW), (b) observing the *Still Open Space* (*Control for Open Space Observation*, COO), and (c) observing the *Still Narrow Space* (*Control for Narrow Space Observation*, CON). Each task was cued by a written instruction. The appearance of the words 'cammina' (i.e., walk), 'premi' (i.e., press), and 'guarda' (i.e., look at) cued participants to perform the different tasks. The present study used a block design with a pseudo-random presentation of the tasks (**Figure 3**). Each task was followed by the correspondent control (e.g., when participants walked on the rolling cylinder while looking at a gray screen, what followed was the control in which participants gently pressed the rolling cylinder with their feet while looking at the gray screen) and was presented four times.

Data Acquisition

MR data were acquired with a 3 Tesla scanner (Discovery MR-750, General Electric, Milwaukee, WI, USA) equipped with a 32-channel receiver head-coil. Functional images were acquired using a T2*-weighted gradient-echo, echo-planar (EPI) pulse sequence (acceleration factor (asset) 2, 32 interleaved transverse slices covering the whole brain, TR = 2000ms, TE = 30ms, flip-angle = 90°, FOV = 240 mm × 240mm, inter-slice gap = 0 mm, slice thickness = 4 mm, in-plane resolution 1.88 mm × 1.88 mm). From each participant, 576 volumes were collected in a single session. Additionally, a 3D structural T1-weighted spoiled gradient (SPGR) echo sequence was acquired.
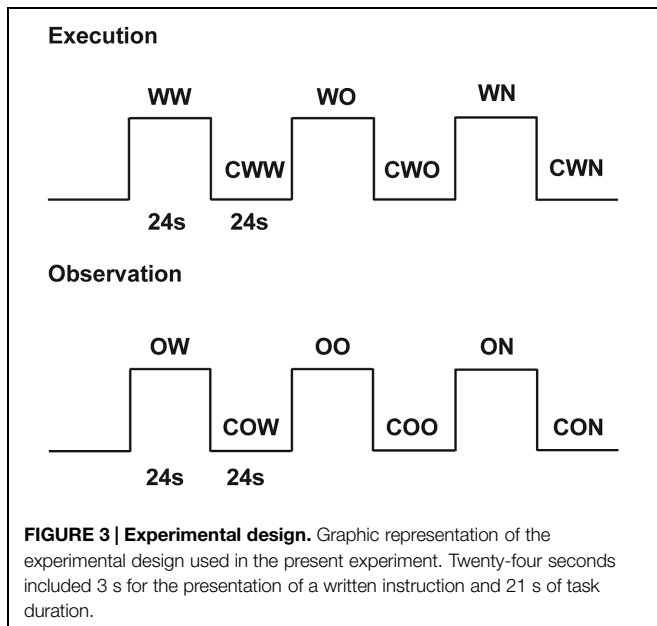
Data Analysis

Data analysis was performed with SPM8 (Statistical Parametric Mapping software by the Wellcome Trust Centre for Neuroimaging, Leopold Muller Functional Imaging Laboratory, University College of London, London, UK; http://www.fil.ion.ucl.ac.uk) running on MATLAB R2011a (The Mathworks,

FIGURE 1 | Stimuli. (A) A still frame taken from the video clip showing an individual lying supine and walking on a rolling cylinder. **(B)** A still frame taken from the video clip showing a countryside view (open space). **(C)** A still frame taken from the video clip showing a corridor (narrow space).

FIGURE 2 | Experimental set up. Picture showing a participant walking on the rolling cylinder while being scanned.

FIGURE 3 | Experimental design. Graphic representation of the experimental design used in the present experiment. Twenty-four seconds included 3 s for the presentation of a written instruction and 21 s of task duration.

Inc., Natick, MA, USA). The mean EPI was first computed for each participant and visually inspected to ensure that none showed artifacts. The first four EPI volumes of each functional run were discarded to allow for T1 equilibration effects. For each subject, all volumes were spatially realigned to the first volume of the run. Next, images were normalized to the EPI SPM template, re-sampled in 2 mm × 2 mm × 2mm voxels using trilinear interpolation in space and spatially smoothed with an 8 mm full-width half-maximum isotropic Gaussian kernel for the group analysis. Two participants showing head movements greater than 2 mm were excluded from all subsequent analyses.

Data were analyzed using a random-effects model (Friston et al., 1999), implemented in a two-level procedure. In the first level, single-subject fMRI data entered an independent General Linear Model (GLM) by design-matrixes modeling the

onsets and durations of 12 experimental factors, 6 related to the experimental tasks, and 6 related to their corresponding controls. For each participant, we generated contrast images displaying the effect of the experimental tasks contrasted with the respective controls: WW-CWW, WO-CWO, WN-CWN, OW-COW, OO-COO, ON-CON. In addition, images displaying the effect of walking in a narrow space contrasted with walking in an open one and vice versa, and images displaying the effect of observing a narrow space contrasted with observing an open one and vice versa were generated: WN-WO, WO-WN, ON-OO, OO-ON. Next, each contrast entered a second-level GLM to obtain: (*i*) SPM{T} maps (one sample *t*-test) related to each task at group-level and (*ii*) SPM{min(T)} maps (*conjunction* analysis) to test for (a) the existence of an observation/execution matching system for walking using the following contrasts (WW-CWW)∩(OW-COW), and (b) the existence of areas specifically involved in coding peripersonal and extrapersonal space, using the following contrasts: (WN-WO)∩(ON-OO) and (WO-WN)∩(OO-ON), respectively (Friston et al., 1999). To this aim, we performed an SPM '*conjunction null*' analysis (Nichols et al., 2005). Given the conservative nature of this analysis (Friston et al., 2005), we report data with a *p*-value < 0.001 uncorrected. A threshold of 10 was applied on cluster dimension. For all analyses, location of the activation foci was determined in the stereotaxic space of the MNI coordinates system. Those cerebral regions for which maps are provided were also localized with reference to cytoarchitectonical probabilistic maps of the human brain, using the SPM-Anatomy toolbox v1.7 (Eickhoff et al., 2005).

Head Movement

Since we asked for actual execution of walking inside the scanner, we were particularly careful in evaluating motion artifacts and minimizing their impact on the results. To this aim, for each subject all volumes were realigned to the first acquired one by applying a 6-parameters (rigid body) spatial transformation computed for each volume using a least-square approach. The mean head movement parameters were: *x*-direction 0.49 mm (±0.33), *y*-direction 0.41 (±0.26), and *z*-direction 1.94 (±1.13). The estimated six spatial transformation parameters computed for each volume entered as regressors in the subsequent model design matrix to de-convolve the head movement effect from the hemodynamic response.

Results

The main results of the conjunction analyses are shown in **Figure 4**. **Table 1** lists the MNI standard brain coordinates of the local maxima of BOLD-signal increases as revealed by all conjunction analyses. Common activations for execution and observation of walking, as revealed by the conjunction analysis (WW-CWW)∩(OW-COW), are shown in **Figure 4A**. Basically, a set of parieto-frontal areas emerged. Frontal activation foci were present in SMA and extended to adjacent dorsal PM cortex, bilaterally. In the right hemisphere a distinct spot in the dorsal premotor cortex was also present. Parietal activation foci were present in right SPL and in IPL bilaterally. In addition, activation
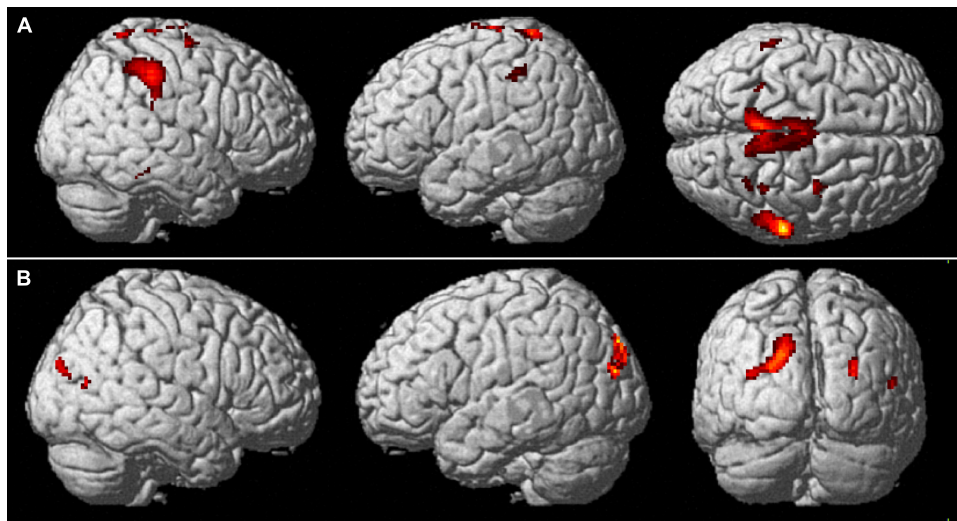
FIGURE 4 | Common brain activations from conjunction null analysis. (A) Upper row: Activations common to walking observation and execution ($p < 0.001$, uncorrected, $k = 10$). Activity is superimposed on a rendered brain viewed from the right (left panel), the left (middle panel), and the top (right panel). **(B)** Lower row: Activations specifically related to the processing of a narrow space as compared to an open one ($p < 0.001$ uncorrected, $k = 10$). Activity is superimposed on a rendered brain viewed from the right (left panel), the left (middle panel), and the posterior (left panel).

foci were present also in the cerebellar vermis and in both cerebellar hemispheres.

Cerebral activations related to the processing of a narrow space with respect to an open one, as revealed by the conjunction analysis (WN-WO)∩(ON-OO), are shown in **Figure 4B**. Basically, a set of occipital and parietal areas emerged. In particular, an occipital activation focus was located in the right middle occipital gyrus (MOG). The most caudal part of the intraparietal sulcus intersecting the transverse occipital sulcus (IntraParietal-TransverseOccipital, IPTO) was active bilaterally, but more largely represented in the left hemisphere where it extended toward the SPL.

Cerebral activations related to the processing of an open space with respect to a narrow one, as assessed by the conjunction analysis (WO-WN)∩(OO-ON), were found in the left inferior occipital gyrus.

Discussion

The present findings support the existence of an observation/execution matching system for walking and the presence of specific brain areas devoted to the coding of near space during walking. We will discuss these points eventually including their potential implication in rehabilitation of walking.

By means of an experimental setting where participants had the possibility to walk on a rolling cylinder while being scanned, we could investigate active walking, in addition to walking observation, while reducing to a minimum movement artifacts. However, it is worth underlining that walking in a scanner remains an approximation of walking in natural contexts, thus preventing us from the possibility of assessing the role of postural adjustments, gravity and so on. That said, the results of the

present study revealed a set of parieto-frontal areas active during both execution and observation of walking, including the dorsal PM, SMA, IPL, and a dorsal sector of SPL. These areas therefore, may be considered part of a wider system recruited during both execution and observation of actions performed with different biological effectors (MNS, Rizzolatti and Craighero, 2004; Fabbri-Destro and Rizzolatti, 2008; Hari and Kujala, 2009). So far, in the monkey there is no evidence of an observation/execution matching system specific for lower limb actions including walking while a mirror mechanism has been described for hand or mouth actions (Gallese et al., 1996; Rizzolatti et al., 1996; Ferrari et al., 2003; Rizzolatti and Craighero, 2004). As for humans, several imaging studies suggest that the MNS is not restricted to hand- and mouth-related actions but it extends also to cover foot actions (Buccino et al., 2001; Wheaton et al., 2004; Sakreida et al., 2005). The present data extend the MNS also to walking and are in line with previous studies assessing the neural substrates of walking imagery (Malouin et al., 2003; Sacco et al., 2006; Bakker et al., 2008; Jahn et al., 2008; Wang et al., 2008; la Fougere et al., 2010; van der Meulen et al., 2012). As a whole, they further support the notion that action execution, action observation and motor imagery share common neural structures (Jeannerod, 2001).

The frontal nodes of the MNS for walking are represented by the SMA and the dorsal PM. Previous studies reported that rostral SMA is particularly active in planning spatiotemporal aspects of action and in updating motor plans for temporally ordered subsequent movements (Roland et al., 1980; Tanji and Kurata, 1985; Shibasaki et al., 1993). The recruitment of SMA in the present study suggests a role for this area in providing proper sequencing and timing of limb movements during actual walking. Dorsal PM is endowed with a motor representation of lower limbs (Kurata, 1989; Godschalk et al., 1995) and a role

TABLE 1 | MNI coordinates of local maxima of the activation foci (conjunction null analysis).

AnatomicalRegion	(ATB*)	Left				(ATB*)	Right			
		x	y	z	Z-score		x	y	z	Z-score
(A) Walking observation/execution										
-Supplementary Motor Area (SMA)	(6)	−2	−18	76	4.68					
-SMA		−4	−38	74	4.85		6	−38	74	4.62
-Precentral Gyrus							38	−8	66	3.67
-Superior Parietal Lobule							32	−52	72	3.43
-Inferior Parietal Lobule	(PF)	−58	−40	48	3.84	(PF)	60	−32	46	5.08
-Cerebellum	(Vermis)	−4	−72	−42	4.62					
-Cerebellum		−26	−46	−28	3.95		20	−36	−28	4.31
(B)Narrow vs. open space walk										
-IntraParietal-TransverseOccipital (IPTO)		−34	−86	18	3.62					
-IPTO	(7, 18)	−14	−90	30	3.74		28	−86	22	3.65
-IPTO		−14	−86	38	3.77					
-Middle Occipital Gyrus (MOG)						(PG)	52	−70	12	3.80
(C)Open vs. narrow space walk										
-Inferior Occipital Gyrus	(17)	−14	−98	−8	4.18					

(A) Conjunction investigating common brain activations during walking execution and walking observation.
(B) and *(C)* Conjunctions which investigated walking in narrow and open space, respectively.
All reported local maxima were significant with p < 0.001 uncorrected and an extent threshold k = 10 voxels.
*ATB: most probable anatomical region in the Anatomy Toolbox 1.7, as reported by Eickhoff et al., 2005.

of this region in locomotor control has long been suggested (Freund and Hummelsheim, 1985). Dorsal PM, therefore, may exert the control of walking especially when it is guided by visual information. In the right hemisphere there also was a distinct dorsal premotor spot that largely coincides with the one described by Buccino et al. (2001) during the observation of foot actions, either object- or non-object-directed.

The posterior nodes of the MNS for walking are represented by sectors of the posterior parietal cortex. The dorsal portion of SPL is involved in integrating proprioceptive information related to the current body position into a motor plan (Andersen et al., 1997; de Lange et al., 2006) and in combining visual and somatosensory information in order to guide spatially directed movements (Andersen et al., 1997; Wenderoth et al., 2006). During walking execution, SPL activation may be related to the processing of visual and somatosensory feedback. During walking observation the same functional activation may represent a re-enactment of the sensory aspects of the observed action. Indeed a mirror mechanism has been described also for sensory information (Keysers et al., 2004; Ebisch et al., 2008; Gazzola and Keysers, 2009). It has been proposed that actual and imagined movements involve prediction of the sensory consequences of the action (Wolpert et al., 1998; Blakemore and Sirigu, 2003). This may be true also for observed actions. Bakker et al. (2008) reported activation of a dorsomedial sector of SPL closely corresponding to that of the present study when participants performed a motor imagery of walking along a narrow path compared to a broad path. The authors interpreted this finding as indicating that during motor imagery sensory information is generated in the absence of concurrent action production.

As for IPL, this cortical sector has long been involved in coding the pragmatic features of an object that are relevant for a

biological effector (for instance, the hand) in order to act properly upon it (Binkofski et al., 1999; Chao and Martin, 2000; Grèzes and Decety, 2002; Grèzes et al., 2003). We suggest that during walking IPL may code the interaction between foot and the surface on which individuals are requested to walk. In other words, this region may provide information on pragmatic features of the surface (for instance, the presence of holes or bumps) relevant for walking properly on it. It is noteworthy that in the present study while actually walking, participants had to interact with a rolling cylinder. In previous studies IPL has been shown to be involved in motor imagery (Malouin et al., 2003) and observation of walking (Iseki et al., 2008).

During both observation and execution of walking, we found a set of activation foci also in the cerebellum which included the vermis and both cerebellar hemispheres. Cerebellar activation during execution and imagery of walking (Fukuyama et al., 1997; la Fougere et al., 2010; van der Meulen et al., 2012) as well as during lower limb movements (Rocca and Filippi, 2010) was previously found. Despite the cerebellum is not considered as a node of the MNS, it might come into play whenever motor sequences such as walking are executed and/or observed (Molinari et al., 2008).

As far as space coding is concerned, our findings show that there are specific regions in the brain involved in the coding of a narrow space as compared to an open one. When considering hand-object interactions, pivotal electrophysiological studies in the monkey showed the existence of bimodal visual and tactile neurons that discharge when the objects are within a reachable distance. Such neurons have been identified in several regions of the monkey brain, including PM and parietal areas, and it has been forwarded that they code for a near space where it is possible for individuals to interact with objects (Rizzolatti

et al., 1997; Fogassi et al., 1999; Graziano, 2001). This space has been called *peripersonal* space to distinguish it from a far space irrelevant for action execution (extrapersonal space). In humans, brain imaging studies have confirmed the presence of a parieto-frontal circuit coding for peripersonal space around the face and hands (Bremmer et al., 2001; Sereno and Huang, 2006; Makin et al., 2007), with additional areas centered in superior parieto-occipital cortex (Quinlan and Culham, 2007; Gallivan et al., 2009) and lateral occipital cortex (Makin et al., 2007). In keeping with fMRI studies, behavioral studies have shown a clear distinction between a peripersonal and an extrapersonal space (for review see Spence et al., 2004, 2008; Brozzoli et al., 2012). All these studies focused on face/hand-object interactions in peripersonal space while less evidence is available for the coding of a motorically relevant space for foot actions. There is a suggestion that a peripersonal space representation would seem to be an efficient organizational principle not only for the upper but also for the lower limb (Graziano et al., 2002; Graziano and Cooke, 2006). In a behavioral study, for instance, Schicke et al. (2009) employed a cross modal congruency task and found that congruency effects did not differ between hand and foot suggesting a representation of peripersonal space also around the feet. It is worth stressing that in our study, when participants were required to observe a narrow space or to walk in it, they perceived this space as peripersonal since the different elements in the seen environment (for instance, the walls of the corridor) were at a reachable distance. In contrast, when participants were required to observe an open space or to walk in it, they perceived this space as extrapersonal since the different elements (for instance, trees and hills in the background) were perceived as too far to interact with them. In the present study, when participants had to process a near space, specific activations were found in the right MOG and in the IPTO bilaterally, but more largely represented in the left hemisphere.

The right MOG corresponds to the area found active by Bakker et al. (2008) in the imagery of walking along a narrow path as compared to a broad one. This area is close to the lateral occipital cortex labeled as *extrastriate body area* (EBA) by Downing et al. (2001) that is recruited during the observation of different body parts even when they imply little motion. Bakker et al. (2008) suggested that MOG activity reflects the generation of accurate predictions of the sensory (presumably visual) consequences of a specific motor plan. As in Bakker et al. (2008), in both space videos our participants did not observe any body parts so the explanation provided by these authors may also fit our data. In more general terms, we suggest that this activation may call for a visual description of the position of body parts including foot in our particular walking set. This interpretation is well supported by the fact that in the present experiment MOG activity appears stronger in the narrow space as compared to an open one. Indeed, in the narrow space a more accurate visual description of one's own body parts is a key requirement to interact properly with the environment. This means that extrastriate regions might be involved in generating visual imagery relevant to motor control (Toni et al., 2002; Astafiev et al., 2004; Helmich et al., 2007), as reported in other sensory domains (Blakemore and Sirigu, 2003).

As for the IPTO, Makin et al. (2007) showed a greater activation of this region in all conditions in which an object moved toward the participant's body (near space) independent of the fact that a biological effector will interact with the object or not. Activation of the IPTO was also found in a PET study by Binkofski et al. (2003) when participants had to reach for a virtual object in a mirror positioned in front of the observer. Since this was a PET study, these findings are not fully comparable with the present ones. Nevertheless, it is worth stressing that also in the condition described by Binkofski et al. (2003) the virtual object was perceived in the near space of participants. Moreover, according to the coordinates provided by Pisella et al. (2009), IPTO is part of a region potentially damaged in patients with optic ataxia (OA). Classically, OA is considered a disorder of reaching objects presented in peripheral vision following an impairment in visuomotor integration. Since IPTO appears to code for a near space as distinct from a far one, it is reasonable that a lesion of this area may contribute to the clinical picture of OA patients. Altogether, our findings and those of the literature reviewed so far strongly suggest a specific role for IPTO in distinguishing a near space from an open and far one. As Makin et al. (2007) found this sector during a task involving the hand while in the present study we consider a task specifically related to foot actions, it is most likely that the representation of near space coded in IPTO is independent of a specific biological effector and may constitute a preliminary processing of space subserving any further visuomotor transformation involving objects located in it.

Greater activation during processing the open space as compared to the narrow one was found in the primary visual cortex. The most likely explanation for this finding is that the landscape depicted in the open space video was full of different elements (grass, road, houses, street lamp, mountains, and so on) which made the scene more vivid and visually complex than the one depicted in the narrow space video.

In our opinion the present findings may have implications in the field of rehabilitation. Indeed, they show that observing walking actions, by triggering the MNS, is effective in recruiting sectors of the cortical motor system involved in the execution of the same motor tasks. As mentioned above, it has been shown that the motor system and in particular the MNS is involved in motor execution, action observation and motor imagery. For several years, motor imagery has been used in the rehabilitation practice and sports (Mulder, 2007). The recruitment of motor representations, driven by motor imagery, can improve the quality of motor performance, even in the absence of an actual execution of action. It has been proposed (Buccino et al., 2006; Small et al., 2013) that, similarly to motor imagery, the careful observation of actions made in an ecological context may be a valid approach in rehabilitation (action observation treatment, AOT), since even action observation has proven to be effective in recruiting the motor representations of the observed actions. During AOT, patients are required to carefully observe and soon afterward execute different daily actions presented through video-clips (Small et al., 2013; Buccino, 2014). So far, AOT has been successfully applied to chronic stroke patients (Ertelt et al., 2007) and to the recovery of daily activities in PD patients

(Buccino et al., 2011). Recently, a case-control study has been conducted on the efficacy of AOT in children affected by cerebral palsy (Buccino et al., 2012).

It is worth stressing that AOT has been shown to be effective also in the rehabilitation of lower limbs motor function. In detail, Pelosin et al. (2010) used AOT in the recovery of walking ability in PD patients with freezing of gait and Bellelli et al. (2010) in the rehabilitation of orthopedic patients that had undergone hip or knee replacement surgery. In our opinion, the findings of the present study, by showing the existence of an observation/execution matching system for lower limbs actions including walking, provide neurophysiological basis for this clinical evidence.

Furthermore, action observation of lower limbs actions has the potential to be exploited also in a variety of fields including educational activities and sport. In fact, to the best of our knowledge, while motor imagery has been used (for review, see Mulder, 2007) as a training strategy for athletes, even at competitive levels, the potential of AOT in this respect has never been systematically investigated.

Acknowledgment

References

Abdollahi, R. O., Jastorff, J., and Orban, G. A. (2013). Common and segregated processing of observed actions in human SPL. *Cereb. Cortex* 23, 2734–2753. doi: 10.1093/cercor/bhs264

Alkadhi, H., Brugger, P., Boendermaker, S. H., Crelier, G., Curt, A., Hepp-Reymond, M. C., et al. (2005). What disconnection tells about motor imagery: evidence from paraplegic patients. *Cereb. Cortex* 15, 131–140. doi: 10.1093/cercor/bhh116

Andersen, R. A., Snyder, L. H., Bradley, D. C., and Xing, J. (1997). Multimodal representation of space in the posterior parietal cortex and its use in planning movements. *Annu. Rev. Neurosci.* 200, 303–330. doi: 10.1146/annurev.neuro.20.1.303

Astafiev, S. V., Stanley, C. M., Shulman, G. L., and Corbetta, M. (2004). Extrastriate body area in human occipital cortex responds to the performance of motor actions. *Nat. Neurosci.* 7, 542–548. doi: 10.1038/nn1241

Bakker, M., de Lange, F. P., Helmich, R. C., Scheeringa, R., Bloem, B. R., and Toni, I. (2008). Cerebral correlates of motor imagery of normal and precision gait. *Neuroimage* 41, 998–1010. doi: 10.1016/j.neuroimage.2008.03.020

Bellelli, G., Buccino, G., Bernardini, B., Padovani, A., and Trabucchi, M. (2010). Action observation treatment improves recovery of postsurgical orthopedic patients: evidence for a top-down effect? *Arch. Phys. Med. Rehabil.* 91, 1489–1494. doi: 10.1016/j.apmr.2010.07.013

Binkofski, F., Buccino, G., Posse, S., Seitz, R. J., Rizzolatti, G., and Freund, H. (1999). A fronto-parietal circuit for object manipulation in man: evidence from an fMRI-study. *Eur. J. Neurosci.* 11, 3276–3286. doi: 10.1046/j.1460-9568.1999.00753.x

Binkofski, F., Butler, A., Buccino, G., Heide, W., Fink, G., Freund, H., et al. (2003). Mirror apraxia affects the peripersonal mirror space. A combined lesion and cerebral activation study. *Exp. Brain Res.* 153, 210–219.

Blakemore, S. J., and Sirigu, A. (2003). Action prediction in the cerebellum and in the parietal lobe. *Exp. Brain Res.* 153, 239–245. doi: 10.1007/s00221-003-1597-z

Bremmer, F., Schlack, A., Shah, N. J., Zafiris, O., Kubischik, M., Hoffmann, K., et al. (2001). Polymodal motion processing in posterior parietal and premotor cortex: a human fMRI study strongly implies equivalencies between humans and monkeys. *Neuron* 29, 287–296. doi: 10.1016/S0896-6273(01)00198-2

Brozzoli, C., Makin, T. R., Cardinali, L., Holmes, N. P., and Farnè, A. (2012). "Peripersonal space. A multisensory interface for body–object interactions," in *The Neural Bases of Multisensory Processes*, Chap. 23, eds M. M. Murray and M. T. Wallace (Boca Raton, FL: CRC Press).

Buccino, G. (2014). Action observation treatment: a novel tool in neurorehabilitation. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 369, 20130185. doi: 10.1098/rstb.2013.0185

Buccino, G., Arisi, D., Gough, P., Aprile, D., Ferri, C., Serotti, L., et al. (2012). Improving upper limb motor functions through action observation treatment: a pilot study in children with cerebral palsy. *Dev. Med. Child Neurol.* 54, 822–828. doi: 10.1111/j.1469-8749.2012.04334.x

Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., et al. (2001). Action observation activates premotor and parietal areas in

a somatotopic manner: an fMRI study. *Eur. J. Neurosci.* 13, 400–404. doi: 10.1046/j.1460-9568.2001.01385.x

Buccino, G., Gatti, R., Giusti, M. C., Negrotti, A., Rossi, A., Calzetti, S., et al. (2011). Action observation treatment improves autonomy in daily activities in Parkinson's disease patients: results from a pilot study. *Mov. Disord.* 26, 1963–1964. doi: 10.1002/mds.23745

Buccino, G., Solodkin, A., and Small, S. L. (2006). Functions of the mirror neuron system: implications for neurorehabilitation. *Cogn. Behav. Neurol.* 19, 55–63. doi: 10.1097/00146965-200603000-00007

Caggiano, V., Fogassi, L., Rizzolatti, G., Their, P., and Casile, A. (2009). Mirror neurons differentially encode the peripersonal and extrapersonal space of monkeys. *Science* 324, 403–406. doi: 10.1126/science.1166818

Chao, L. L., and Martin, A. (2000). Representation of manipulable man-made objects in the dorsal stream. *Neuroimage* 12, 478–484. doi: 10.1006/nimg.2000.0635

Christensen, L. O., Johannsen, P., Sinkjaer, T., Petersen, N., Pyndt, H. S., and Nielsen, J. B. (2000). Cerebral activation during bicycle movements in man. *Exp. Brain Res.* 135, 66–72. doi: 10.1007/s002210000493

Costantini, M., Ambrosini, E., Tieri, G., Sinigaglia, C., and Committeri, G. (2010). Where does an object trigger an action? An investigation about affordances in space. *Exp. Brain Res.* 207, 95–103. doi: 10.1007/s00221-010-2435-8

Cramer, S. C., Lastra, L., Lacourse, M. G., and Cohen, M. J. (2005). Brain motor system function after chronic, complete spinal cord injury. *Brain* 128, 2941–2950. doi: 10.1093/brain/awh648

Cramer, S. C., Orr, E. L., Cohen, M. J., and Lacourse, M. G. (2007). Effects of motor imagery training after chronic, complete spinal cord injury. *Exp. Brain Res.* 177, 233–242. doi: 10.1007/s00221-006-0662-9

Decety, J. (1996). Do imagined and executed actions share the same neural substrate? *Brain Res. Cogn. Brain Res.* 3, 87–93. doi: 10.1016/0926-6410(95)00033-X

de Lange, F. P., Helmich, R. C., and Toni, I. (2006). Posture influences motor imagery: an fMRI study. *Neuroimage* 33, 609–617. doi: 10.1016/j.neuroimage.2006.07.017

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. doi: 10.1126/science.1063414

Ebisch, S. J., Perrucci, M. G., Ferretti, A., Del Gratta, C., Romani, G. L., and Gallese, V. (2008). The sense of touch: embodied simulation in a visuotactile mirroring mechanism for observed animate or inanimate touch. *J. Cogn. Neurosci.* 20, 1611–1623. doi: 10.1162/jocn.2008.20111

Eickhoff, S. B., Stephan, K. E., Mohlberg, H., Grefkes, C., Fink, G. R., Amunts, K., et al. (2005). A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *Neuroimage* 25, 1325–1335. doi: 10.1016/j.neuroimage.2004.12.034

Enzinger, C., Ropele, S., Fazekas, F., Loitfelder, M., Gorani, F., Seifert, T., et al. (2008). Brain motor system function in a patient with complete spinal cord injury following extensive brain-computer interface training. *Exp. Brain Res.* 190, 215–223. doi: 10.1007/s00221-008-1465-y

Ertelt, D., Small, S., Solodkin, A., Dettmers, C., McNamara, A., Binkofski, F., et al. (2007). Action observation has a positive impact on

rehabilitation of motor deficits after stroke. *Neuroimage* 36, 164–173. doi: 10.1016/j.neuroimage.2007.03.043

Fabbri-Destro, M., and Rizzolatti, G. (2008). Mirror neurons and mirror systems in monkeys and humans. *Physiology (Bethesda)* 23, 171–179. doi: 10.1152/physiol.00004.2008

Ferrari, P. F., Gallese, V., Rizzolatti, G., and Fogassi, L. (2003). Mirror neurons responding to the observation of ingestive and communicative mouth actions in the monkey ventral premotor cortex. *Eur. J. Neurosci.* 17, 1703–1714. doi: 10.1046/j.1460-9568.2003.02601.x

Fogassi, L., and Luppino, G. (2005). Motor functions of the parietal lobe. *Curr. Opin. Neurobiol.* 15, 626–631. doi: 10.1016/j.conb.2005.10.015

Fogassi, L., Raos, V., Franchi, G., Gallese, V., Luppino, G., and Matelli, M. (1999). Visual responses in the dorsal premotor area F2 of the macaque monkey. *Exp. Brain Res.* 128, 194–199. doi: 10.1007/s002210050835

Freund, H. J., and Hummelsheim, H. (1985). Lesions of premotor cortex in man. *Brain* 108, 697–733. doi: 10.1093/brain/108.3.697

Friston, K. J., Holmes, A. P., Price, C. J., Büchel, C., and Worsley, K. J. (1999). Multisubject fMRI studies and conjunction analyses. *Neuroimage* 10, 385–396. doi: 10.1006/nimg.1999.0484

Friston, K. J., Penny, W. D., and Glaser, D. E. (2005). Conjunction revisited. *Neuroimage* 25, 661–667. doi: 10.1016/j.neuroimage.2005.01.013

Fukuyama, H., Ouchi, Y., Matsuzaki, S., Nagahama, Y., Yamauchi, H., Ogawa, M., et al. (1997). Brain functional activity during gait in normal subjects: a SPECT study. *Neurosci. Lett.* 228, 183–186. doi: 10.1016/S0304-3940(97)00381-9

Gallese, V., Fadiga, L., Fogassi, L., and Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain* 119, 593–609. doi: 10.1093/brain/119.2.593

Gallese, V., and Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends Cogn. Sci.* 2, 493–501. doi: 10.1016/S1364-6613(98)01262-5

Gallivan, J. P., Cavina-Pratesi, C., and Culham, J. C. (2009). Is that within reach? fMRI reveals that the human superior parieto-occipital cortex encodes objects reachable by the hand. *J. Neurosci.* 29, 4381–4391. doi: 10.1523/JNEUROSCI.0377-09.2009

Gazzola, V., and Keysers, C. (2009). The observation and execution of actions share motor and somatosensory voxels in all tested subjects: single-subject analyses of unsmoothed fMRI data. *Cereb. Cortex* 19, 1239–1255. doi: 10.1093/cercor/bhn181

Gerardin, E., Sirigu, A., Lehéricy, S., Poline, J. B., Gaymard, B., Marsault, C., et al. (2000). Partially overlapping neural networks for real and imagined hand movements. *Cereb. Cortex* 10, 1093–1104. doi: 10.1093/cercor/10.11.1093

Godschalk, M., Mitz, A. R., van Duin, B., and van der Burg, H. (1995). Somatotopy of monkey premotor cortex examined with microstimulation. *Neurosci. Res.* 23, 269–279. doi: 10.1016/0168-0102(95)00950-7

Graziano, M. S. (2001). A system of multimodal areas in the primate brain. *Neuron* 29, 4–6. doi: 10.1016/S0896-6273(01)00174-X

Graziano, M. S. A., and Cooke, D. F. (2006). Parieto-frontal interactions, personal space, and defensive behavior. *Neuropsychologia* 44, 2621–2635. doi: 10.1016/j.neuropsychologia.2005.09.011

Graziano, M. S. A., Taylor, C. S., Moore, T., and Cooke, D. F. (2002). The cortical control of movement revisited. *Neuron* 3, 349–362. doi: 10.1016/S0896-6273(02)01003-6

Greenstein, J., Gastineau, E., Siegel, B., Macsata, R., Conklin, J., and Maurer, A. (1995). Cerebral hemisphere activation during human bipedal locomotion. *Hum. Brain Mapp.* 320.

Grèzes, J., and Decety, J. (2002). Does visual perception of object afford action? Evidence from a neuroimaging study. *Neuropsychologia* 4, 212–222.

Grèzes, J., Tucker, M., Armony, J., Ellis, R., and Passingham, R. E. (2003). Objects automatically potentiate action: an fMRI study of implicit processing. *Eur. J. Neurosci.* 17, 2735–2740. doi: 10.1046/j.1460-9568.2003.02695.x

Gustin, S. M., Wrigley, P. J., Henderson, L. A., and Siddall, P. J. (2010). Brain circuitry underlying pain in response to imagined movement in people with spinal cord injury. *Pain* 148, 438–445. doi: 10.1016/j.pain.2009.12.001

Hari, R., and Kujala, M. V. (2009). Brain basis of human social interaction: from concepts to brain imaging. *Physiol. Rev.* 89, 453–479. doi: 10.1152/physrev.00041.2007

Hashimoto, T. (2006). Speculation on the responsible sites and pathophysiology of freezing of gait. *Parkinsonism Relat. Disord.* 12, S55–S62. doi: 10.1016/j.parkreldis.2006.05.017

Helmich, R. C., de Lange, F. P., Bloem, B. R., and Toni, I. (2007). Cerebral compensation during motor imagery in Parkinson's disease. *Neuropsychologia* 45, 2201–2215. doi: 10.1016/j.neuropsychologia.2007.02.024

Hotz-Boendermaker, S., Funk, M., Summers, P., Brugger, P., Hepp-Reymond, M. C., Curt, A., et al. (2008). Preservation of motor programs in paraplegics as demonstrated by attempted and imagined foot movements. *Neuroimage* 39, 383–394. doi: 10.1016/j.neuroimage.2007.07.065

Iacoboni, M., and Dapretto, M. (2006). The mirror neuron system and the consequences of its dysfunction. *Nat. Rev. Neurosci.* 7, 942–951. doi: 10.1038/nrn2024

Iseki, K., Hanakawa, T., Shinozaki, J., Nankaku, M., and Fukuyama, H. (2008). Neural mechanisms involved in mental imagery and observation of gait. *Neuroimage* 41, 1021–1031. doi: 10.1016/j.neuroimage.2008.03.010

Ishii, K., Senda, M., Toyama, H., Oda, D., Ishii, K., Ishiwata, K., et al. (1995). Brain function in bipedal gait: a PET study. *Hum. Brain Mapp.* 3:s321.

Jahn, K., Deutschländer, A., Stephan, T., Kalla, R., Wiesmann, M., Strupp, M., et al. (2008). Imaging human supraspinal locomotor centers in brainstem and cerebellum. *Neuroimage* 39, 786–792. doi: 10.1016/j.neuroimage.2007.09.047

Jeannerod, M. (2001). Neural simulation of action: a unifying mechanism for motor cognition. *Neuroimage* 14, S103–S109. doi: 10.1006/nimg.2001.0832

Johansson, G., von Hofsten, C., and Jansson, G. (1980). Event perception. *Annu. Rev. Psychol.* 31, 27–63. doi: 10.1146/annurev.ps.31.020180.000331

Keysers, C., Wicker, B., Gazzola, V., Anton, J. L., Fogassi, L., and Gallese, V. (2004). A touching sight: SII/PV activation during the observation and experience of touch. *Neuron* 42, 335–346. doi: 10.1016/S0896-6273(04)00156-4

Kurata, K. (1989). Distribution of neurons with set- and movement-related activity before hand and foot movements in the premotor cortex of rhesus monkey. *Exp. Brain Res.* 77, 245–256. doi: 10.1007/BF00274982

Lafleur, M., Jackson, P., Malouin, F., Richards, C. L., Evans, A. C., and Doyon, J. (2002). Motor learning produces parallel dynamic functional changes during the execution and imagination of sequential foot movements. *Neuroimage* 16, 142–157. doi: 10.1006/nimg.2001.1048

la Fougere, C., Zwergal, A., Rominger, A., Foerster, S., Fesl, G., Dieterich, M., et al. (2010). Real versus imagined locomotion: a [18F]-FDG PET-fMRI comparison. *Neuroimage* 50, 1589–1598. doi: 10.1016/j.neuroimage.2009.12.060

Maffei, V., Giusti, M. A., Macaluso, E., Laquaniti, F., and Viviani, P. (2015). Unfamiliar walking movements are detected early in the visual stream: an fMRI study. *Cereb. Cortex* 25, 2022–2034. doi: 10.1093/cercor/bhu008

Makin, T. R., Holmes, N. P., and Zohary, E. (2007). Is that near my hand? Multisensory representation of peripersonal space in human intraparietal sulcus. *J. Neurosci.* 27, 731–740.

Malouin, F., Richards, C. L., Jackson, P. L., Dumas, F., and Doyon, J. (2003). Brain activations during motor imagery of locomotor-related tasks: a PET study. *Hum. Brain Mapp.* 19, 47–62. doi: 10.1002/hbm.10103

Miyai, I., Tanabe, H. C., Sase, I., Eda, H., Oda, I., Konishi, I., et al. (2001). Cortical mapping of gait in humans: a near-infrared spectroscopic topography study. *Neuroimage* 14, 1186–1192. doi: 10.1006/nimg.2001.0905

Molinari, M., Chiricozzi, F. R., Clausi, S., Tedesco, A. M., De Lisa, M., and Leggio, M. G. (2008). Cerebellum and detection of sequences, from perception to cognition. *Cerebellum* 7, 611–615. doi: 10.1007/s12311-008-0060-x

Mulder, T. (2007). Motor imagery and action observation: cognitive tools for rehabilitation. *J. Neural. Transm.* 114, 1265–1278. doi: 10.1007/s00702-007-0763-z

Nichols, T., Brett, M., Andersson, J., Wager, T., and Poline, J. B. (2005). Valid conjunction inference with the minimum statistic. *Neuroimage* 25, 653–660. doi: 10.1016/j.neuroimage.2004.12.005

Orr, E. L., Lacourse, M. G., Cohen, M. J., and Cramer, S. C. (2008). Cortical activation during executed, imagined, and observed foot movements. *Neuroreport* 19, 625–630. doi: 10.1097/WNR.0b013e3282fbf9e0

Pavlova, M. A. (2012). Biological motion processing as a hallmark of social cognition. *Cereb. Cortex* 22, 981–995. doi: 10.1093/cercor/bhr156

Pelosin, E., Avanzino, L., Bove, M., Stramesi, P., Nieuwboer, A., and Abbruzzese, G. (2010). Action observation improves freezing of gait in patients with Parkinson's disease. *Neurorehabil. Neural Repair* 24, 746–752. doi: 10.1177/1545968310368685

Pisella, L., Sergio, L., Blangero, A., Torchin, H., Vighetto, A., and Rossetti, Y. (2009). Optic ataxia and the function of the dorsal stream: contributions

to perception and action. *Neuropsychologia* 47, 3033–3044. doi: 10.1016/j.neuropsychologia.2009.06.020

Porro, C. A., Francescato, M. P., Cettolo, V., Diamond, M. E., Baraldi, P., Zuiani, C., et al. (1996). Primary motor and sen sory cortex activation during motor performance and motor imagery: a functional magnetic resonance imaging study. *J. Neurosci.* 16, 7688–7698.

Quinlan, D. J., and Culham, J. C. (2007). fMRI reveals a preference for near viewing in the human parietooccipital cortex. *Neuroimage* 36, 167–187. doi: 10.1016/j.neuroimage.2007.02.029

Rizzolatti, G., and Craighero, L. (2004). The mirror-neuron system. *Annu. Rev. Neurosci.* 27, 169–192. doi: 10.1146/annurev.neuro.27.070203.144230

Rizzolatti, G., Fadiga, L., Fogassi, L., and Gallese, V. (1996). Premotor cortex and the recognition of motor actions. *Brain Res.* 3, 131–141.

Rizzolatti, G., Fadiga, L., Fogassi, L., and Gallese, V. (1997). The space around us. *Science* 11, 190–191. doi: 10.1126/science.277.5323.190

Rocca, M. A., and Filippi, M. (2010). FMRI correlates of execution and observation of foot movements in left-handers. *J. Neurol. Sci.* 288, 34–41. doi: 10.1016/j.jns.2009.10.013

Roland, P. E., Larsen, B., Lassen, N. A., and Skinhøj, E. (1980). Supplementary motor area and other cortical areas in organization of voluntary movements in man. *J. Neurophysiol.* 43, 118–136.

Sacco, K., Cauda, F., Cerliani, L., Mate, D., Duca, S., and Geminiani, G. C. (2006). Motor imagery of walking following training in locomotor attention. The effect of the "tango lesson". *Neuroimage* 32, 1441–1449.

Sakreida, K., Schubotz, R. I., Wolfensteller, U., and von Cramon, D. Y. (2005). Motion class dependency in observers' motor areas revealed by functional magnetic resonance imaging. *J. Neurosci.* 25, 1335–1342. doi: 10.1523/JNEUROSCI.4170-04.2005

Salsone, M., Bagnato, A., Novellino, F., Cascini, G. L., Paglionico, S., Cipullo, S., et al. (2009). Cardiac MIBG scintigraphy in primary progressive freezing gait. *Parkinsonism Relat. Disord.* 15, 365–369. doi: 10.1016/j.parkreldis.200 8.08.008

Schicke, T., Bauer, F., and Röder, B. (2009). Interactions of different body parts in peripersonal space: how vision of the foot influences tactile perception at the hand. *Exp. Brain Res.* 192, 703–715. doi: 10.1007/s00221-00 8-1587-2

Sereno, M. I., and Huang, R. S. (2006). A human parietal face area contains aligned head-centered visual and tactile maps. *Nat. Neurosci.* 9, 1337–1343. doi: 10.1038/nn1777

Shibasaki, H., Sadato, N., Lyshkow, H., Yonekura, Y., Honda, M., Nagamine, T., et al. (1993). Both primary motor cortex and supplementary motor area play an important role in complex finger movement. *Brain* 116, 1387–1398.

Small, S. L., Buccino, G., and Solodkin, A. (2013). Brain repair after stroke–a novel neurological model. *Nat. Rev. Neurol.* 9, 698–707. doi: 10.1038/nrneurol.2013.222

Solodkin, A., Hlustik, P., Chen, E. E., and Small, S. L. (2004). Fine modulation in network activation during motor execution and motor imagery. *Cereb. Cortex* 14, 1246–1255. doi: 10.1093/cercor/bhh086

Spence, C., Pavani, F., Maravita, A., and Holmes, N. (2004). Multisensory contributions to the 3-D representation of visuotactile peripersonal space in humans: evidence from the cross modal congruency task. *J. Physiol. Paris* 98, 171–189. doi: 10.1016/j.jphysparis.2004.03.008

Spence, C., Pavani, F., Maravita, A., and Holmes, N. P. (2008). "Multisensory interactions," in *Haptic Rendering: Foundations, Algorithms, and Applications*, eds M. C. Lin and M. A. Otaduy (Wellesley, MA: A.K. Peters Ltd), 21–52.

Stephan, K. M., Fink, G. R., Passingham, R. E., Silbersweig, D., Ceballos-Baumann, A. O., Frith, C. D., et al. (1995). Functional anatomy of the mental representation of upper extremity movements in healthy subjects. *J. Neurophys.* 73, 373–386.

Tanji, J., and Kurata, K. (1985). Contrasting neuronal activity in supplementary and precentral motor cortex of monkeys. I. Responses to instructions determining motor responses to forthcoming signals of different modalities. *J. Neurophysiol.* 53, 129–141.

Tashiro, M., Itoh, M., Fujimoto, T., Fujiwara, T., Ota, H., Kubota, K., et al. (2001). 18F-FDG PET mapping of regional brain activity in runners. *J. Sports Med. Phys. Fitness* 41, 11–17.

Toni, I., Shah, N. J., Fink, G. R., Thoenissen, D., Passingham, R. E., and Zilles, K. (2002). Multiple movement representations in the human brain: an event-related fMRI study. *J. Cogn. Neurosci.* 14, 769–784. doi: 10.1162/08989290260138663

van der Meulen, M., Allali, G., Rieger, S. W., Assal, F., and Vuilleumier, P. (2012). The influence of individual motor imagery ability on cerebral recruitment during gait imagery. *Hum. Brain Mapp.* 35, 455–470. doi: 10.1002/hbm.22192

Wang, C., Wai, Y., Kuo, B., Yeh, Y.-Y., and Wang, J. (2008). Cortical control of gait in healthy humans: an fMRI study. *J. Neural. Transm.* 115, 1149–1158. doi: 10.1007/s00702-008-0058-z

Wenderoth, N., Toni, I., Bedeleem, S., Debaere, F., and Swinnen, S. P. (2006). Information processing in human parieto-frontal circuits during goal-directed bimanual movements. *Neuroimage* 301, 264–278. doi: 10.1016/j.neuroimage.2005.11.033

Wheaton, K. J., Thompson, J. C., Syngeniotis, A., Abbott, D. F., and Puce, A. (2004). Viewing the motion of human body parts activates different regions of premotor, temporal, and parietal cortex. *Neuroimage* 22, 277–288. doi: 10.1016/j.neuroimage.2003.12.043

Wolpert, D. M., Goodbody, S. J., and Husain, M. (1998). Maintaining internal representations: the role of the human superior parietal lobe. *Nat. Neurosci.* 1, 529–533. doi: 10.1038/2245

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 SM cortex: an EEG and fMRI study. *Neuroimage* 49, 2596–2606. doi: 10.1016/j.neuroimage.2009.10.028

The influence of object shape and center of mass on grasp and gaze

Loni Desanghere[1,2]* and Jonathan J. Marotta[1]

[1] *Perception and Action Laboratory, Department of Psychology, University of Manitoba, Winnipeg, MB, Canada,*
[2] *Postgraduate Medical Education, College of Medicine, University of Saskatchewan, Saskatoon, SK, Canada*

Recent experiments examining where participants look when grasping an object found that fixations favor the eventual index finger landing position on the object. Even though the act of picking up an object must involve complex high-level computations such as the visual analysis of object contours, surface properties, knowledge of an object's function and center of mass (COM) location, these investigations have generally used simple symmetrical objects – where COM and horizontal midline overlap. Less research has been aimed at looking at how variations in object properties, such as differences in curvature and changes in COM location, affect visual and motor control. The purpose of this study was to examine grasp and fixation locations when grasping objects whose COM was positioned to the left or right of the objects horizontal midline (Experiment 1) and objects whose COM was moved progressively further from the midline of the objects based on the alteration of the object's shape (Experiment 2). Results from Experiment 1 showed that object COM position influenced fixation locations and grasp locations differently, with fixations not as tightly linked to index finger grasp locations as was previously reported with symmetrical objects. Fixation positions were also found to be more central on the non-symmetrical objects. This difference in gaze position may provide a more holistic view, which would allow both index finger and thumb positions to be monitored while grasping. Finally, manipulations of COM distance (Experiment 2) exerted marked effects on the visual analysis of the objects when compared to its influence on grasp locations, with fixation locations more sensitive to these manipulations. Together, these findings demonstrate how object features differentially influence gaze vs. grasp positions during object interaction.

Keywords: visuomotor control, fixations, gaze locations, grasp locations, irregular non-symmetrical objects

Introduction

We move about and interact with objects in our environment so effortlessly that the complexities of these interactions are rarely noticed. Although the integration of various senses, such as visual and tactile feedback when locating and picking up objects and vestibular information for balance (for review see Kandel et al., 2000), plays key roles in our interactions, we primarily rely on our sense of vision to accurately carry out our movements, with eye movements typically preceding hand movements in both pointing (Abrams et al., 1990; Bekkering et al., 1994; van Donkelaar et al., 2004) and object manipulation tasks (Land et al., 1999; Johansson et al., 2001; Land and Hayhoe, 2001; Hayhoe et al., 2003; Hayhoe and Ballard, 2005). When you want to pick up an object, it is usually a simple matter to look where you remember leaving it, reach out to its location, and

accurately pick it up. However, we do not typically grasp objects in arbitrary locations or look at random parts of that object during these actions. Instead, how we interact with an object depends on both the ongoing and planned behavior of the individual (Hamed et al., 2002) and the cognitive demands of the task (Yarbus, 1967).

For example, where we look on an object will vary depending on whether we are directly interacting with that object (e.g., picking it up; Johansson et al., 2001; de Grave et al., 2008; Brouwer et al., 2009; Desanghere and Marotta, 2011; Prime and Marotta, 2013), performing a series of movements (e.g., making sandwiches or tea; Ballard et al., 1992; Smeets et al., 1996; Land et al., 1999; Hayhoe et al., 2003), or completing a perceptual task such as visual search (Findlay, 1997; Zelinsky et al., 1997; Klein, 2000; Araujo et al., 2001), viewing objects (He and Kowler, 1991; Kowler and Blaser, 1995; McGowan et al., 1998; Melcher and Kowler, 1999; Vishwanath et al., 2000; Vishwanath and Kowler, 2003, 2004), or size estimation (Desanghere and Marotta, 2011). Indeed, visuomotor control is a complex, interactive process between perception and action; reflective in the separate yet interconnected neural pathways dedicated to these processes (for review see Milner and Goodale, 1995; Goodale, 1998, 2014; Schenk and McIntosh, 2010). Not only do we use vision to identify objects with which we are interacting, vision also provides us with feedback about the approaching hand (toward the object) to enable online corrections (Binsted et al., 2001; Riek et al., 2003), as well as provides information about where the contact location on the object is relative to the arm's motor system (Land and Hayhoe, 2001; Soechting et al., 2001). Research has shown that eye movements are typically initiated toward the object 40–100 ms prior to movement onset (Prablanc et al., 1979; Biguer et al., 1982; Land et al., 1999) with fixations linked to where participants grasp an object (i.e., they look at the location where they place their index finger during a precision grasp; Brouwer et al., 2009; Desanghere and Marotta, 2011; Prime and Marotta, 2013), and, when manipulating the objects, linked to forthcoming grasp sites, obstacles, and landing sites where objects are subsequently grasped, moved around, and placed, respectively (Johansson et al., 2001).

How fixations change throughout a reach-to-grasp movement and exactly what object properties are fixated has yet to be fully explored. A growing body of research has investigated *where* on an object people are fixating during basic reaching and grasping movements to simple objects. For example, de Grave et al. (2008) examined fixation locations during a reaching and grasping task to objects that were either fully visible or that had the index finger, thumb, or both grasp locations partially occluded. They found that first and second fixations on the objects were above the object's center of mass (COM), as well as above the visible COM (calculation of the COM based on the visible surface area of the object) in the case of partly occluded objects. In both instances fixation locations were toward index finger grasp location. However, similar to Johansson et al. (2001) where participants were instructed to grasp the object at a specific location, participants in this experiment also had specific grasp locations where they had to place their index finger and thumb on relatively simple objects such as a triangle or cross. In a

later investigation using similar objects, Brouwer et al. (2009) contrasted fixation locations when participants were reaching out to grasp an object vs. when they were asked to simply view that object. They found that during first fixations to the objects there was no difference between tasks; during both grasping and viewing participants were looking closer to the COM. During second fixations, however, fixations while grasping were found to be significantly higher up on the objects (toward index finger location) than those found during viewing. Desanghere and Marotta (2011) and later Prime and Marotta (2013) showed the opposite pattern in fixations when grasping simple symmetrical objects (squares and rectangles), with fixation locations first directed towards the grasp site for the index finger, and then directed lower toward the object's COM just prior to contact with the object. In these experiments, participant's grasp axis (the imaginary line joining the contact points of the index finger and thumb on the object) coincided with the object's COM location.

These studies all suggest that where we look on an object plays a key role in real-time grasping movements, with grasp and gaze locations sensitive to COM location and linked to the eventual index finger grasp location. In everyday life; however, we are not always reaching out to grasp simple symmetrical objects. Our visuomotor system must deal with complex calculations such as surface curvature and differences in an object's COM. Indeed, the role of vision in grasping is not only to activate appropriate grasp schemas, but also to determine accurate positioning of the fingers on the objects (Jeannerod et al., 1995; Smeets and Brenner, 1999). However, it is not always a simple relationship between fixation location and focal attention (i.e., what someone is concentrating on), as visual-spatial attention can be directed either overtly by actively fixating the eyes (the fovea) onto a specific area or covertly, by allocating cognitive resources to process information that is located in another region of space (Posner, 1980; Irwin, 2004). In this way, we are able to successfully reach out and pick up objects outside of foveal vision. Despite this; however, research has shown that when reaching out to pick up objects (Brouwer et al., 2009; Desanghere and Marotta, 2011; Prime and Marotta, 2013), participants' fixation locations do not stray off the object and are linked to the most relevant dimensions of the objects (index finger grasp location during a precision grasp), even in instances where digit trajectories of the thumb are made more variable (Cavina-Pratesi and Hesse, 2013). In other words, despite participants' ability to fixate any part of the desired object or any part of the scene in front of them and still complete the task, participants in these experiments reliably fixate the eventual index finger grasp location on the object with which they are interacting.

As we can see, the relationship between index finger grasp location and fixation locations are well documented, demonstrating a reliable relationship between grasp locations and focal attention. However, the effects that object complexity and changes in COM may have on eye-hand coordination are under-explored. Although it has been shown that alterations in object shape affect various reaching and grasping kinematics such as minimum grip force (Jenmalm et al., 1998), maximum grip aperture (Eloka and Franz, 2011), and digit placement (Jeannerod, 1988; Goodale et al., 1994; Lederman and Wing,

2003; Marotta et al., 2003; Kleinholdermann et al., 2007), the relationship between where we look relative to where we grasp on irregular objects remains unclear. Given that current research suggests that both the relevant dimension of an object and the overall shape of the object facilitate ongoing movements during reaching and grasping tasks (Eloka and Franz, 2011), further investigations between eye-hand coordination when grasping more complex objects needs to be carried out.

The primary goal of this research was to examine the variability in fixation locations across several conditions on irregular shaped objects and explore the relationship of these fixation locations to grasp locations on those same objects. In Experiment 1, we investigated fixation and grasp locations to contoured objects that had an asymmetrical design and compared this behavior to grasps made to symmetrical objects (squares and rectangles). Both object types had identical maximum horizontal and vertical dimensions. In Experiment 2, we explored the effects of object shape and changes in COM location on fixation and grasp locations. Based on the literature it was expected that participant's grasp axis and fixation locations would be shifted away from the horizontal center of the objects, toward the object's COM (Experiments 1 and 2), and this distance would increase with increases in COM locations from that point (Experiment 2), and in both instances be linked to the eventual index finger grasp location. If object irregularity is affecting visual analysis of the object, we expect to see differences in fixation locations relative to symmetrical objects (Experiment 1) and grasp locations (Experiments 1 and 2). If fixation locations are identical across conditions (e.g., linked to index finger location and not influenced by object irregularity) this would suggest that the visual analysis of irregular shaped objects during object manipulation is taking in only the most relevant dimensions of the objects needed for grasping.

Experiment 1

Materials and Methods

Participants

Fourteen undergraduate psychology students (nine female) between the ages of 18 and 30 ($M = 21$ years-old) were recruited for participation in this study. All participants were shown to be strongly right-handed as determined by a modified version of the Edinburg Handedness Inventory (Oldfield, 1971), and had normal or corrected-to-normal-vision. This research was approved by the Psychology/Sociology Human Research Ethics Board (PSREB) at the University of Manitoba.

Stimulus and Procedure

Participants were instructed to reach out "quickly but naturally" with their index finger and thumb and pick up randomly inter-mixed symmetrical (Efron blocks; Efron, 1968) and asymmetrical (Blake shapes; Blake, 1992) objects (all blocks weighed approximately 10 g). The Efron blocks differed in shape but were equal in surface area, and had the following horizontal and vertical dimensions: (1) 15.2 cm × 4.2 cm, (2) 12.2 cm × 5.2 cm, (3) 10.2 cm × 6.2 cm, (4) 9.0 cm × 7.1 cm,

and (5) 8.0 cm × 8.0 cm (see **Figure 1**). The Blake shapes had smoothly bounded contours and lacked clear symmetry (see **Figure 1**). Thus, the positioning of stable grasp locations on asymmetrical Blake shapes requires the analysis of the entire shape (Goodale et al., 1994). Each of the five Efron blocks were matched with Blake shapes with identical maximum vertical and
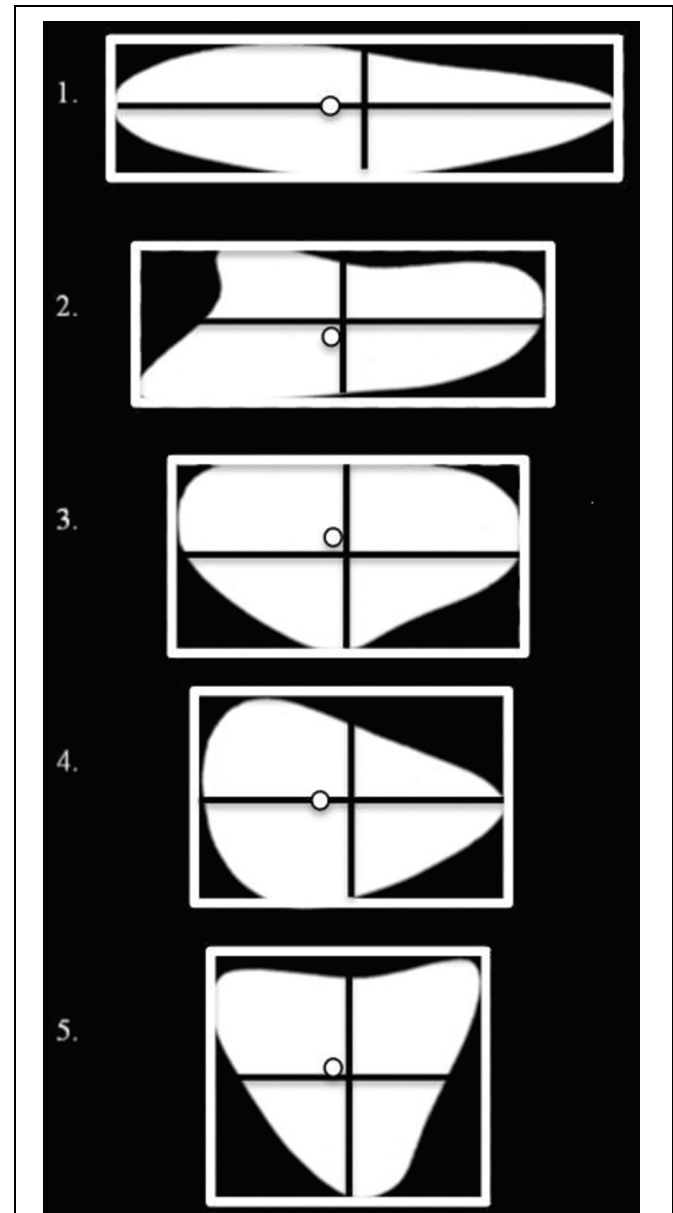

FIGURE 1 | Displays the Efron Shapes (dashed lines) and Blake shapes used in Experiment 1. The Blake shapes in this Figure demonstrate COML objects. Maximum horizontal and vertical object dimensions are as follows: (1) 15.2 cm × 4.2 cm, (2) 12.2 cm × 5.2 cm, (3) 10.2 cm × 6.2 cm, (4) 9.0 cm × 7.1 cm, (5) 8.0 cm × 8.0cm. The COM location for each asymmetrical object is represented by a white circle on the object, with the following X and Y distances relative to the center of the object which is represented by the intersection of the black lines (1) 1 cm, 0 cm; (2) 0.3 cm, −0.5 cm; (3) 0.03 cm, 0.05 cm; (4) 1 cm, 0 cm; (5) 0.5 cm, 0.3 cm.

horizontal dimensions. In addition, each Blake shape was either presented with the COM oriented to the left or right of the object's horizontal midline (objects were the mirror image of themselves), resulting in two asymmetrical object groups: Blake shapes with their COM shifted to the left (COML), and Blake shapes with their COM shifted to the right (COMR; see **Figure 1** for COM positions), and a third symmetrical group of Efron blocks.

Reach-to-grasp movements were recorded with an Optotrak Certus 3-D recording system (150 Hz sampling rate, spatial accuracy up to 0.01 mm; Northern Digital, Waterloo, ON, Canada). Two IREDs were fastened onto the participants' index finger (positioned on the left side of the cuticle), thumb (positioned on the right side of the cuticle), and wrist (positioned on the radial portion of the wrist) of their right hand. An Eye-link II (250 Hz sampling rate, spatial resolution <0.5°; SR Research Ltd., Osgoode, ON, Canada) was used to record eye-movements in both tasks. Kinematic information from both the Optotrak and the Eyelink II was integrated into a common frame of reference via MotionMonitor software (Innovative Sports technology, Chicago, IL, USA). The Motion Monitor system integrates eye, head, and hand data in a common frame of reference.

Both eyes were calibrated using a nine point calibration/validation procedure on the computer monitor, after which, a black display board was fastened over the calibration area. The asymmetrical and symmetrical objects were suspended via small magnets to this display during the grasping paradigm. To ensure accurate calibrations of less than 1° error and reliability of binocular eye data, accuracy checks both immediately following calibration and after the completion of the experiment were taken by having participants fixate a marker on the display while positional eye data was obtained.

At the start of each trial, participants held their right hand stationary on the start button with their index finger and thumb together and their eyes closed. The experimenter signaled the beginning of each trial with verbal instructions for the participant to open their eyes. At that point, participants reached out as quickly, but as naturally as they could, and grabbed the object with their index finger and thumb (objects were positioned 30 cm from the start button), and placed it on the table in front of them. Each grasp was positioned vertically, such that participant's index finger made contact with the object's top edge, and their thumb made contact along the bottom edge. After the completion of the trial, participants returned to the starting position with their index finger and thumb on the start button and their eyes closed, until given the verbal command to start the next trial.

The shapes were always presented with their longest axis on the horizontal plane such that, for the asymmetrical objects, their COM was oriented to the left or right of the actual center of the object on any given trial. For the purpose of these experiments, the 'center' of an object corresponds to the horizontal and vertical midpoint of each object, the halfway point calculated from the maximum horizontal and vertical dimensions of the block. The COM refers to the point where all of the mass of the object is concentrated, based on the averaging of the surface area of the objects (see **Figure 1**). Each object was suspended in such a manner that every object's vertical and horizontal center was

aligned with the board's center. All horizontal (X-axis) and vertical (Y-axis) coordinates were calculated from this location. For example, negative fixation or grasp locations represented locations to the left of or below the object's horizontal and vertical center, respectively. To protect from false start or end times, the beginning of all trials started when wrist velocity reached 5 cm/sec and ended when wrist velocity decreased to 10 cm/sec. Each object was randomly presented five times, for a total of 75 experimental trials. Sessions took ∼ 1 h to complete.

Data Analysis

The main goal of Experiment 1 was to investigate where participants looked when grasping asymmetrical shapes, when compared to symmetrical shapes, and whether fixations were linked to grasp locations. Thus, we were mainly concerned with fixation and grasp locations on the objects, and how these locations differed between conditions (symmetrical Efron blocks vs. asymmetrical Blake shapes (COML, COMR)). For the purpose of clarification in the results, object shapes are referred to individually by object type and size [COML(1-5), COMR(1-5), symmetrical shapes(1-5)] or collectively based on object category (collapsed across all sizes; COML, COMR, and symmetrical shapes).

Fixation locations (both X and Y positions) and durations were determined by a dispersion algorithm (see Salvucci and Goldberg, 2000), with a minimum duration threshold of 150 ms and a maximum dispersion threshold of 1 cm. The dispersion algorithm identifies fixations from the raw eye position data points when consecutive data points are located within a specified spatial window (maximum dispersion threshold) for a minimum period of time (minimum duration threshold). Fixations were calculated from the point when participants first opened their eyes until they made contact with the object. All X and Y coordinates of the gaze were relative to the center of the object.

For all analyses, significance levels of $p < 0.05$ were used. Analyses were carried out on the mean values computed across repeated trials in a given condition. For any main effects or interactions, Bonferroni adjusted planned comparisons were carried out. To explore differences between fixation locations (first and second fixations) across object types (COML, COMR, and symmetrical objects), and the five object sizes (X and Y dimensions: (1) 15.2 cm × 4.2 cm, (2) 12.2 cm × 5.2 cm, (3) 10.2 cm × 6.2 cm, (4) 9.0 cm × 7.1 cm, and (5) 8.0 cm × 8.0 cm; see **Figure 1**), a 2 fixation × 3 object type × 5 object size repeated measure analysis of variance (rmANOVA) was carried out for fixation positions along the X- and Y- axes separately. To explore whether fixation locations were linked to grasp axis locations for each object type, a 2 location (first fixation location by grasp axis location) × 5 object size rmANOVA was carried out for each object type (COML, COMR, and symmetrical objects) relative to X-axis locations.

For the kinematic data, rmANOVA's, with within subject variables of object type and object size, were carried out on the following dependent variables: grasp line, maximum grip aperture between the index finger and thumb (MGA), time to MGA, peak wrist velocity, and reach duration A participant's 'grasp line' location, an imaginary line connecting
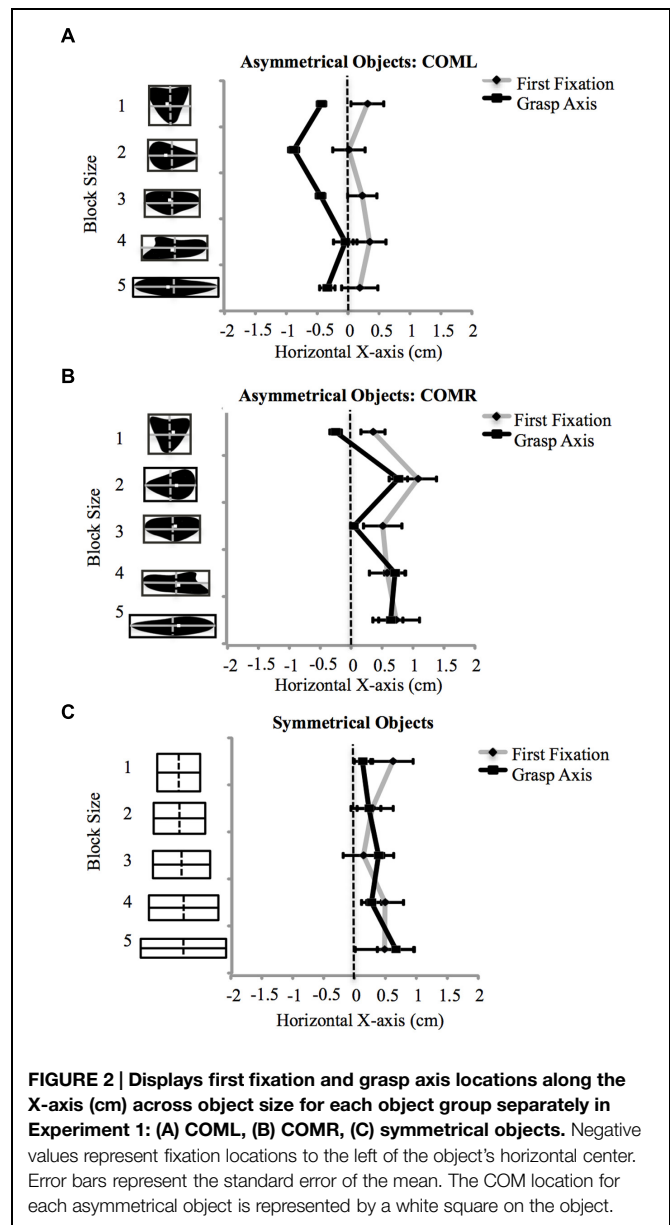
the contact points of the index finger and thumb on the object, was calculated by determining the perpendicular distance between the 'grasp line' and the horizontal midpoint of the object.

Results

Fixation Data

A first fixation was detected in 97% of all experimental trials. Ninety percent of those trials contained more than one fixation. Fixations were not detected in 3% of trials due to loss of eye data (e.g., loss of corneal reflection or IRED interference resulted in loss of data or inaccurate fixation points outside of the calibrated region), and these trials were excluded from any further analyses. A 2 fixation × 3 object type × 5 object size rmANOVA revealed the expected influence of COM location on fixation locations (i.e., fixations were drawn toward the COM location) along the X-axis [$F(2,26) = 12.46$, $p < 0.001$]. That is, fixation locations were significantly more to the left ($M = 0.13$ cm to the right of the center, $SE = 0.23$) when grasping the COML objects when compared to the COMR objects ($M = 0.68$ cm to the right of the center, SE = 0.25; $p < 0.001$). No differences in fixation locations to symmetrical objects were observed ($M = 0.33$ cm to the right of the object's center, $SE = 0.33$) when compared to asymmetrical fixation locations. No significant differences between first and second fixation locations along the X-axis were apparent in any object category (COML, COMR, symmetrical objects). A main effect of object type showed some variability in fixation locations within the different object categories [$F(8,104) = 6.20$, $p < 0.001$]. For the COML objects, fixation locations were significantly more to the left when grasping COML 4 when compared to COML 3 (see **Figure 2A**). For the COMR objects, fixation locations were significantly more to the right when grasping COMR 4 when compared to COMR 2, 3, and 5 objects (see **Figure 2B**). No significant differences in fixation locations were observed across objects within the symmetrical object category sizes (see **Figure 2C**).

Along the Y axis, a 2 fixation × 3 object type × 5 object size rmANOVA revealed a significant main effect of object size [$F(4,52) = 8.82$, $p < 0.001$]. In general, as object size increased in height, fixation locations moved progressively higher. *Post hoc* comparisons showed these differences to be significant between the tallest object categories (5: $M = 2.51$ cm, $SE = 0.40$) and 4: $M = 2.34$ cm, $SE = 0.36$) with the two shortest object categories (2: $M = 1.86$ cm, $SE = 0.31$ and 1: $M = 1.89$ cm, $SE = 0.28$; p's < 0.05). A significant main effect of object type (COML, COMR, symmetrical objects) [$F(2,26) = 9.55$, $p = 0.001$] showed differences in the Y-axis locations to asymmetrical objects relative to symmetrical shapes. Fixation locations to both asymmetrical COML and COMR objects were significantly lower on the objects (M's = 2.05 cm ($SE = 0.34$) and 2.00 cm ($SE = 0.35$) above the center, respectively) compared to fixations to symmetrical objects ($M = 2.37$ cm above center, $SE = 0.34$). A significant object type by object size interaction [$F(8,104) = 3.23$, $p = 0.02$] showed that Y fixation locations were located significantly higher when grasping symmetrical shapes 3 and 4, than when grasping COMR 3, COML 4, or COMR 4 objects (p's < 0.05; see **Figure 3**).

FIGURE 2 | Displays first fixation and grasp axis locations along the X-axis (cm) across object size for each object group separately in Experiment 1: (A) COML, (B) COMR, (C) symmetrical objects. Negative values represent fixation locations to the left of the object's horizontal center. Error bars represent the standard error of the mean. The COM location for each asymmetrical object is represented by a white square on the object.

Kinematic Data

Of all experimental trials, 4% were removed due to loss of IRED signal from the camera (due to obstruction). Results again showed the expected influence of COM location on grasp location (i.e., grasp locations were drawn to COM location). A 3 (object type) by 5 (object size) rmANOVA showed a significant main effect of object type (COML, COMR, symmetrical objects) [$F(2,26) = 28.74$, $p < 0.001$]. Grasp axis locations for COML objects were significantly more to the left ($M = 0.43$ cm to the left of the object's center, $SE = 0.06$) when compared to COMR and symmetrical objects (p's < 0.05; M's = 0.41 cm ($SE = 0.09$) and 0.29 cm ($SE = 0.19$) to the right of the object's center, respectively). No significant differences were observed between the symmetrical and COMR objects.
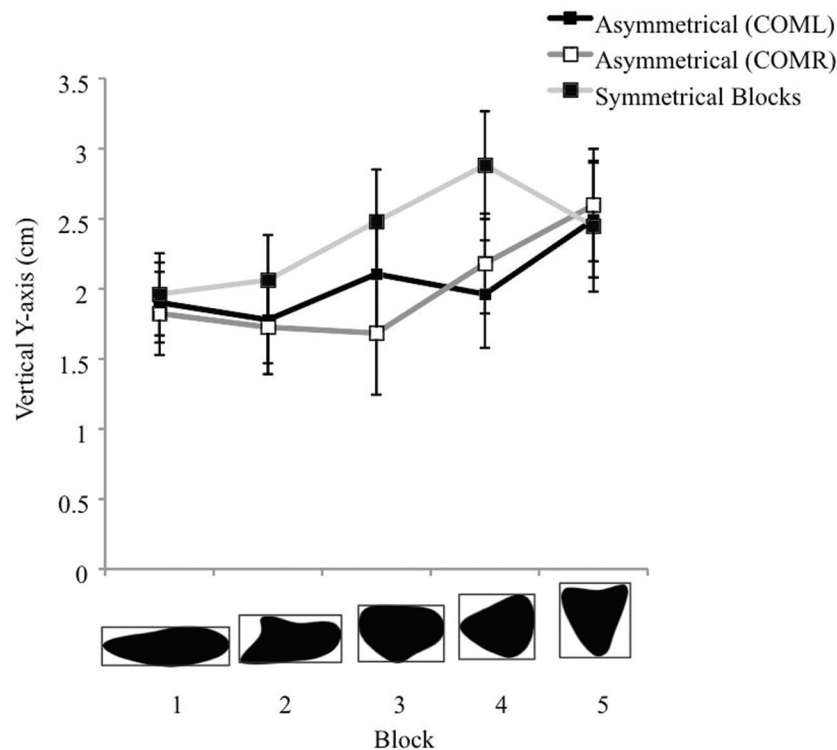
FIGURE 3 | Displays the vertical fixation locations (cm) when grasping the Asymmetrical (COML, COMR) and symmetrical objects in Experiment 1. Positive values represent fixation locations above the object's vertical center. Error bars represent the standard error of the mean.

An object type by object size interaction [$F(8,104) = 11.03$, $p < 0.001$] showed differences in grasp axis locations within each object category. For COML objects, grasp axis locations were significantly more to the left when grasping COML four when compared to all other objects (p's < 0.05; see **Figure 2A**). For COMR objects, grasp axis locations were significantly more to the left when grasping COMR 5 (the shortest in width) when compared to COMRs 4, 2, and 1 (p's < 0.05). In addition, grasp axis locations for COMR 3 were positioned significantly more to the left of grasp axis locations to COMRs 2 and 4 (p's < 0.05; see **Figure 2B**). For the symmetrical objects, no differences in grasp axis locations were observed (see **Figure 2C**).

A 3 (object type) by 5 (object size) rmANOVA on MGA showed a significant main effect of object size [$F(4,52) = 121.70$, $p < 0.001$]. Across object categories, MGA increased with object height. Planned *post hoc* comparisons revealed significant increases in MGA between all object sizes except object size 3 with 4 (p's < 0.05). On average, participants obtained MGA 75% through the reach-to-grasp movement. Across objects, no significant main effects or interactions were observed for peak velocity or total reach time (p's > 0.05). On average, participants' peak velocity was 113 cm/sec ($SE = 5$) and their completed reach-to-grasp movement was 594 ms ($SE = 38$).

Fixation Locations (X-axis) vs. Grasp Axis Locations

A 2 location (first fixation location by grasp axis location) × 5 object size rmANOVA revealed that overall, fixation locations

supported grasp axis locations for all object categories except COML objects [$F(1,13) = 8.90$, $p = 0.01$]. *Post hoc* comparisons revealed that across objects, grasp axis locations to COML objects were significantly more to the left of first fixation locations on the same objects (p's < 0.05; see **Figure 2A**). A location by object size interaction was observed for COMR objects [$F(4,52) = 3.38$, $p = 0.02$]. *Post hoc* analysis revealed that fixation locations supported grasp locations for all objects except COMR 5, where first fixations were significantly more to the right of grasp axis locations (p's < 0.05; see **Figure 2B**). Finally, a significant location by object size interaction was observed for the symmetrical objects [$F(4,52) = 4.62$, $p = 0.003$], however, *post hoc* comparisons did not reveal any significant differences in first fixation and grasp axis locations across these objects (p's > 0.05; see **Figure 2C**).

Experiment 2

The results from Experiment 1 demonstrate grasp and fixation locations were influenced by COM location. Grasp and fixation locations to asymmetrical objects with their COM to the left of the object's midline were found to be more to the left of grasp and fixation locations to asymmetrical objects whose COM was to the right of the object's midline, grasp and fixation locations for symmetrical objects were in between. In all instances, there were no differences in grasp axis locations and fixation locations

along the X-axis, except when grasping the COML objects and the smallest width asymmetrical (COMR 5) object. This experiment showed both a dissociation in fixation locations and grasp locations on asymmetrical objects due to COM position as well as a tendency to fixate areas closer to an object's COM when grasping asymmetrical vs. symmetrical shapes. However, despite COM location exerting a large effect on where we looked and grasped objects, we are still unsure how systematic changes in COM location (i.e., an increase in distance of the COM of an object from its horizontal center) would influence visuomotor control. In Experiment 2, we wanted to further explore the relationship between COM location and visuomotor control by investigating how fixation and grasp locations are affected by changes in COM distance. In other words, will fixation and grasp positions be influenced by systematic changes in COM distance from an object's midline? To explore this, the COM of three different objects were dissociated from each object's horizontal midline at three different distances and fixation and grasp locations were recorded. It was expected that participant's grasp axis and fixation locations would be shifted away from the horizontal center of the objects, towards the object's COM, and this distance would increase with increases in COM locations from that point.

Materials and Methods
Participants
Fifteen undergraduate psychology students (11 female) between the ages of 18 and 32 ($M = 20$ years-old) were recruited for participation in this study. All participants were right-handed as determined by a modified version of the Edinburg handedness inventory (Oldfield, 1971) and had normal or corrected-to-normal-vision. This research was approved by the PSREB at the University of Manitoba.

Stimulus and Procedure
All equipment, procedures, and instructions were identical as that described in Experiment 1. The stimuli used in this task were one asymmetrical object (modeled after the Blake shapes) and two differently shaped symmetrical objects with one axis of symmetry. These three distinct objects were selected to explore the effects that COM distance, across a variety of shapes, exerted on grasp and fixation selection. Each object was presented in three variations, with the COM at 0.5 cm, 1 cm, and 1.5 cm from the center of the object (total of nine objects; due to changes in COM, slight variations in shape within object categories occur; see **Figure 4**). Each object was presented 16 times [eight presentations with the COM oriented to the left (COML) and eight presentations with the COM oriented to the right (COMR) of the subjects midline], for a total of 144 trials. Sessions took approximately one and a half hours to complete.

Analysis
To explore the immediate influence of COM position on fixation locations, only first fixations were analyzed in this experiment. To investigate whether changes in COM location were affecting grasp axis locations and gaze fixation locations along the X- and Y- axis, rmANOVA's, with factors of COM position (COML,

COMR), COM distance (COM 0.5 cm, 1 cm, and 1.5 cm away from horizontal center), and object type (three different objects: object 1, 2, and 3), were performed on all dependent measures. To explore whether fixation locations supported grasp axis locations for each object type across changes in COM distance and COM position, a 2 location (first fixation location X-axis, grasp axis location) × 3 object type × 3 COM distance × 2 COM position rmANOVA was carried out.

Results
Fixations
On average participants made 2.03 fixations per trial. In 98% of all trials a first fixation was detected. Fixations were not detected in 2% of trials due to loss of eye data (e.g., loss of corneal reflection or IRED interference), and these trials were excluded from any further analyses.

Object type exerted a significant main effect on both X [$F_{(2,28)} = 3.86$, $p = 0.03$] and Y [$F_{(2,28)} = 4.42$, $p = 0.02$] fixation locations. While fixation locations to object 1 were positioned more to the right ($M = 0.61$ cm to the right of the object's midline, $SE = 0.25$) of fixations to object 2 and 3 (M's = 0.42 cm ($SE = 0.27$) and 0.44 cm ($SE = 0.23$) to the right of the object's midline, respectively), *post hoc* analysis only showed significant differences along the vertical axis; first fixations to object 3 ($M = 1.77$ cm above the object's center, $SE = 0.36$) were significantly higher when compared to object 2 ($M = 1.54$ cm above the object's center, $SE = 0.32$; $p < 0.05$). First fixations to object 1 were located 1.71 cm ($SE = 0.36$) above the object's center.

Expected differences in fixation locations based on COM position (left vs. right) along the X-axis were observed [$F_{(1,14)} = 21.95$, $p < 0.001$]. First fixations were significantly more to the left ($M = 0.17$ cm to the right of the object's horizontal midline, $SE = 0.24$) for COML objects when compared to COMR objects ($M = 0.81$ cm to the right of the midline, $SE = 0.27$). An object type by COM position interaction [$F_{(2,28)} = 4.49$, $p = 0.02$] showed that fixations were significantly more to the left for object 3 compared to object 1 for COML objects; for COMR objects, fixations to object 3 were significantly more to the right when compared to object 2 (p's < 0.05; see **Figure 5**).

Center of mass distance also influenced fixation locations. A COM distance by COM position interaction for the X-axis fixation locations [$F_{(2,28)} = 12.28$, $p < 0.001$] revealed that fixation locations in the COML condition moved increasingly leftward as the COM moved increasingly left (note: participants started with a rightward bias; see **Figure 6**). Significant differences were observed between COM distance .5 cm with COM distance 1 cm and 1.5 cm (p's < 0.05). For COMR objects, fixation locations moved increasingly rightward as the COM moved increasingly right. Significant differences were observed between COM distance .5 cm with COM distance 1 cm and 1.5 cm (p's < 0.05; see **Figure 6**). Along the Y axis, a main effect of COM distance [$F_{(2,28)} = 3.90$, $p = 0.03$] showed that as COM locations were positioned further from the center (resulting in decreases in object height at the center of the objects) fixation locations moved closer to the object's vertical center [M's = 1.80 cm ($SE = 0.36$),
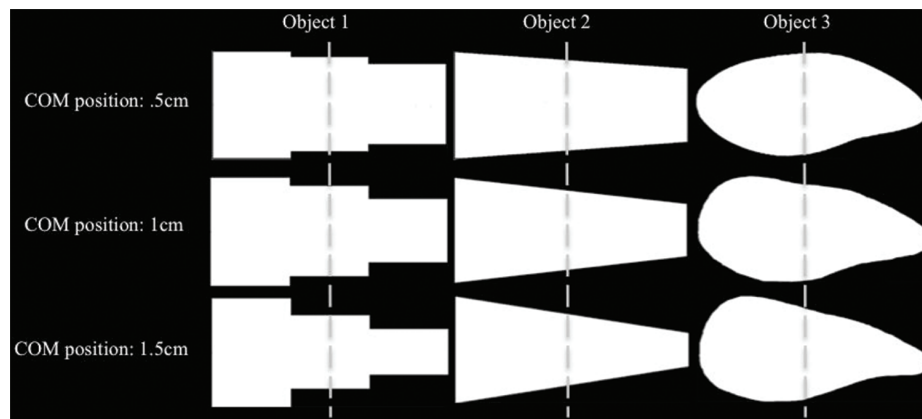
FIGURE 4 | Displays the two symmetrical and one asymmetrical object used in Experiment 2. Each object shape was manipulated such that the COM of each object was located at three distances from the object's horizontal center: 0.5 cm, 1 cm, and 1.5 cm. The dashed line transects the object's horizontal center.
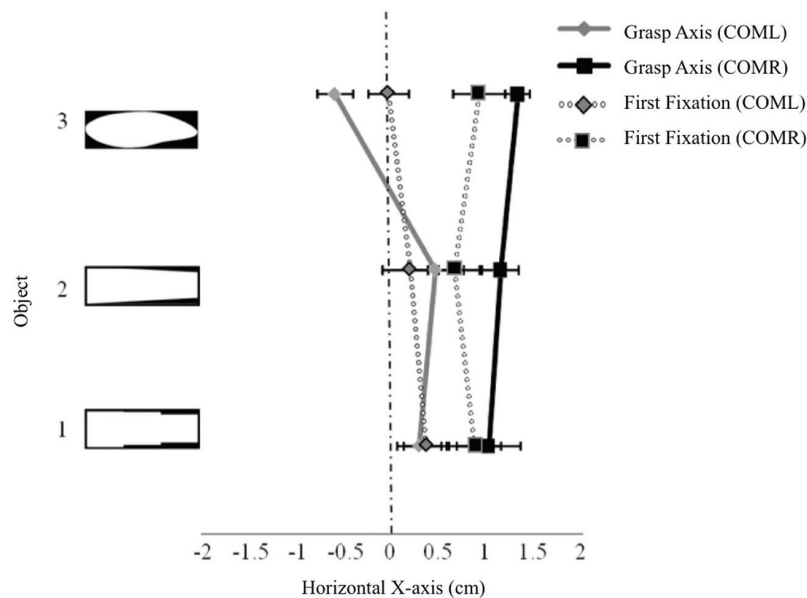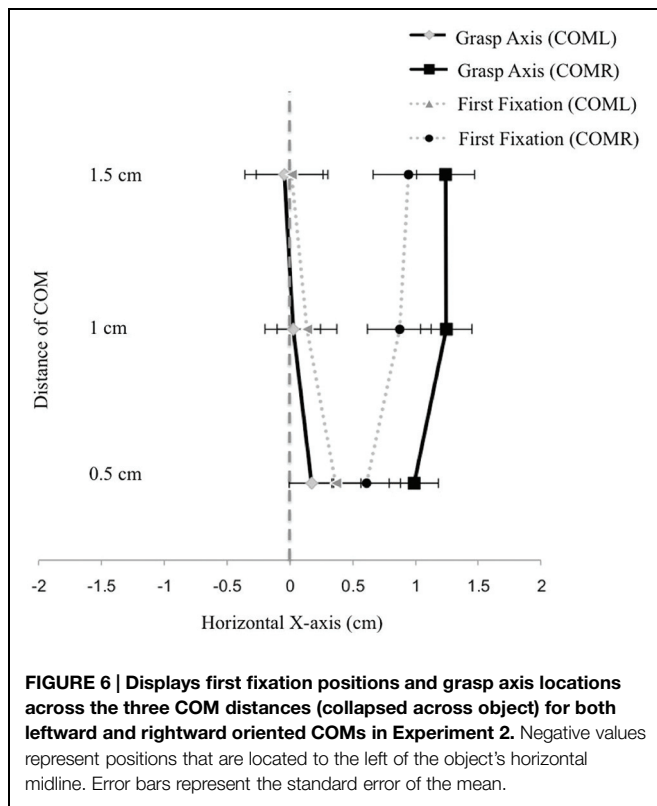

FIGURE 5 | Shows the mean first fixation and grasp axis locations along the horizontal X-axis collapsed across COM distance for the three different object types in Experiment 2. Negative values represent positions that are located to the left of the object's horizontal midline. Error bars represent the standard error of the mean.

1.65 cm ($SE = 0.34$), and 1.58 cm ($SE = 0.33$), respectively]. Significant differences were observed between the tallest object size (COM distance of 0.5 cm) and the shortest object size (COM distance of 1.5 cm; $p < 0.05$).

Grasp Location

Object shape had a significant effect on grasp location [$F(2,28) = 7.33, p = 0.003$], participants' grasp axis locations were significantly farther away from the object's center for object 2 ($M = 0.80$ cm to the right of the object's midline, $SE = 0.21$) compared to object 3 ($M = 0.36$ cm to the right of the object's midline, $SE = 0.14$; $p < 0.05$). A significant main effect of COM position [$F(1,14) = 30.88, p < 0.001$] and an object by

COM position interaction [$F(2,28) = 13.53, p < 0.001$] were also observed. Grasp axis locations were significantly more to the left for COML objects ($M = 0.05$ cm to the right of the midline, $SE = 0.23$) when compared to COMR objects ($M = 1.16$ cm to the right of the midline, $SE = 0.20$). The object by COM position interaction showed no differences between objects in grasp axis locations when the objects were oriented to the right of the center (see **Figure 5**). With COML objects, grasp axis locations for object 3 were significantly more to the left than grasp axis locations for objects 1 and 2 (p's < 0.05; see **Figure 5**). No significant main effects of COM distance or any object or COM position by COM distance interactions were observed (p's > 0.05).

FIGURE 6 | Displays first fixation positions and grasp axis locations across the three COM distances (collapsed across object) for both leftward and rightward oriented COMs in Experiment 2. Negative values represent positions that are located to the left of the object's horizontal midline. Error bars represent the standard error of the mean.

First Fixation Location (X-axis) vs. Grasp Axis Location

Results revealed a significant COM location by object type by COM position three-way interaction [$F(2,28) = 6.50$, $p = 0.01$]. *Post hoc* analysis showed that first fixations and grasp axis locations (collapsed across COM distance) for all objects in both COM positions (COML, COMR) were the same, except for object 3. When object 3's COM was oriented to the left of the center, fixation locations were found to be positioned significantly more to the right of grasp axis locations (see **Figure 5**).

Discussion

The characteristics of reaching and grasping objects have been well documented (e.g., Jeannerod, 1986; Gentilucci et al., 1991; Jakobson and Goodale, 1991; Paulignan et al., 1991; Galletti et al., 2003; Castiello, 2005). Traditionally, however, the primary concern in the reaching and grasping literature has been with how the opening of the hand is coordinated with the hand's approach towards the target objects, typically using regular shaped objects or objects where the grasp points were controlled. Few studies have examined the selection of grasp locations when grasping irregularly shaped objects and to our knowledge, no studies have examined the selection of grasp and gaze behaviors during various manipulations of object COM. The purpose of this research was to examine the variability in fixation locations across several conditions on irregular asymmetrical objects and explore the relationship of these fixations to grasp locations on those same objects. In Experiment 1, we investigated fixation and grasp locations to contoured objects that had an

asymmetrical design and compared this behavior to grasps made to symmetrical objects that had identical maximum horizontal and vertical dimensions. In Experiment 2, we explored the effects of object shape and COM location on fixation and grasp locations. The combined results from these studies demonstrate several significant effects of object properties on grasp and fixation locations, including: COM position (COML vs. COMR) influences where we grasp and where we look when picking up an object (Experiments 1 and 2); fixations are less linked to grasp locations when we are grasping asymmetrical objects with the COM oriented to the left of the object's midline (Experiments 1 and 2); object irregularity results in more central fixations (Experiment 1), and; increasing COM distance from the objects horizontal midline affects grasp and fixation locations differently (Experiment 2). These findings will be discussed in turn.

Results from Experiments 1 and 2 demonstrated that both fixation and grasp locations were influenced by COM location (COML vs. COMR). That is, grasp and fixation locations were drawn towards COM location, resulting in positional differences in visuomotor control between leftward and rightward oriented COM positions. This manipulation, however, also systematically changed other factors of the objects, which arguably could have influenced grasp and fixation behavior as well. For example, the side of the object where the COM was located also appears larger than the other side. This difference in object size could have biased the perception of the ease of grasp and influenced results. Despite this caveat, COM position, rather than differences in object width, does seem to be the determining factor mediating grasp and fixation locations, consistent with previous reaching and grasping findings (e.g., Jeannerod, 1988; Goodale et al., 1994; Lederman and Wing, 2003; Marotta et al., 2003; Kleinholdermann et al., 2007). For example, in Experiment 1, for asymmetrical objects 1 and 3, the center of the object is approximately the same width as at the COM location. However, we still see a bias in grasp location depending on the position of the COM (Left vs. Right). If object width was influencing grasp positions due to ease of grasp, then we would not expect such a large influence of COM position with these objects as the center would be just as likely to be grasped in both COM orientations. Additionally, in Experiment 2 (for objects 1 and 2), if participants were biased to grasp the larger parts of the objects we would expect to see grasp locations much more influenced by these areas. Rather, we saw grasp locations much closer to the object's COM despite very large widths present on one side (especially object 1). Together these finding advance on the existing literature by including the link between fixation and grasping behavior to irregular shaped asymmetrical objects and highlight the importance of COM position for both behaviors. This influence was apparent in both studies, despite small deviations in COM position (within 1.5 cm) from the blocks midline.

The manipulation of COM location also revealed several interesting differences in eye-hand behavior. For instance, fixation and grasp locations for COMR objects, while more to the right, were not found to be significantly different in position than those to symmetrical shapes. Potentially, this lack of difference is due to the slight rightward grasp and fixation biases when picking

up symmetrical objects – results demonstrated in previous studies (Desanghere and Marotta, 2011; Prime and Marotta, 2013). Additionally, in both instances overall fixation locations were linked to grasp location as demonstrated in previous studies (de Grave et al., 2008; Brouwer et al., 2009; Desanghere and Marotta, 2011; Prime and Marotta, 2013; Bulloch et al., 2015). When the COM of the asymmetrical objects were positioned to the left of the horizontal midline (COML), despite grasp positions that were to the left of the center of the objects, fixations were again found to be in close proximity to fixations when grasping symmetrical shapes (i.e., to the right of the object's midline); but still significantly more to the left than fixations to COMR objects, thus demonstrating a COM position influence on where we look. In this condition, these rightward fixation locations did not support grasp locations as with symmetrical and COMR objects. These results suggest a rightward fixation bias when interacting with both symmetrical and asymmetrical shapes, regardless of the position of the object's COM. Rightward fixation biases have been shown in other studies when participants are instructed to look at the mid point of an object. For example, previous research has shown that when participants "visually bisect" complex stimuli, the subjective midpoint is placed to the right of the object's true center (Elias et al., 2005; Rhode and Elias, 2007). In addition, Handy et al. (2003) showed that visual spatial attention was drawn to graspable objects (tools) in the right visual hemifield. These results suggest visual field asymmetries in the processing of action-related attributes and spatial attention, with attention to specific object features aiding in recognition of the motor affordance (Handy et al., 2003).

Interestingly, recent research has demonstrated an effect of handedness on grasp point selection when picking up symmetrical objects (Paulun et al., 2014). Consistent with previous research (e.g., Desanghere and Marotta, 2011; Prime and Marotta, 2013), Paulun et al. (2014) demonstrated a slight rightward grasp bias relative to the object's COM, when participants were picking up symmetrical objects. Conversely, a slight leftward grasp bias was demonstrated when participants grasped the objects with their left hand. These authors suggest that the variation in grasp point selection is the result of a compromise between obtaining maximum stability (grasp points near the COM) and a slight lateral deviation toward the side of the grasping hand, potentially to increase the visibility of the object as a whole while lifting it (Paulun et al., 2014). Whether fixations would support this leftward grasp bias when grasping with the left hand, or whether fixations would be drawn to the right side of the object as demonstrated in the present experiments is yet to be determined. Indeed, the results from the present study do suggest a persistent rightward visuospatial bias, despite grasp locations to the left of an objects midline when the COM of the object is oriented in that direction. For example, and as previously mentioned, fixations to COML objects in Experiment 1 did not coincide with where participants were grasping the objects. Grasp locations were drawn to COM position (positioned to the left of the horizontal midline); however, a dissociation between fixation locations and grasp positions were observed in this condition, with fixation locations to the right of the objects

center. This effect was also demonstrated in Experiment 2, where a rightward fixation bias (relative to the center of the object) and a difference in this position relative to grasp locations for COML objects (object 3) were also shown. Overall, these results are similar to Prime and Marotta (2013) who demonstrated a decoupling in fixation locations with grasp locations when performing a memory guided grasping task to symmetrical shapes. They conclude that the purpose of initial fixations for the purpose of memory guided grasping is to provide the visuomotor system with a general perceptual analysis of the blocks properties. The present results suggest that visual attention and grasping movements become loosely coupled in some conditions, with grasp locations largely mediated by COM location and fixations only supporting grasp locations when they are to the right of the object's horizontal midline. When grasp locations are to the left of the object's midline, fixations are drawn to the right of center; similar to findings during perceptual tasks.

In addition to this slight rightward bias, our results also demonstrated that more irregular structures are eliciting fixation locations that are more central on the objects (lower on the objects when compared to fixations to symmetrical shapes; Experiment 1), regardless of COM position. Overall, object irregularity is resulting in more centralized fixation locations on the objects, perhaps to maintain maximum visibility of the object's shape, where placement of both index finger and thumb is important for grasp stability on the irregular shaped objects. Indeed, allocating attention to a specific location on an object, in this case a more central location, has been shown to result in faster and more accurate processing of form information in regions of space surrounding that location (Bashinski and Bacharach, 1980; Hoffman and Nelson, 1981; Downing, 1988). This more central view would provide a more holistic representation during grasping, taking the whole object into consideration. Definitely, when we reach out to pick up an object, we are able to do this with great precision, regardless of the object contours. To do this, information about intrinsic (e.g., size, shape) and extrinsic (e.g., distance, orientation) features need to be transformed in order to develop a motor plan for movement execution (Jeannerod, 1988). Our results suggest that the visual analysis of complex objects is different from that of symmetrical shapes. While research supports that when grasping simple objects, fixation locations are tightly linked to index finger grasp location, the more complex structures in these experiments are eliciting fixation locations that are more central on the objects (lower on the objects when compared to fixations to symmetrical shapes) and close to the horizontal midpoint of the block regardless of COM position.

In Experiment 2 we also showed that increasing COM distance from the objects horizontal midline affects grasp and fixation locations differently. In this experiment, the influence of COM position on grasp and fixation locations were explored through manipulations of COM distance from the horizontal midline of the objects (up to 1.5 cm from the object's horizontal center). These manipulations did not have the expected systematic influence on grasp point selection. Consistent with Experiment 1, changes in COM position affected both grasp and gaze positions, however, increased distances from that point did not

exert a further influence on grasp locations. Regardless of the COM position, and collapsed across all object types, significant increases in distance for grasp locations from the center of the objects were not observed with increased distances of COM. Unlike grasp locations, fixations were influenced by systematic changes in COM location, with fixations to objects with COM distances of 1 cm and 1.5 cm away from the center, significantly further away from fixation locations to the objects with the COM closest to the object's center (0.5 cm) in both COM positions. Interestingly, a rightward fixation bias was again present for all objects. This research suggests that fixations are influenced by COM position and distance, but tend to still maintain a "central" position, with a slight rightward bias relative to the center of the object. The maintenance of these centralized fixation locations were also observed along the vertical axis as well, with fixation locations moving progressively closer to the objects center with systematic increases in COM position (resulting in a decrease in object height at the block's midpoint). Again, supporting the notion that when grasping irregular shapes, a central view provides a more holistic representation that may be needed for monitoring both index and thumb placement on irregular objects.

Taken together, our findings not only highlight the importance of including irregular, non-symmetrical objects in visuomotor paradigms but also reveal how object features differentially influence gaze vs. grasping during object interaction. The importance in allocating visual attention to an object's COM or the exact grasp location of the index finger on asymmetrical objects may become more important when interacting with heavier objects. When we reach out to pick up an object, the anticipated mass of that object automatically influences anticipatory grip forces (i.e., grip force is scaled proportional to the expected weight of an object, which is based on its size and type of material; Gordon et al., 1991). In other words, when we reach out to grasp an object using our index finger and thumb, the opposing digits exert a grip force (forces that are equal and opposite) to hold the object level. If the COM of the object is to one side of our grasp location, this offset position results in a turning force or torque. If we are to successfully grasp this object, we can either increase our grip force in order to generate the torsional friction needed to offset the rotation of the object around the grasp axis and keep the object level during the lift, or move our grasp axis to intersect the object's COM (Wing and Lederman, 1998; Endo et al., 2011). Since the objects in this study were relatively lightweight, easily compensated for by increases in grip force, the placement of our grasp axis with varying degrees of COM changes becomes less important for grasp location selection. Despite this, however, it is apparent that we attend to changes in COM location, despite these locations not necessarily dictating our exact grasp location. While visual attention to asymmetrical objects remains relatively central (with a slight rightward bias), we are clearly sensitive to slight changes in COM location. Unlike grasping, in which changes in COM can be easily compensated for in grip force, the eyes have to attend to changes in relevant properties of the objects to help mediate this process. As you can imagine, the placement of our grasp would become progressively linked to an object's COM if these changes also coincided with an increase in object mass and, if this were to occur, it would be important to be attending to these locations. Further research is needed to explore these factors as well as the persistent rightward visuospatial bias observed when grasping objects and the influence of handedness on gaze and grasp point selection during object manipulation.

References

Abrams, R. A., Meyer, D. E., and Kornblum, S. (1990). Eye-hand coordination: oculomotor control in rapid aimed limb movements. *J. Exp. Psychol. Hum. Percept. Perform.* 16, 248–267. doi: 10.1037/0096-1523.16.2.248

Araujo, C., Kowler, E., and Pavel, M. (2001). Eye movements during visual search: the costs of choosing the optimal path. *Vis. Res.* 41, 3613–3625. doi: 10.1016/S0042-6989(01)00196-1

Ballard, D. H., Hayhoe, M. M., Li, F., and Whitehead, S. D. (1992). Hand-eye coordination during sequential tasks. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 337, 331–338. doi: 10.1098/rstb.1992.0111

Bashinski, H. S., and Bacharach, V. R. (1980). Enhancement of perceptual sensitivity as the result of selectively attending to spatial 1ocations. *Percept. Psychophys.* 28, 241–280. doi: 10.3758/BF03204380

Bekkering, H., Adam, J. J., Kingma, H., Huson, A., and Whiting, H. T. (1994). Reaction time latencies of eye and hand movements in single- and dual-task conditions. *Exp. Brain Res.* 97, 471–476. doi: 10.1007/BF00241541

Biguer, B., Jeannerod, M., and Prablanc, C. (1982). The coordination of eye, head, and arm movements during reaching at a single visual target. *Exp. Brain Res.* 46, 301–304. doi: 10.1007/BF00237188

Binsted, G., Chua, R., Helsen, W., and Elliott, D. (2001). Eye–hand coordination in goal-directed aiming. *Hum. Mov. Sci.* 20, 563–585. doi: 10.1016/S0167-9457(01)00068-9

Blake, A. (1992). Computational modelling of hand–eye coordination. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 337, 351–360. doi: 10.1098/rstb.1992.0113

Brouwer, A. M., Franz, V. H., and Gegenfurtner, K. R. (2009). Differences in fixations between grasping and viewing objects. *J. Vis.* 9, 11–24. doi: 10.1167/9.1.18

Bulloch, M. C., Prime, S. L., and Marotta, J. J. (2015). Anticipatory gaze strategies when grasping moving objects. *Exp. Brain Res.* doi: 10.1007/s00221-015-4413-7 [Epub ahead of print],

Castiello, U. (2005). The neuroscience of grasping. *Nat. Rev. Neurosci.* 6, 726–736. doi: 10.1038/nrn1775

Cavina-Pratesi, C., and Hesse, C. (2013). Why do the eyes prefer the index finger? Simultaneous recording of eye and hand movements during precision grasping. *J. Vis.* 13, 1–15. doi: 10.1167/13.5.15

de Grave, D. D. J., Hesse, C., Brouwer, A. M., and Franz, V. H. (2008). Fixation locations when grasping partly occluded objects. *J. Vis.* 8, 1–11. doi: 10.1167/8.7.5

Desanghere, L., and Marotta, J. J. (2011). "Graspability" of objects affect gaze patterns during perception and action tasks. *Exp. Brain Res.* 212, 177–187. doi: 10.1007/s00221-011-2716-x

Downing, C. J. (1988). Expectancy and visual-spatial attention: effects on perceptual quality. *J. Exp. Psychol.* 14, 188–202.

Efron, R. (1968). *What is Perception?* New York, NY: Humanitites Press.

Elias, L. J., Robinson, B., and Saucier, D. M. (2005). Rightward biases in free-viewing visual bisection tasks: implications for leftward responses biases on similar tasks. *Brain Cogn.* 59, 296–298. doi: 10.1016/j.bandc.2004.02.021

Eloka, O., and Franz, V. H. (2011). Effects of object shape on the visual guidance of action. *Vis. Res.* 51, 925–931. doi: 10.1016/j.visres.2011.02.002

Endo, S., Wing, A. M., and Bracewell, R. M. (2011). Haptic and visual influences on grasp point selection. *J. Mot. Behav.* 43, 427–431. doi: 10.1080/00222895.2011.621996

Findlay, J. M. (1997). Saccade target selection during visual search. *Vis. Res.* 37, 617–631. doi: 10.1016/S0042-6989(96)00218-0

Galletti, C., Kutz, D. F., Gamberini, M., Breveglieri, R., and Fattori, P. (2003). Role of the medial parieto-occipital cortex in the control of reaching and grasping movements. *Exp. Brain Res.* 153, 158–170. doi: 10.1007/s00221-003-1589-z

Gentilucci, M., Castiello, U., Corradini, M. L., Scarpa, M., Umilta, C., and Rizzolatti, G. (1991). Influence of different types of grasping on the transport component of prehension movements. *Neuropsychologia* 29, 361–378. doi: 10.1016/0028-3932(91)90025-4

Goodale, M. A. (1998). "Vision for perception and vision for action in the primate brain," in *Sensory Guidance of Movement*, eds. G. R. Bock and J. A. Goode (London: John Wiley & Sons Ltd), 21–34.

Goodale, M. A. (2014). How (and why) the visual control of action differs from visual perception. *Proc. R. Soc.* 281:20140337. doi: 10.1098/rspb.2014.0337

Goodale, M. A., Meenan, J. P., Bülthoff, H. H., Nicolle, D. A., Murphy, K. J., and Racicot, C. I. (1994). Separate neural pathways for the visual analysis of object shape in perception and prehension. *Curr. Biol.* 4, 604–610. doi: 10.1016/S0960-9822(00)00132-9

Gordon, A. M., Forssberg, H., Johansson, R. S., and Westling, G. (1991). Integration of sensory information during the programming of precision grip: comments on the contributions of size cues. *Exp. Brain Res.* 85, 226–229. doi: 10.1007/BF00230004

Hamed, S., Duhamel, J. R., Bremmer, F., and Graf, W. (2002). Visual receptive field modulation in the lateral intraparietal area during attentive fixation and free gaze. *Cereb. Cortex* 12, 234–245. doi: 10.1093/cercor/12.3.234

Handy, T. C., Grafton, S. T., Shroff, N. M., Ketay, S., and Gazzaniga, M. (2003). Graspable objects grab attention when the potential for action is recognized. *Nat. Neurosci.* 6, 421–427. doi: 10.1038/nn1031

Hayhoe, M., and Ballard, D. (2005). Eye movements in natural behavior. *Trends Cogn. Sci.* 9, 188–194. doi: 10.1016/j.tics.2005.02.009

Hayhoe, M. M., Shrivastava, A., Mruczek, R., and Pelz, J. B. (2003). Visual memory and motor planning in a natural task. *J. Vis.* 3, 49–63. doi: 10.1167/3.1.6

He, P. Y., and Kowler, E. (1991). Saccadic localization of eccentric forms. *J. Opt. Soc. Am.* 8, 440–449. doi: 10.1364/JOSAA.8.000440

Hoffman, J. E., and Nelson, B. (1981). Spatial selectivity in visual search. *Percept. Psychophys.* 30, 283–290. doi: 10.3758/BF03214284

Irwin, D. E. (2004). *The Interface of Language, Vision and Action: Eye Movements and the Visual World*. New York, NY: Psychology Press, 105–133.

Jakobson, L. S., and Goodale, M. A. (1991). Factors affecting higher-order movement planning: a kinematic analysis of human prehension. *Exp. Brain Res.* 86, 199–208. doi: 10.1007/BF00231054

Jeannerod, M. (1986). Mechanisms of visuomotor coordination: a study in normal and brain-damaged subjects. *Neuropsychologia* 24, 41–78. doi: 10.1016/0028-3932(86)90042-4

Jeannerod, M. (1988). *The Neural and Behavioural Organization of Goal-Directed Movements*. Oxford: Clarendon.

Jeannerod, M., Arbib, M. A., Rizzolatti, G., and Sakata, H. (1995). Grasping objects: the cortical mechanisms of visuomotor transformation. *Trends Neurosci.* 18, 314–320. doi: 10.1016/0166-2236(95)93921-J

Jenmalm, P., Goodwin, A. W., and Johansson, R. S. (1998). Control of grasp stability when humans lift objects with different surface curvatures. *J. Neurophysiol.* 79, 1643–1652.

Johansson, R. S., Westling, G., Backstrom, A., and Flanagan, J. R. (2001). Eye-hand coordination in object manipulation. *J. Neurosci.* 21, 6917–6932.

Kandel, E., Schwartz, J., and Jessell, T. (2000). *Principles of Neural Science*, 4 Edn. New York, NY: McGraw-Hill Professional.

Klein, R. M. (2000). Inhibition of return. *Trends Cogn. Sci.* 4, 138–147. doi: 10.1016/S1364-6613(00)01452-2

Kleinholdermann, U., Brenner, E., Franz, V. H., and Smeets, J. B. (2007). Grasping trapezoidal objects. *Exp. Brain Res.* 180, 415–420. doi: 10.1007/s00221-007-0867-6

Kowler, E., and Blaser, E. (1995). The accuracy and precision of saccades to small and large targets. *Vis. Res.* 35, 1741–1754. doi: 10.1016/0042-6989(94)00255-K

Land, M., and Hayhoe, M. (2001). In what ways do eye movements contribute to everyday activities? *Vis. Res.* 41, 3559–3565. doi: 10.1016/S0042-6989(01)00102-X

Land, M., Mennie, N., and Rusted, J. (1999). The roles of vision and eye movements in the control of activities of daily living. *Perception* 28, 1311–1328. doi: 10.1068/p2935

Lederman, S. J., and Wing, A. M. (2003). Perceptual judgment, grasp point selection and object symmetry. *Exp. Brain Res.* 152, 156–165. doi: 10.1007/s00221-003-1522-5

Marotta, J. J., McKeeff, T. J., and Behrmann, M. (2003). Hemispatial neglect: its effects on visual perception and visually guided grasping. *Neuropsychologia* 41, 1262–1271. doi: 10.1016/S0028-3932(03)00038-1

McGowan, J. W., Kowler, E., Sharma, A., and Chubb, C. (1998). Saccadic localization of random dot targets. *Vis. Res.* 38, 895–909. doi: 10.1016/S0042-6989(97)00232-0

Melcher, D., and Kowler, E. (1999). Shapes, surfaces and saccades. *Vis. Res.* 39, 2929–2946. doi: 10.1016/S0042-6989(99)00029-2

Milner, A. D., and Goodale, M. A. (1995). *The Visual Brain in Action*, Vol. 27. Oxford: Oxford University Press.

Oldfield, R. C. (1971). The assessment and analysis of handedness: the edinburgh inventory. *Neuropsychologia* 9, 97–113. doi: 10.1016/0028-3932(71)90067-4

Paulignan, Y., Jeannerod, M., MacKenzie, C., and Marteniuk, R. (1991). Selective perturbation of visual input during prehension movements. The effects of changing object size. *Exp. Brain Res.* 87, 407–420.

Paulun, V. C., Kleinholdermann, U., Gegenfurtner, K. R., Smeets, B. J., and Brenner, E. (2014). Center or side: biases in selecting grasp points on small bars. *Exp. Brain Res.* 232, 2061–2072. doi: 10.1007/s00221-014-3895-z

Posner, M. I. (1980). Orienting of attention. *Q. J. Exp. Psychol.* 32, 3–25. doi: 10.1080/00335558008248231

Prablanc, C., Echallier, J. F., Komilis, E., and Jeannerod, M. (1979). Optimal response of eye and hand motor systems in pointing at a visual target. Spatio-temporal characteristics of eye and hand movements and their relationships when varying the amount of visual information. *Biol. Cybern.* 35, 113–124.

Prime, S. L., and Marotta, J. J. (2013). Gaze strategies during visually-guided versus memory-guided grasping. *Exp. Brain Res.* 225, 291–305. doi: 10.1007/s00221-012-3358-3

Rhode, L., and Elias, L. J. (2007). Visual bisection of freely-viewed asymmetrical stimuli. *Laterality* 12, 347–355.

Riek, S., Tresilian, J. R., Mon-Williams, M., Coppard, V. L., and Carson, R. G. (2003). Bimanual aiming and overt attention: one law for two hands. *Exp. Brain Res.* 153, 59–75. doi: 10.1007/s00221-003-1581-7

Salvucci, D. D., and Goldberg, J. H. (2000). "Identifying fixations and saccades in eye-tracking protocols," in *Proceedings of the Eye Tracking Research and Applications Symposium*, (New York, NY: ACM Press), 71–78.

Schenk, T., and McIntosh, R. D. (2010). Do we have independent visual streams for perception and action? *Cogn. Neurosci.* 1, 52–62. doi: 10.1080/17588920903388950

Smeets, J. B., and Brenner, E. (1999). A new view on grasping. *Motor Control* 3, 237–271.

Smeets, J. B., Hayhoe, M. M., and Ballard, D. H. (1996). Goal-directed arm movements change eye-head coordination. *Exp. Brain Res.* 109, 434–440. doi: 10.1007/BF00229627

Soechting, J. F., Engel, K. C., and Flanders, M. (2001). The Duncker illusion and eye-hand coordination. *J. Neurophysiol.* 85, 843–854.

van Donkelaar, P., Siu, K. C., and Walterschied, J. (2004). Saccadic output is influenced by limb kinetics during eye-hand coordination. *J. Mot. Behav.* 36, 245–252. doi: 10.3200/JMBR.36.3.245-252

Vishwanath, D., and Kowler, E. (2003). Localization of shapes: eye movements and perception compared. *Vis. Res.* 43, 1637–1653. doi: 10.1016/S0042-6989(03)00168-8

Vishwanath, D., and Kowler, E. (2004). Saccadic localization in the presence of cues to three-dimensional shape. *J. Vis.* 4, 445–458. doi: 10.1167/4.6.4

Vishwanath, D., Kowler, E., and Feldman, J. (2000). Saccadic localization of occluded targets. *Vis. Res.* 40, 2797–2811. doi: 10.1016/S0042-6989(00)00118-8

Wing, A. M., and Lederman, S. J. (1998). Anticipating load torques produced by voluntary movements. *J. Exp. Psychol. Hum. Percept. Perform.* 24, 1571–1581. doi: 10.1037/0096-1523.24.6.1571

Yarbus, A. L. (1967). *Eye Movements and Vision.* New York, NY: Plenum Press.

Zelinsky, G. J., Rao, R. P. N., Hayhoe, M. M., and Ballard, D. (1997). Eye movements reveal the spatiotemporal dynamics of visual search. *Psychol. Sci.* 8, 448–453. doi: 10.1111/j.1467-9280.1997.tb00459.x

Independent effects of 2-D and 3-D locations of stimuli in a 3-D display on response speed in a Simon task

Hiroyuki Umemura*

Medical and Biological Engineering Research Group, Biomedical Research Institute, National Institute of Advanced Industrial Science and Technology, Ikeda, Japan

The Simon Effect is a phenomenon in which reaction times are usually faster when the stimulus location and the response correspond, even if the stimulus location is irrelevant to the task. Recent studies have demonstrated the Simon effect in a three-dimensional (3-D) display. The present study examined whether two-dimensional (2-D) and 3-D locations simultaneously affected the Simon effect for stimuli in which a target and fixation were located on the same plane (ground or ceiling) at different 3-D depths, and the perspective effect produced a difference in the 2-D vertical location of the target stimulus relative to the fixation. The presence of the ground and ceiling plane was controlled to examine the contextual effects of background. The results showed that the 2-D vertical location and 3-D depth simultaneously affected the speed of responses, and they did not interact. The presence of the background did not affect the magnitude of either the 2-D or the 3-D Simon effect. These results suggest that 2-D vertical location and 3-D depth are coded simultaneously and independently, and both affect response selection in which 2-D and 3-D representations overlap.

Keywords: Simon effect, 3-D, stimulus–response compatibility, reaction time, binocular disparity

Introduction

When people perform a choice reaction time (RT) task, the time needed to make a response varies with the compatibility between the stimulus and the response. In the spatial case of this stimulus–response (S–R) compatibility effect, RTs are usually faster and responses are more accurate when the stimulus occurs in the same relative location as the response. The Simon Effect is a specific case of the S–R compatibility effect in which the stimulus location is irrelevant to the task (Simon and Rudell, 1967). For example, participants are instructed to press a right key whenever they observe a red target and a left key in response to a white target. Even though stimulus location is entirely task-irrelevant, responses are typically faster when response keys spatially correspond to the stimulus location: red on the right or white on the left. The Simon effect has been extensively investigated, not only because the effects are useful in the design of man–machine interfaces, but also because they provide important insights on attentional operations, the representation of space and body, the cognitive representation of intentional action, and decision making and action execution (Kornblum et al., 1990; Hommel, 2011).

Most studies on the Simon effect have been conducted with two-dimensional (2-D) stimulus displays. Recently, Rigon et al. (2011) showed that the Simon effect is not confined to 2-D displays but can be observed for stimulus locations in depth in three-dimensional (3-D) space defined by

binocular disparity (provided by a color anaglyph system or a VR headset). Participants responded faster when the stimulus location (near or far) and the location of response keys were identical. Stins and Michaels (2010) investigated the effect of 3-D background on the S–R compatibility effect. They used a joystick to collect responses and found that RT was affected by 3-D orientation of the background, as defined by a texture gradient (i.e., monocular depth cue). They found an S–R compatibility effect between the 3-dimensional 'far' location and the response made by moving a joystick toward 'far' (but the effect was not observed when the target was displayed at a near distance in their study). In these and other studies on the S–R compatibility effect in 3-D locations (Chan and Chan, 2010), the effects of 2-D location were regarded as a disturbance. Therefore, 2-D locations were counterbalanced and were not analyzed. As a result, it is not known whether 2-D and 3-D locations simultaneously affect response speed.

It is important to examine whether the 3-D and 2-D locations simultaneously influence the spatial correspondence effect, because this can help explain how the representations of target location and response are coded. In the formulation of the Simon effect, it is assumed that a spatial code is generated for the stimulus location which is irrelevant to the task, and that the Simon effect occurs at the response selection stage (Stoffer and Umiltà, 1997). A stimulus automatically activates the response code that spatially corresponds to it if there is sufficient similarity between the spatial stimulus dimension and the spatial response dimension (Kornblum et al., 1990). When the activated response code is different from the response code required by the task, a conflict is generated that requires time to resolve.

The present study examined how 2-D (vertical) location and 3-D location are coded and how these representations interact. One hypothesis is that both the 2-D and 3-D locations of the stimulus influence the speed of responses. Previous research has shown that multiple spatial codes influence the speed of a stimulus identification task. Lamberts et al. (1992) compared the spatial S–R compatibility among eight different positions which were obtained as a result of orthogonal manipulation of hemispace, visual hemifield within hemispace, and relative position within hemifield. They found compatibility effects based on both hemifield and relative position. Hommel and Lippa (1995) showed that object-based spatial stimulus codes are formed automatically and thus influence the speed of response selection. They had participants respond to a stimulus superimposed on the eyes of an image of a face. They found that when the response location was identical to the position of the eye (e.g., 'left' response for stimuli on the left eye), the response was relatively fast even if the stimulus was aligned vertically by tilting the face 90 or 270°. This means that the tilted image of face provided the object based spatial relationship. Proctor et al. (2003) used a display in which a target stimulus appeared in one of four corners and response keys were placed diagonally (top-right vs. bottom-left or top-left vs. bottom-right); they obtained a Simon effect for both horizontal and vertical dimensions concurrently. Based on these studies, it is predicted that both the 2-D location and the 3-D location should form different spatial stimulus codes, and both should produce the

Simon effect. The situation of the present experiment was similar to the study of Proctor et al. (2003). A joystick ordinarily used to navigate a plane in a 3-D flight simulator was chosen as the response device. The response made with the device implicates not only a 3-D spatial code but also a 2-D spatial code. Because most interfaces assigned to the control of a 2-D display use an arrangement in which inclining the joystick (or pushing a button) away from the user corresponds to the upper 2-D direction, the action of inclination to indicate "far" also contains the response code "upper" in the 2-D vertical location (**Figure 1**).

An alternative possibility is that the 2-D location has no or little effect on the speed of response. This hypothesis seems unlikely because the 2-D Simon effect has been repeatedly reported. Consider, however, the display in **Figures 2A,B**, in which a red object and a gray object are located at the same vertical height (i.e., on the same horizontal plane) in a 3-D scene. Because they differ in depth relative to the observer, they appear to have different vertical locations on the 2-D image due to projection. The present study focused on this 2-D vertical difference produced by projection from a 3-D scene onto a 2-D image. In such a case, it might be possible that the 2-D Simon effects are decreased, or in the extreme case, disappear, if the locations objects are coded in 3-D representation when they are embedded in 3-D space. Furthermore, the relative impact of 2-D and 3-D representations on the Simon effect may be modified by additional information about the context in which the stimuli are embedded in the 3-D environment. The present study therefore included a condition with a background which is composed of textured ground and ceiling planes. (**Figure 2C**). These planes are parallel to the horizontal plane, and positing a fixation and target objects on the same plane should strengthen a context in which the two objects are at the same vertical positions in 3-D space. If the existence of the background provides contextual information, as reported in previous studies (Hommel and Lippa, 1995; Stins and Michaels, 2010), the relative effects of 2-D and
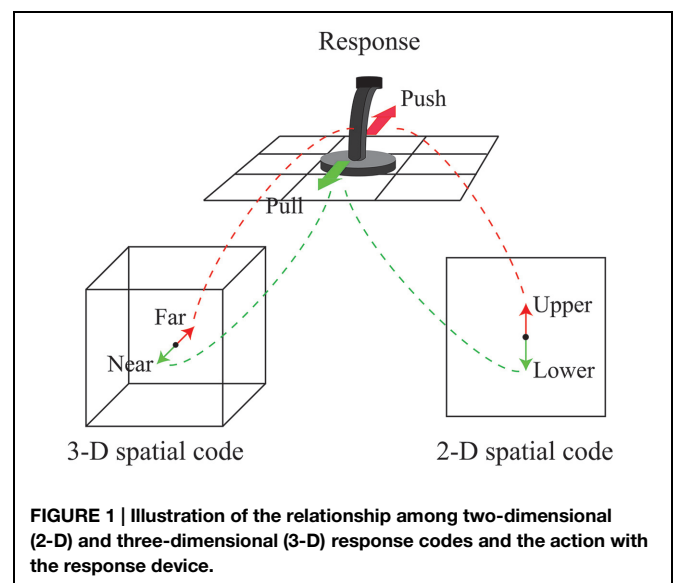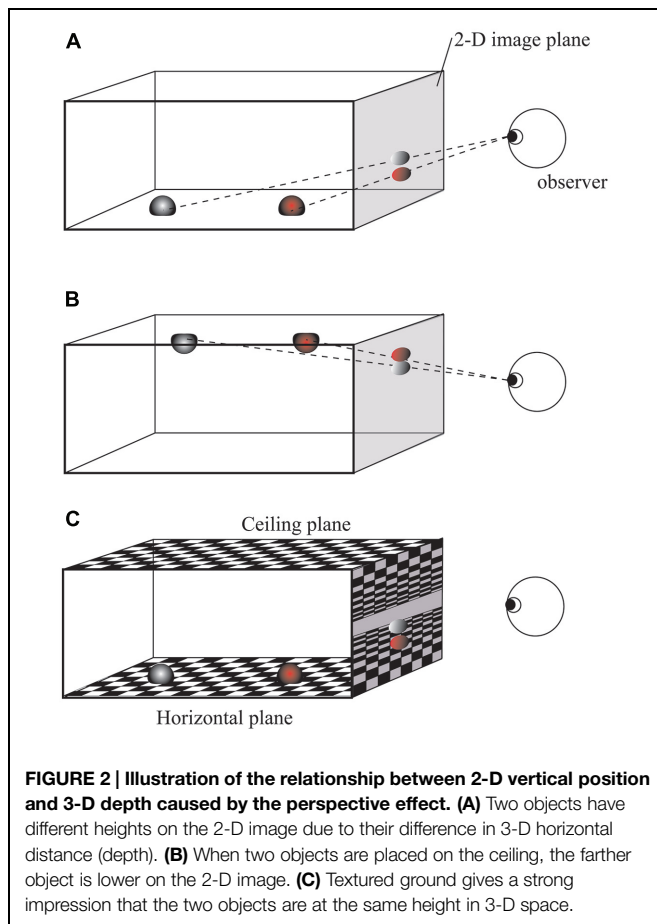

FIGURE 1 | Illustration of the relationship among two-dimensional (2-D) and three-dimensional (3-D) response codes and the action with the response device.

FIGURE 2 | Illustration of the relationship between 2-D vertical position and 3-D depth caused by the perspective effect. (A) Two objects have different heights on the 2-D image due to their difference in 3-D horizontal distance (depth). **(B)** When two objects are placed on the ceiling, the farther object is lower on the 2-D image. **(C)** Textured ground gives a strong impression that the two objects are at the same height in 3-D space.

3-D information should be altered. To examine this, ground and ceiling planes were present in Experiment 1 and absent in Experiment 2. In both experiments, 2-D vertical positions of the targets were produced by effects of perspective. When the background was absent, the impression that the fixation and target stimuli were located at the same height in the 3-D scene was reduced.

Materials and Methods

Participants

Eighteen people (12 males and 6 females) participated in both Experiments 1 and 2. All were between 20 and 30 years of age and had normal or corrected-to-normal vision. None were aware of the purpose of the present experiment. Written informed consent was obtained from all participants before the experiment. All the experimental procedures were approved by the Ethics Committee for Human and Animal Research of the National Institute of Advanced Industrial Science and Technology (AIST).
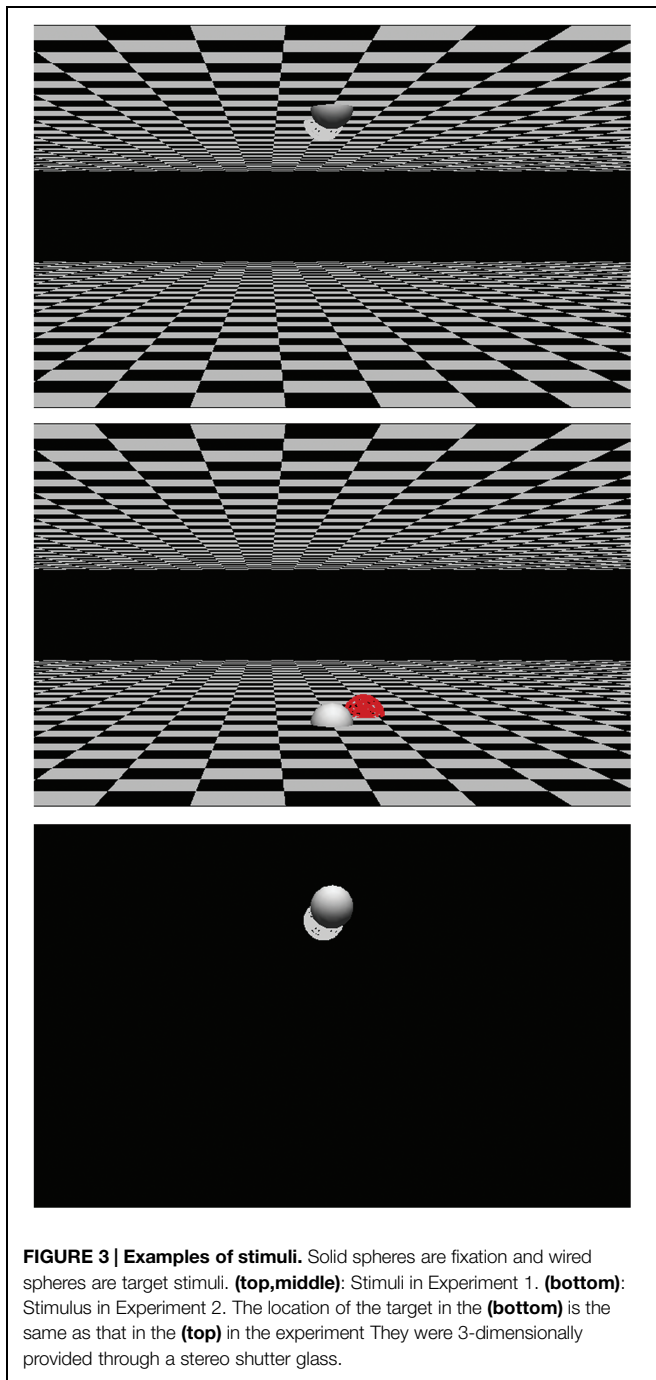
Apparatus and Setups

Experiments were conducted in a dark room. A Windows PC was used to control stimulus presentation on a CRT monitor (Sony 24″ GDMFW900) placed at 75 cm distance from the observer.

The height of the center of the display and the eye-height of each participant was adjusted by a chin-rest. To display the stimuli with binocular disparity, a shutter goggle (Stereographics, Crystal eyes) was used. The CRT's refresh rate was 85 Hz and its resolution was 1024 × 768. A joystick, Thrustmaster T-Flight stick X, was used for making responses. This joystick had a conventional design of a stick for aircraft, with a stem about 18 cm in height and 4 cm in diameter. The joystick was placed in front of the participant and the two response directions (near–far) were along their midline. Participants grasped the joystick stem with their dominant hand and held its base down with their non-dominant hand.

Stimuli and Tasks

In both Experiments 1 and 2, participants were required to respond to the color of the target stimulus. The target stimulus was presented at far or near depths relative to a fixation stimulus. The 2-D vertical location of the target stimulus (upper or lower) was produced according to the perspective effect (**Figure 3**). The location of the stimulus and the color of the stimulus were independently determined, and participants were required to ignore the location of the stimulus.

The positions of the fixation point and a target stimulus were determined in a 3-D space. The following procedure was used to draw a stimulus display in Experiment 1, in which the background (ground plane and ceiling plane) was present. Except for the presence of the background, the same procedure was used in Experiment 2. A ground plane and a ceiling plane were drawn 10 cm (in the 3-D scene) below and above eye height. They extended 50 cm away from the fixation and the observer (*far side* of the fixation) and 15 cm from the fixation toward the observer (*near side*). On the 2-D image, the planes created rectangles at a height of 5° of visual angle, with a 3° gap at the center of the image (**Figure 3**). These planes were textured with a checkerboard, and the gap was black. The fixation and target stimulus were placed on the same plane. The fixation was a solid sphere with a radius of 1 cm in 3-D space (0.8° of visual angle) and was placed on the ground or ceiling plane. It was located 75 cm away from the observer in 3-D and 7.6° below or above the center of the CRT. The target stimulus was a wire-framed sphere colored red or white. Its size was as same as that of the fixation stimulus in 3-D space, but this varied on the 2-D display according to perspective. The target stimulus was located on the near side or far side of the fixation. The 2-D vertical position of the target was determined by the combination of the depth of the target and the plane on which the target and the fixation were located. For example, when the fixation and the target stimulus were placed on the ceiling plane, the target stimulus on the far side appeared below the fixation on the 2-D image (**Figure 3**, top). In this case, if the correct response was pulling the joystick toward the observer (near), the 2-D vertical location was compatible but the 3-D depth was incompatible. The relationships among the depth/plane combinations and response-location compatibilities are summarized in **Table 1**. The distance of the target stimuli from the fixation in 3-D space varied from 4 to 8 cm far or near. As a result, the vertical distance between the target stimuli and fixation stimuli on 2-D display

FIGURE 3 | Examples of stimuli. Solid spheres are fixation and wired spheres are target stimuli. **(top,middle)**: Stimuli in Experiment 1. **(bottom)**: Stimulus in Experiment 2. The location of the target in the **(bottom)** is the same as that in the **(top)** in the experiment They were 3-dimensionally provided through a stereo shutter glass.

varied from 0.5 to 1.4° when the target was positioned on the near side in 3-D, and 0.4 to 1.0° at the upper when the target was positioned on the far side in 3-D. Horizontal distance also varied from −1.1 to 1.1° (2.5 cm in 3-D), but this was not analyzed.

On each trial, the fixation was displayed on the ground or ceiling plane (these planes were invisible in Experiment 2). The participants were required to gaze at the fixation and initiated a trial by pressing a button on the joystick. After 1.5 s, a target stimulus was displayed. Participants responded to

its color by pushing (inclining toward the far side) or pulling (inclining toward the near side) the joystick. The correspondence between the direction and the color was counterbalanced among participants; that is, half of the participants responded to the red stimulus by pushing the joystick and responded to the white stimulus by pulling. Participants were required to respond as fast and as accurately as possible.

In Experiments 1 and 2, participants were presented with each combination of two target depths (near or far) and two planes (ground or ceiling) 32 times, for a total of 128 trials. A rest break was provided after half of the trials were completed. All participants conducted both experiments; half participated in Experiment 1 first.

Results

Mean RTs and percent errors (PEs) are summarized in **Table 2**. RTs were categorized by three factors: the conflict between response direction and 2-D vertical location, the conflict between response and 3-D depth, and response direction (see **Table 1**). Here, the 2-D vertical location 'upper' was considered to be consistent with the response action of 'push,' and 'lower' with 'pull.' RTs for incorrect responses and for trials on which they were shorter than 150 ms or longer than 2000 ms were excluded. Repeated-measure ANOVAs with three within-subjects factors (two 3-D depths, two 2-D vertical locations, and two response directions) were conducted on RTs and PEs.

ANOVA for the RTs in Experiment 1, in which the ground and ceiling planes were displayed (**Table 2**, top), revealed significant main effects of 3-D depth [$F_{(1,17)} = 25.456$, $p < 0.001$, $\eta_p^2 = 0.6$] and 2-D vertical location [$F_{(1,17)} = 10.491$, $p < 0.005$, $\eta_p^2 = 0.382$], but not response direction [$F_{(1,17)} = 0.634$, $p = 0.437$, $\eta_p^2 = 0.036$]. All interactions between and among these factors were not significant. ANOVA for the PEs revealed significant main effects of 3-D depth [$F_{(1,17)} = 4.776$, $p = 0.043$, $\eta_p^2 = 0.219$] but not 2-D vertical location [$F_{(1,17)} = 0.676$, $p = 0.422$, $\eta_p^2 = 0.038$], and response direction [$F_{(1,18)} = 0.295$, $p = 0.594$, $\eta_p^2 = 0.017$]. Although significant main effects were observed only for effects of 3-D depth, the pattern of effects for PEs and RTs did not conflict.

In Experiment 2, in which the ground and ceiling planes were invisible (**Table 2**, bottom), ANOVA for the RTs revealed significant main effects of 3-D depth [$F_{(1,17)} = 15.201$, $p = 0.001$, $\eta_p^2 = 0.472$] and 2-D vertical location [$F_{(1,17)} = 18.523$, $p < 0.001$, $\eta_p^2 = 0.521$], but not response direction [$F_{(1,17)} = 0.826$, $p = 0.376$, $\eta_p^2 = 0.046$]. All interactions of these factors were not significant. ANOVA for the PEs revealed no significant main effects of 3-D depth [$F_{(1,17)} = 2.181$, $p = 0.158$, $\eta_p^2 = 0.114$], 2-D vertical position [$F_{(1,17)} = 1.527$, $p = 0.233$, $\eta_p^2 = 0.082$], or response direction [$F_{(1,17)} = 0.595$, $p = 0.451$, $\eta_p^2 = 0.07$]. Although no significant main effects were observed, the pattern of effects for PEs and RTs did not conflict.

To examine the effect of the presence of the background, a repeated measures ANOVA with three within-subjects factors

TABLE 1 | The relationships among response-location compatibilities and combinations of the depth of the target and the plane on which the target and the fixation were located.

Plane	Ground				Ceiling			
Target position relative to fixation	**Far**		**Near**		**Far**		**Near**	
Response direction	**Far**	**Near**	**Far**	**Near**	**Far**	**Near**	**Far**	**Near**
2-D location-response compatibility	+	−	−	+	−	+	+	−
3-D location-response compatibility	+	−	−	+	+	−	−	+

The symbols '+' indicates compatibility between the stimulus location and response and '−' indicates incompatibility.

TABLE 2 | Mean RTs, (in milliseconds), standard deviations (SDs), differences (ΔRT = RT for incompatible – RT for compatible), and percent errors (PEs) in Experiments 1 and 2 as a function of 3-D depth-response correspondence, 2-D vertical location-response correspondence, and correct response action.

	RT (SD)			PE	
	3-D location-response compatibility		ΔRT(3D)	**3-D location-response compatibility**	
2-D location-response compatibility	**+**	**−**		**+**	**−**
Experiment 1 (background was present)					
	Correct response: Near				
+	459 (51)	496 (67)	37	3.1	4.9
−	474 (58)	507 (53)	33	0.7	7.1
ΔRT(2D)	15	11			
	Correct response: Far				
+	464 (48)	494 (49)	30	1.6	5.4
−	484 (50)	509 (63)	25	4	6.7
ΔRT(2D)	20	15			
Experiment 2 (background was absent)					
	Correct response: Near				
+	462 (54)	481 (67)	19	3.6	6.1
−	477 (56)	515 (84)	38	4.7	4.2
ΔRT(2D)	15	34			
	Correct response: Far				
+	468 (56)	493 (58)	25	1.1	4
−	483 (70)	510 (73)	27	4.1	6.1
ΔRT(2D)	15	17			

The symbol '+' indicates compatibility between the stimulus location and response and '−' indicates incompatibility.

(presence of background, 3-D depth, and 2-D vertical location) was conducted on the RTs merged across the two response directions in each of the experiments; response directions were non-significant in the previous ANOVAs. These merged RTs are shown in **Figure 4**. The results of the ANOVA showed no significant effect of the presence of the background [$F(1,17) = 0.001$, $p > 0.5$]. The main effects of 2-D vertical location and 3-D depth were significant [$F(1,17) = 29.802$, $p < 0.001$, $F(1,17) = 26.236$, $p < 0.001$], and no significant interactions were observed for 2-D vertical location and 3-D depth [$F(1,17) = 0.302$, $p > 0.5$], presence of the background and 2-D vertical location [$F(1,17) = 0.676$, $p = 0.422$], and presence of the background and 3-D depth [$F(1,17) = 0.381$, $p > 0.5$]. This indicates that the presence of the background did not affect the 2-D and 3-D Simon effects.

Discussion

Experiments 1 and 2 both clearly showed that 3-D depth and 2-D vertical locations of the target affected the speed of responses. The lack of interaction between 3-D depth and 2-D vertical location indicated that the incompatibility of 2-D vertical location and response direction and those in 3-D depth had additive effects on the Simon effect. This suggests that the codes of the 2-D vertical location and the 3-D depth were independently and simultaneously formed, and these independently prolonged the response if each or both of the location codes conflicted with the response made by the joystick, in which 2-D vertical location and 3-D depth overlapped. Previous research has suggested that the human brain codes several spatial aspects (Lamberts et al., 1992; Hommel and Lippa, 1995; Proctor et al., 2003; Valle-Inclán et al.,

FIGURE 4 | Mean reaction times (RTs) for Experiment 1 (black bars) and Experiment 2 (white bars) as a function of location-response compatibility. Error bars indicate *SEs*. The symbol '+' indicates compatibility between the stimulus location and response, and '−' means incompatibility.
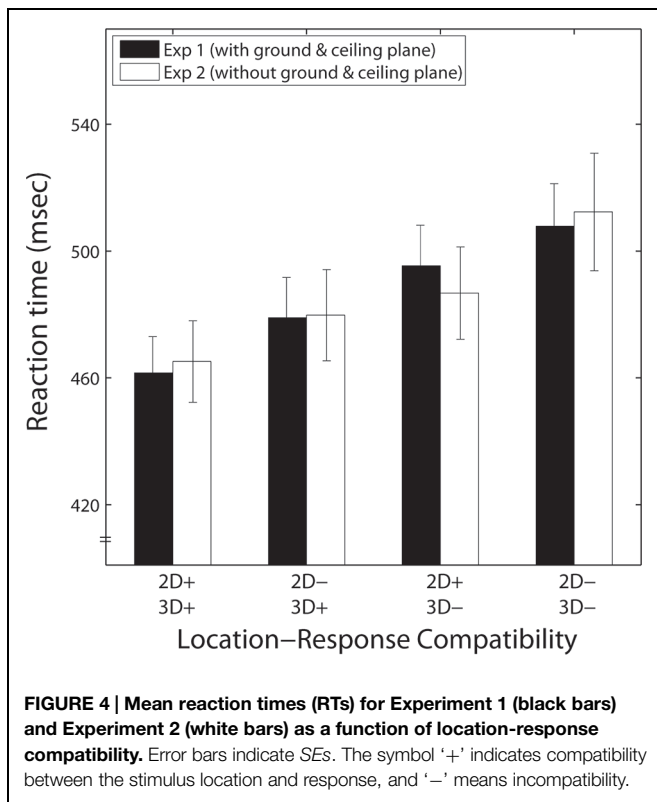
2003; Rubichi et al., 2006). 3-D depth should form another such spatial dimension, and a stimulus in 3-D space should be defined by a combination of these codes.

Generally, Experiments 1 and 2 produced similar patterns of response. One of the predicted outcomes was that the 2-D vertical locations would be less effective when the ceiling and ground planes were simultaneously displayed. That is, the existence of the ground and ceiling planes should decrease the Simon effect for the 2-D vertical locations, because these planes provide the context that the fixation and the target lay on the same plane. This was not observed. The presence of the background had no influence on the speed of response. The absence of these effects suggests that 2-D information is coded based on the retinal image rather than reconstructed 3-D scene. These results are interesting because previous research showed that contextually given location affected the speed of responses (Hommel and Lippa, 1995). This difference may have resulted because background played different roles in these studies. The context given by the background used in Hommel and Lippa (1995) (a picture of tilted face) contained a different (new) spatial frame, and this was superimposed on the original one (i.e., body-centered); spatial representations are coded based on both the new context and original frames. On the other hand, the context given by the background in the present study did not add a new frame. The context might weakened the relative impression of 2-dimensionality, but the compatibility effect, once triggered with the retinal image, could not be canceled by the contextual information acquired during the 3-D reconstruction processing that follows.

In both experiments, the effects of depth were larger than those of 2-D vertical location. The effect of 3-D depth, or the mean RT difference for trials in which 3-D depth and response direction conflicted versus those with no conflict [mean ΔRT(3D) in **Table 2**] was 31 ms in Experiment 1 and 27 ms in Experiment 2. On the other hand, the effect of 2-D location [mean ΔRT(2D) in **Table 2**] was 15 ms in Experiment 1 and 20 ms in Experiment 2. Although the size of the Simon effect was larger for 3-D depth than for 2-D vertical location, the relative strength between 2-D and 3-D locations cannot be determined, because the present experiment did not match the distances of these two directions. Shifts in 2-D vertical position were very small because they were produced by changes in relative height due to perspective. Proctor et al. (2003) examined whether the relative magnitudes of the horizontal and vertical Simon effects could be systematically altered by manipulating the relative distance of the horizontal and vertical dimensions. They showed that the magnitude of the vertical Simon effect changes with the manipulation of the distance. This means that further experiments are required to determine the relative strengths of the 2-D vertical Simon effect and the 3-D depth Simon effect. What is important in the present results is that the 2-D Simon effect was observed even with a small shift in 2-D vertical location.

In the present study, there was no significant difference between responses corresponding to 'near' and 'far' locations. Stins and Michaels (2010), who investigated the effect of the presence of a 2-D textured background on the S–R compatibility effect, reported an asymmetrical effect of depth direction. They found an S–R compatibility effect between the 3-dimensional 'far' location and the response made by moving a joystick away (toward 'far'), but the effect was not observed when the target was displayed at a near distance. This difference probably arose from the richness of 3-D depth information in the display. Unlike the display used by Stins and Michaels (2010), the display in the present experiment involved binocular disparity, relative size, and occlusion when necessary. The display of Stins and Michaels (2010) could not provide sufficient depth information in near space because the texture was sparse there. The richness of the 3-D information in the present display may also have contributed to the absence of a significant effect of the presence of the background. Even if the presence of background had no contextual effect, it could have been used as a cue for 3-D depth. However, it seems that binocular disparity and relative size provided sufficient 3-D information in the present experiment.

The main finding of the present study is that conflicts in 3-D depth and 2-D vertical location independently and simultaneously affected performance. The results are consistent with previous studies which showed the effect of multiple frames. With respect to the formation of spatial coding, there is a debate between the referential coding account (Umiltà and Nicoletti, 1985; Hommel, 1993, 2011) and the attentional shift account (Nicoletti and Umiltà, 1994; Proctor and Lu, 1994; Stoffer and Umiltà, 1997) about when and how the code is formed. Although the present experiments did not intend to reveal which of these accounts is preferable, the results are suggestive. The referential coding account (Umiltà and Nicoletti, 1985; Hommel, 1993, 2011) assumes that relative spatial coding is accomplished by

relating the target stimulus to reference frames or reference objects. The attentional shift account holds that a spatial code is generated through a shift of attention to the location of the target stimulus (Nicoletti and Umiltà, 1994; Stoffer and Umiltà, 1997). One important difference between the two accounts is that the attentional shift account cannot explain a Simon effect that occurs in multiple frames, because it assumes that there can be only one relative spatial code that is automatically generated for one attentional shift (Stoffer and Umiltà, 1997). Therefore, the attentional shift account seems to require some modification to provide an account for the present results. While, the referential-coding account can readily provide an account for the present results, because the account allows parallel coding of a stimulus if multiple frames of references are available.

It has been suggested that the difficulty of accounting for more than one active spatial code (e.g., spatial codes for horizontal and vertical locations) can be resolved by assuming that there are as many attention shifts as relative spatial codes (Stoffer and Umiltà, 1997). This seems applicable to the present results because the movement of attention in depth is known to occur (de Gonzaga Gawryszewski et al., 1987). However, the attentional shift account is based on the premotor theory of attention, in which programming of a saccadic eye movement toward a position is assumed to be necessary in order to shift attention (Rizzolatti et al., 1987, Umiltà et al., 1991, Stoffer and Umiltà, 1997, Van der Lubbe and Abrahamse, 2011). When attention shifts toward a stimulus, the program for the saccadic eye movement is prepared, and this oculomotor program for the saccade becomes a spatial code of the stimulus. Rigon et al. (2011) argued that the existence of the 3-D Simon effect does not support this account, because the saccadic eye movement should occur in the horizontal and vertical dimensions, but not in depth; therefore programming of vergence eye movements are necessary to account for the 3-D Simon effect. The present results also support this view. Reconciliation of the attentional shift account with the 3-D Simon effect may be based on programming of eye movements, including both saccadic and vergence eye movements. Saccadic eye movements should generate 2-D horizontal and vertical spatial codes, and vergence eye movements should generate 3-D depth codes. Yet it is unlikely that only one spatial code can account for the Simon effect. Thus, it seems valid to assume the simultaneous and parallel existence of 2-D and 3-D representations.

The present study reported that 2-D vertical location and 3-D depth are coded simultaneously and independently, and both of them affect response selection, in which 2-D and 3-D representations overlap. The present study used a CRT display with stereo shutter glasses, because these were suitable to control the experiment. The present results, however, should be confirmed in additional experiments in a real environment or in one using virtual reality techniques in which participants can move their heads.

Acknowledgment

References

Chan, A. H. S., and Chan, K. W. L. (2010). Three-dimensional spatial stimulus-response (S-R) compatibility for visual signals with hand and foot controls. *Appl. Ergo.* 41, 840–848. doi: 10.1016/j.apergo.2010.03.005

de Gonzaga Gawryszewski, L., Riggio, L., Rizzolatti, G., and Umiltá, C. (1987). Movements of attention in the three spatial dimensions and the meaning of "neutral" cues. *Neuropsychologia* 25, 19–29. doi: 10.1016/0028-3932(87)90040-6

Hommel, B. (1993). The role of attention for the Simon effect. *Psychol. Res.* 55, 208–222. doi: 10.1007/BF00419608

Hommel, B. (2011). The Simon effect as tool and heuristic. *Acta Psychol.* 136, 189–202. doi: 10.1016/j.actpsy.2010.04.011

Hommel, B., and Lippa, Y. (1995). S-R compatibility effects due to context-dependent spatial stimulus coding. *Psychon. Bull. Rev.* 2, 370–374. doi: 10.3758/BF03210974

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. doi: 10.1037/0033-295X.97.2.253

Lamberts, K., Tavernier, G., and d'Ydewalle, G. (1992). Effects of multiple reference points in spatial stimulus-response compatibility. *Acta Psychol.* 79, 115–130. doi: 10.1016/0001-6918(92)90028-C

Nicoletti, R., and Umiltà, C. (1994). Attention shifts produce spatial stimulus codes. *Psychol. Res.* 56, 144–150. doi: 10.1007/BF00419701

Proctor, R. W., and Lu, C. H. (1994). Referential coding and attention-shifting accounts of the Simon effect. *Psychol. Res.* 56, 185–195. doi: 10.1007/BF00419706

Proctor, R. W., Vu, K.-P. L., and Nicoletti, R. (2003). Does right-left prevalence occur for the Simon effect? *Percept. Psychophys.* 65, 1318–1329. doi: 10.3758/BF03194855

Rigon, J., Massaccesi, S., and Umiltà, C. (2011). A Simon effect for depth in three-dimensional displays. *Am. J. Psychol.* 124, 395–403. doi: 10.5406/amerjpsyc.124.4.0395

Rizzolatti, G., Riggio, L., Dascola, I., and Umiltá, C. (1987). Reorienting attention across the horizontal and vertical meridians: evidence in favor of a premotor theory of attention. *Neuropsychologia* 25, 31–40. doi: 10.1016/0028-3932(87)90041-8

Rubichi, S., Vu, K.-P. L., Nicoletti, R., and Proctor, R. W. (2006). Spatial coding in two dimensions. *Psychon. Bull. Rev.* 13, 201–216. doi: 10.3758/BF03193832

Simon, J. R., and Rudell, A. P. (1967). Auditory S-R compatibility: the effect of an irrelevant cue on information processing. *J. Appl. Psychol.* 51, 300–304. doi: 10.1037/h0020586

Stins, J. F., and Michaels, C. F. (2010). Stimulus-response compatibilityis information-action compatibility. *Ecol. Psychol.* 9, 25–45. doi: 10.1207/s15326969eco0901_2

Stoffer, T. H., and Umiltà, C. (1997). Spatial stimulus coding and the focus of attention in S-R compatibility and the Simon effect. *Adv. Psychol.* 118, 181–208. doi: 10.1016/S0166-4115(97)80033-7

Umiltà, C., and Nicoletti, R. (1985). "Attention and coding effects in S-R compatibility due to irrelevant spatial cues," in *Attention and Performance XI*, eds M. I. Posner and O. S. M. Marin (Hillsdale, NJ: Erlbaum), 457–471.

Umiltà, C., Riggio, L., Dascola, I., and Rizzolatti, G. (1991). Differential effects of central and peripheral cues on the reorienting of spatial attention. *Eur. J. Cog. Psychol.* 3, 247–267. doi: 10.1080/09541449108406228

Valle-Inclán, F., Hackley, S. A., and De Labra, C. (2003). Stimulus-response compatibility between stimulated eye and response location: implications for attentional accounts of the Simon effect. *Psychol. Res.* 67, 240–243. doi: 10.1007/s00426-003-0131-z

Van der Lubbe, R. H. J., and Abrahamse, E. L. (2011). The premotor theory of attention and the Simon effect. *Acta Psychol.* 136, 259–64. doi: 10.1016/j.actpsy.2010.09.007

Beauty and cuteness in peripheral vision

Kana Kuraguchi* and Hiroshi Ashida

Department of Psychology, Graduate School of Letters, Kyoto University, Kyoto, Japan

Guo et al. (2011) showed that attractiveness was detectable in peripheral vision. Since there are different types of attractiveness (Rhodes, 2006), we investigated how beauty and cuteness are detected in peripheral vision with a brief presentation. Participants ($n = 45$) observed two Japanese female faces for 100 ms, then were asked to respond which face was more beautiful (or cuter). The results indicated that both beauty and cuteness were detectable in peripheral vision, but not in the same manner. Discrimination rates for judging beauty were invariant in peripheral and central vision, while discrimination rates for judging cuteness declined in peripheral vision as compared with central vision. This was not explained by lower resolution in peripheral vision. In addition, for male participants, it was more difficult to judge cuteness than beauty in peripheral vision, thus suggesting that gender differences can have a certain effect when judging cuteness. Therefore, central vision might be suitable for judging cuteness while judging beauty might not be affected by either central or peripheral vision. This might be related with the functional difference between beauty and cuteness.

Keywords: attractiveness, peripheral vision, face, beauty, cuteness

Introduction

It is well known that an attractive face captures attention (Shimojo et al., 2003; Leder et al., 2010) and Guo et al. (2011) showed that attractiveness is even detectable in peripheral vision. They also discussed that low spatial frequency information can be used for judging attractiveness. These findings have suggested that judging attractiveness is possible even though available visual information can be limited. This is related to the idea that attractiveness is important for mate selection. Attractiveness, however, has both sexual and non-sexual aspects, such as attractiveness as a potential ally and cuteness in addition to sexual attractiveness (Rhodes, 2006). For example, we can point out the mere exposure effect (Zajonc, 1968; Peskin and Newell, 2004), that self-resemblance of same gender faces increase the attractiveness (DeBruine, 2004), and that smiling increases the attractiveness rating (Reis et al., 1990). These are mentioned as social attractiveness, which leads to establishment of the relationship of trust and aid with the other person, rather than sexual attractiveness. In fact, the dominant visual field (hemisphere) differs by the type of judged attractiveness. Sexual attractiveness (for a date situation) is more related to the left visual field (right hemisphere) while non-sexual attractiveness (as a lab partner) is more related to the right visual field (left hemisphere; Franklin and Adams, 2010). The attractiveness as a lab partner in Franklin and Adams (2010) is considered to be social attractiveness. Thus, it should be investigated how the different types of attractiveness are detected in peripheral vision.

Beauty consists of averageness, symmetry, and sexual dimorphism, all of which might show the quality of one's genes (Thornhill and Moller, 1997) and the state of one's health (Rhodes et al., 2001).

As a result, such aspects of beauty provide important information for mate selection and function as an innate component of attractiveness. Conversely, cuteness represents the attractiveness of infants (Karraker and Stern, 1990) and is related to the baby schema concept (Hildebrandt and Fitzgerald, 1979; Alley, 1981). This concept elicits caregiving behaviors, which has been proven effective even for adult faces (Keating et al., 2003). In this regard, cuteness functions as social attractiveness in interactions with other people. Accordingly, it is possible that beauty and cuteness might reveal different aspects of attractiveness (Geldart, 2010). Attractive female faces, however, possess both neonate and sexually dimorphic features (Cunningham, 1986; Pfluger et al., 2012), therefore the criteria of judging beauty and cuteness might be overlapping. Kuraguchi et al. (2015) showed that beauty and cuteness might represent different aspects of attractiveness even though similar facial features affected both judgments to some degree. Therefore, the first aim of this study is to investigate whether beauty and cuteness are distinguished through availability in peripheral vision. We also aimed to extract the difference between beauty and cuteness from participant's natural responses without the definition of two word meanings, in order to investigate whether sexual and social attractiveness is distinguished even in the expression used in daily life.

Moreover, the effect of gender was not considered by Guo et al. (2011). While one study showed that participants can assess the attractiveness of both male and female faces in a similar manner regardless of their gender or sexual orientation (Kranz and Ishai, 2006), it has been argued that facial attractiveness provides the adaptive benefits (i.e., signals of good genes) for different-sex observers, but not for same-sex observers (Senior, 2003). It was also reported that women were more sensitive in perceiving cuteness than men (Sprengelmeyer et al., 2009). Therefore, our second aim is to examine whether there are gender differences in judging beauty and cuteness in peripheral vision.

Furthermore, Japanese people tend to confuse beauty with cuteness (Daibo, 2007). However, as mentioned earlier, these two characteristics are considered to represent different aspects of attractiveness. It was also reported that beauty and cuteness were distinguished in the North American culture (Geldart, 2010). It is possible that peripheral viewing might reveal the difference between beauty and cuteness for Japanese people. Therefore, our third aim is to investigate how Japanese people perceive beauty and cuteness in peripheral vision.

Experiment 1

Methods
Participants
The participants consisted of 45 Japanese students (18–25 years; 22 males and 23 females) with normal vision or corrected to normal vision, who were naïve to the experimental purposes and asked to view the stimulus presentation. The distance between the eyes and the display was 56 cm, and a chin rest was used to stabilize the head. We obtained written informed consent from all the participants, and each individual was paid according to

the standards of Kyoto University. During the experiment, their fixations were monitored with the Tobii T120 Eye Tracker.
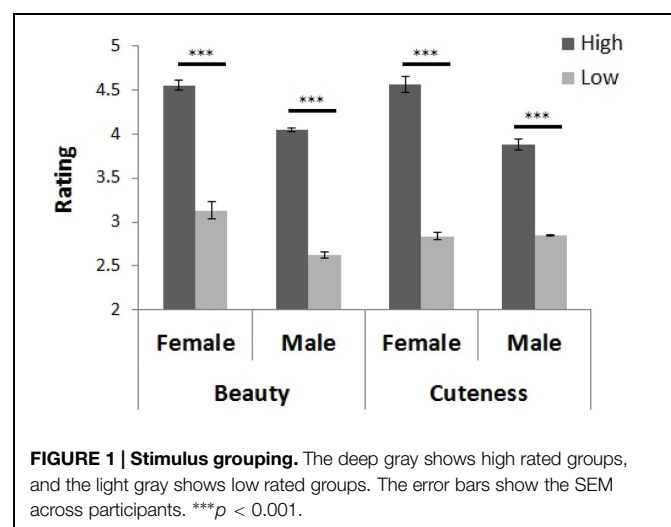
Stimuli
Visual stimuli were presented on a 17-inch LCD monitor (Tobii T120, 1280 × 1024 pixels, 60 Hz), and the average luminance of the stimuli was 9.94 cd/m^2. Images of 10 Japanese female faces (18–24 years) were presented in a visual angle of 9° and by 6°. All the faces included frontal views with neutral expressions. The images were gray scale and cropped to remove external features (e.g., hair style) and they were classified into high and low groups of five faces each through a preliminary test for both beauty and cuteness in which a separate group of participants ($n = 29$) rated the images on a 6-point scale. These groups were used for the discrimination rate analysis. As **Figure 1** shows, we found significant differences in the mean rating scores of five faces between the high and low groups in both beauty [$t(4) = 14.168, p < 0.001$] and cuteness [$t(4) = 12.947, p < 0.001$]. The stimulus groups for beauty and cuteness judgments actually included exactly the same faces for female participants, and only one different face for male participants.
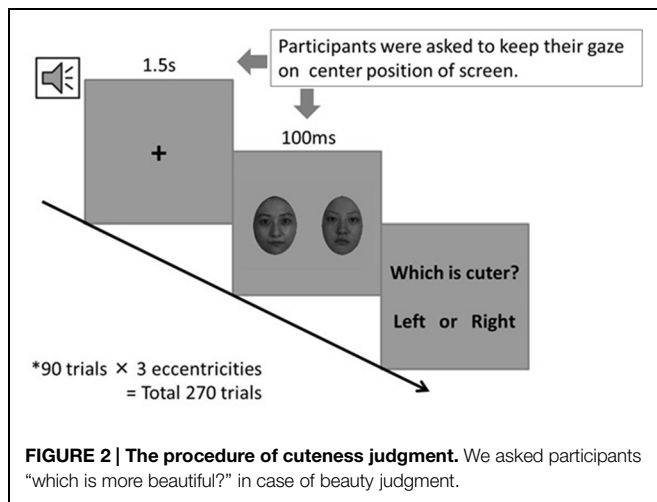
Judgment Conditions
Judgments included two conditions, beauty or cuteness, as a between-participant factor (beauty: 9 males and 10 females; cuteness: 13 males and 13 females). The participants were asked to choose the more beautiful or cuter face out of the two facial images simultaneously presented, and respond by pressing a key after the faces disappeared.

Procedure
Each trial began with a warning tone, followed by a central fixation cross shown for 1.5 s. In the same manner as Guo et al. (2011), a pair of faces was presented for 100 ms to the left and right with equal distances from the central fixation cross. The participants were asked to maintain their visual fixation and respond by pressing one of the two keys to indicate which face was more beautiful or cuter after the faces disappeared. There was no need

FIGURE 1 | Stimulus grouping. The deep gray shows high rated groups, and the light gray shows low rated groups. The error bars show the SEM across participants. ***$p < 0.001$.

FIGURE 2 | The procedure of cuteness judgment. We asked participants "which is more beautiful?" in case of beauty judgment.

for a quick response. **Figure 2** presents the aforementioned procedure. Pairs of faces were made by combining round-robbin ($_{10}C_2 = 45$ pairs), with the presented sides (left/right) counterbalanced (90 patterns in total). The data of high–low pairs were analyzed, and the other half (high–high and low–low combinations) were discarded.

The viewing eccentricity was 2°, 5°, and 10° (from the center to the inner edge of the faces), to probe foveal, parafoveal, and peripheral vision, respectively, as in Guo et al. (2011). The number of trials was 90 per eccentricity (270 in total) for each participant, tested in a random order. After this main experiment, all the participants performed a self-paced task without fixation and with unlimited presentation time. In the self-paced task, a pair of faces was presented to the left and right with equal distances (2°) from the central fixation cross, and the participants were asked to respond by pressing one of the two keys to indicate which face was more beautiful or cuter while looking at faces.

Results

Consistency of Rating Beauty and Cuteness

In order to confirm the consistency for rating beauty and cuteness between the main and preliminary experiments, the facial images were ranked based on the results of the self-paced judgments by using paired comparison (Thurston's method). In addition, the ranks of the faces were compared to those from the preliminary experiment. Significant rank correlations were found for judging both beauty and cuteness regardless of the participant's gender ($ps < 0.001$). Accordingly, the groupings of the high and low rated for both the judgments were consistent. Furthermore, the high- and low-rated faces did not exchange with one another for both beauty and cuteness.

Analysis of the Discrimination Rates

We checked the participants' fixation in stimulus presentation, and discarded all data of the participants whose fixation exceeded 1° for more than 5% of the total looking time. We therefore discarded the data from 14 participants whose fixation was unstable, and analyzed the data of 31 participants (beauty: nine males and nine females; cuteness: six males and seven females).

Discrimination rate was defined as the rate of responses that were congruent with the pre-defined high and low groups. For beauty judgment, all the discrimination rates were significantly above the chance level (50%), regardless of the participant's gender [males: 2° $t(9) = 3.56$, $p = 0.006$, 5° $t(9) = 3.63$, $p = 0.005$, 10° $t(9) = 4.87$, $p < 0.001$; females: 2° $t(9) = 6.09$, $p < 0.001$, 5° $t(9) = 6.98$, $p < 0.001$, 10° $t(9) = 5.54$, $p < 0.001$]. For cuteness judgment, all the discrimination rates of the female participants were significantly above chance level [2° $t(9) = 4.04$, $p < 0.001$, 5° $t(9) = 4.96$, $p < 0.001$, 10° $t(9) = 3.80$, $p = 0.004$], but the discrimination rate of male participants at the eccentricity of 10 was not above-chance [2° $t(9) = 4.92$, $p < 0.001$, 5° $t(9) = 4.64$, $p = 0.001$, 10° $t(9) = 1.55$, $p = 0.155$].
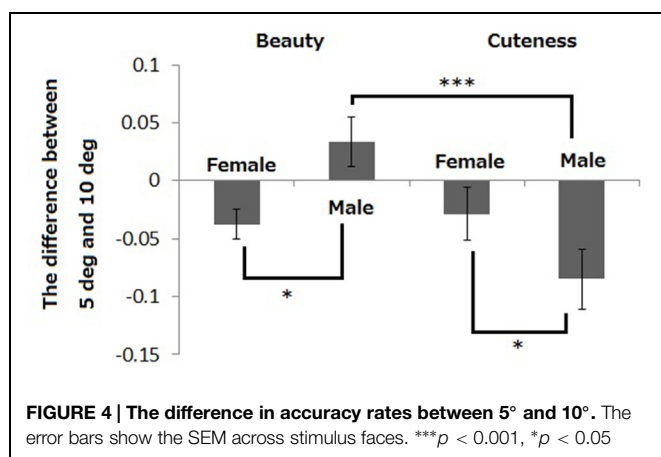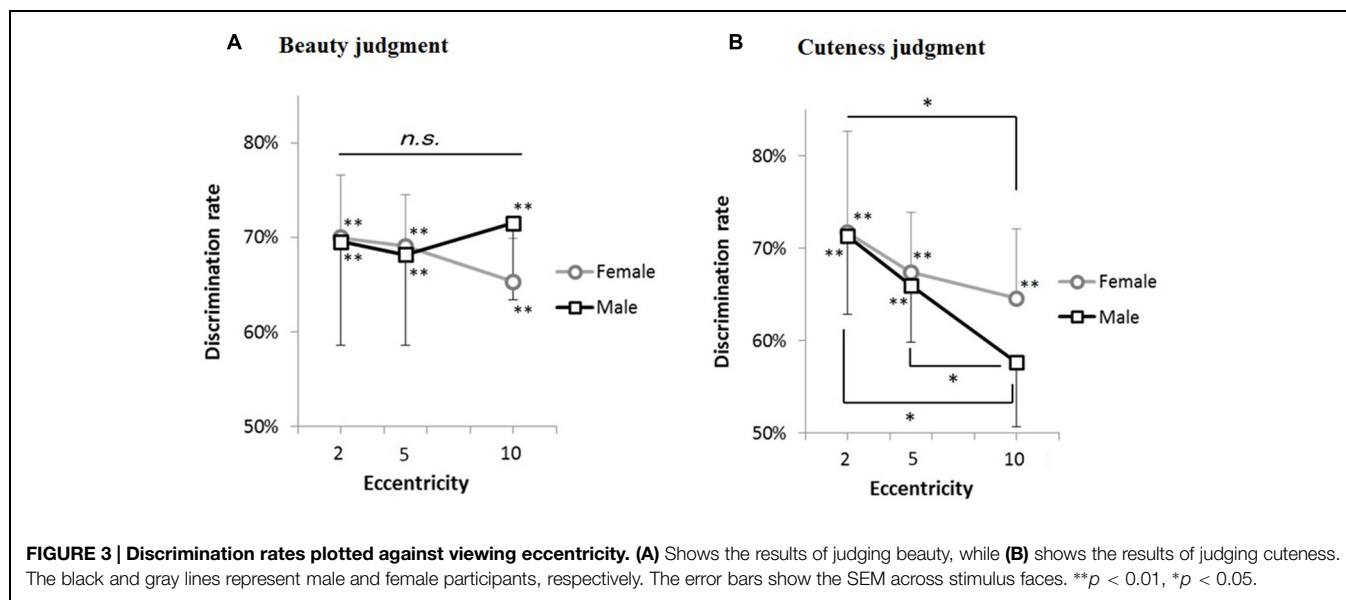
We then conducted a three-way ANOVA (2: judgment, 2: gender difference of participants, 3: eccentricity). The main effect of eccentricity [$F(2,72) = 11.28$, $p < 0.001$], the interaction between judgment and eccentricity [$F(2,72) = 6.84$, $p = 0.001$], and the interaction among judgment, gender, and eccentricity [$F(2,72) = 4.59$, $p = 0.013$] were significant. A Mendoza's multisample sphericity test revealed that sphericity assumption was satisfied ($p = 0.25$). A simple main effect test for the interaction between judgment and eccentricity revealed the effect of eccentricity on judging cuteness [$F(2,72) = 17.80$, $p < 0.001$]. A multiple comparison test (Ryan's method) revealed significant differences between 2° and 5°, between 2° and 10°, and between 5° and 10° ($ps < 0.05$). Simple interaction between judgment and eccentricity of the male participants was found [$F(2,72) = 10.94$, $p < 0.001$]. We also found the simple–simple main effect of judgment for the male participants in the visual angle of 10 [$F(1,108) = 6.19$, $p = 0.014$], the effect of eccentricity on judging cuteness for the male participants [$F(2,72) = 15.71$, $p < 0.001$], and the effect of eccentricity on judging cuteness for the female participants [$F(2,72) = 4.17$, $p = 0.019$]. A multiple comparison test (Ryan's method) revealed significant differences between 2° and 10°, and between 5° and 10° for the male participants ($ps < 0.05$), and significant differences between 2° and 10° for the female participants ($p < 0.05$). No such simple effects were found for beauty judgments. These results are summarized in **Figure 3**.

For further support of the gender difference in the decline from 5° to 10°, we conducted another statistical analysis on the difference between the results of 5° and 10°. A two-way ANOVA (2: judgment type, 2: participants' gender) revealed significant interaction [$F(1,9) = 8.68$, $p = 0.016$]. Following simple main effect analyses revealed the effect of judgments on male participants [$F(1,18) = 13.26$, $p = 0.001$], and the effect of gender differences both on judging cuteness [$F(1,18) = 4.61$, $p = 0.045$] and beauty [$F(1,18) = 7.32$, $p = 0.014$]. Gender difference is evident in judging not only cuteness but also beauty (see **Figure 4**).

Discussion

Beauty was judged correctly in all eccentricities above the chance level, regardless of the participant's gender. This result showed that beauty is detectable in peripheral vision, thus replicating the results of Guo et al. (2011), and beauty judgment was hardly affected by eccentricity.

However, the judgment of cuteness was affected by eccentricity. We found a significant difference between central vision (2°)

FIGURE 3 | Discrimination rates plotted against viewing eccentricity. (A) Shows the results of judging beauty, while **(B)** shows the results of judging cuteness. The black and gray lines represent male and female participants, respectively. The error bars show the SEM across stimulus faces. $**p < 0.01$, $*p < 0.05$.

FIGURE 4 | The difference in accuracy rates between 5° and 10°. The error bars show the SEM across stimulus faces. $***p < 0.001$, $*p < 0.05$

and peripheral vision (10°), regardless of the participant's gender. Judging cuteness in central vision is more accurate than in peripheral vision, even though judging cuteness is partly possible in peripheral vision. In addition, central vision could be more suitable for judging cuteness, and based on the fixation data, more participants were excluded in judging cuteness than in judging beauty, which also supports our finding that judging cuteness is difficult in peripheral vision.

Furthermore, gender difference was found in accuracy rates between 5° and 10°. In judging cuteness, the performance of males significantly declined more than that of females. In judging beauty, the performance of females declined more than that of males, while the overall decline was not significant for either gender (**Figure 3**). Significant difference was also found between beauty and cuteness in the performance of males. Accordingly, males were able to judge beauty but not cuteness in peripheral vision, while females were able to judge both beauty and cuteness. This also highlights the difference between beauty and cuteness in peripheral vision.

Then what is the cause of such a difference in central and peripheral vision? An obvious factor is blurred retinal images at the periphery. Therefore, in Experiment 2, we showed blurred images of faces in central vision that matched the perceptual blur at each eccentricity in order to investigate whether the aforementioned results can be explained solely by the blurred image.

Experiment 2

We tested the effect of blurred faces on judging beauty or cuteness.

Methods
Participants
The participants consisted of 31 Japanese students (18–36 years: 16 males, 15 females), with normal vision or corrected to normal vision, who were naïve to the experimental purposes and asked to view the stimulus presentation. No one had participated in Experiment 1. The distance between the eyes and the display was 48 cm and a chin rest was used to stabilize the head. We obtained written informed consent from all the participants, and each individual was paid according to the standards of Kyoto University.

Stimuli
The 10 facial images used in Experiment 1 were blurred by convolution with a 2-D Gaussian kernel of variable SD by using GNU Octave. All the faces were presented at the center of an LCD screen (Mitsubishi 23′ LCD) with a visual angle of 9° by 6° as in Experiment 1. Eye movement was not monitored. The luminance profile of the monitor was measured and was taken into account in blurring the images.

First, we conducted an experiment to estimate the points of subjective equality (PSE) for blurred faces that correspond to

each eccentricity. A separate group of 14 Japanese students (19–25 years: six males, eight females) participated. SuperLab 4.5 for Windows (Cedrus, Inc.,) was used to control the experiment. We presented one of the blurred images at the center of the screen and the original (not blurred) image at one of the eccentricities (2°, 5°, and 10°), either to the left or to the right, for 100 ms. Participants were asked to compare the two faces and judge if the central image appeared clearer than the peripheral one. Eight levels of blurred images were made for the two faces. Each image was repeated 10 times at the three eccentricities and at the two sides in a random order (960 trials in total). The PSEs were calculated for individual participants as the 50% level of the psychometric function that was estimated by the probit analysis, using the glm() function of R language. On average, the Gaussian half width at half maximum (HWHM) of 31.72 cycle/face-width (c/fw) corresponded to the eccentricity of 2°, 30.68 c/fw corresponded to the eccentricity of 5°, and 28.97 c/fw corresponded to the eccentricity of 10 (see **Figure 5**).

Judgment Conditions

Judgments were made for two conditions, beauty, or cuteness, as the factor between participants. Participants were asked to judge the beauty or cuteness on a 6-point scale, ranging 1 (e.g., not cute) to 6 (e.g., very cute), and respond by pressing a key after the faces disappeared.

Procedure

Each trial began with a central fixation and the participants pressed a key while viewing this fixation. A blurred face or not-blurred face was then presented at the center of screen for 100 ms. Participants were asked to fix their eyes on the face stimulus and judge beauty or cuteness on a 6-point scale presented on the display after the face disappeared, and respond by pressing one of the six keys. The number of trials was 120 trials per participant [10 images × 4 eccentricity equivalents (0°, 2°, 5°, and 10°) × 3], tested in a random order. SuperLab 4.5 for Windows (Cedrus, Inc.,) was used to control the experiment.

Results

We conducted a three-way ANOVA [2: stimulus group (high rating/low rating), 2: participants' gender, 4: blurring (eccentricity equivalents)] on the mean rating values of five images for beauty

and cuteness ratings. In cuteness judgment, the main effect of stimulus group was significant [($F(1,13) = 160.473$, $p < 0.001$], whereas the other effects or interactions were not significant ($p > 0.10$, see **Figure 3**). Moreover, in beauty judgment, the main effect of stimulus group was significant [$F(1,14) = 201.763$, $p < 0.001$], whereas the other effects or interactions were not significant [$p > 0.10$, see **Figure 3**].

Discussion

The pattern of results in **Figure 6** is clearly distinct from those in **Figure 3**. First, the results of beauty and cuteness judgments in Experiment 2 were similar to each other, which indicate that image blur is not the primary cause of the differences between beauty and cuteness in peripheral vision. Second, participants could judge facial beauty and cuteness of the blurred faces as well as the original ones, regardless of the participant's gender. This also suggests that image blur is not the primary cause of lower cuteness ratings in the periphery.

General Discussion

Judging Beauty and Cuteness in Peripheral Vision

In Experiment 1, we confirmed that beauty is detectable in both central and peripheral vision (Guo et al., 2011), while it also revealed that central vision is more suitable for judging cuteness. However, judging cuteness is more difficult in peripheral vision, especially for male participants.

In Experiment 2, we showed that this difficulty in judging cuteness did not vary with the level of blurred faces at each eccentricity. In addition, no gender difference was found. The procedural difference of presentation method between Experiment 1 and 2 (comparing two faces vs. rating single face) might affect the results. However, significant rank correlations between Experiment 1 (self-paced judgment) and Experiment 2 (each level of blurring) were found for judging both beauty and cuteness regardless of the participant's gender ($ps < 0.001$). Therefore, it is hardly to say that the difference of stimulus presentation affect the judgments. Another important feature of peripheral vision was the weaker response to color, but this was not relevant since we used grayscale images. Therefore, the difficulty in judging cuteness in peripheral vision and the related gender difference cannot be explained by the image properties. For example, a simple hypothesis that cuteness depends on higher spatial frequency information more than beauty should be rejected. Beauty reflects averageness and symmetry (Rhodes, 2006). These sets of information might be readily available in peripheral vision as beauty judgment was not much affected. Since the high-beauty faces were almost the same as the high-cuteness ones, no difference should have been observed between beauty and cuteness if the participants had relied upon the same accessible arrangement of facial features. The difference between beauty and cuteness found in this study, accordingly, indicates that beauty and cuteness judgment should rely on different features, and that cuteness relies on features that are not accessible in peripheral vision.
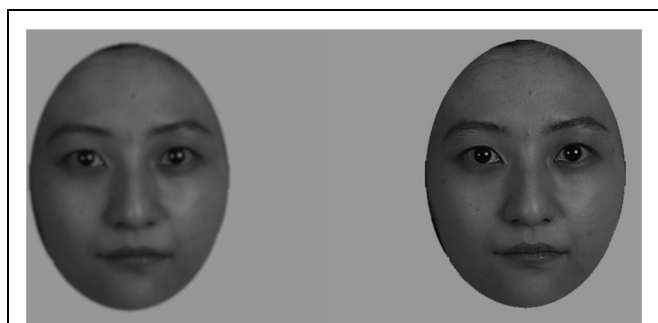

FIGURE 5 | Sample facial images. (Left) A blurred image corresponding to the eccentricity of 10°. **(Right)** The image without blur.
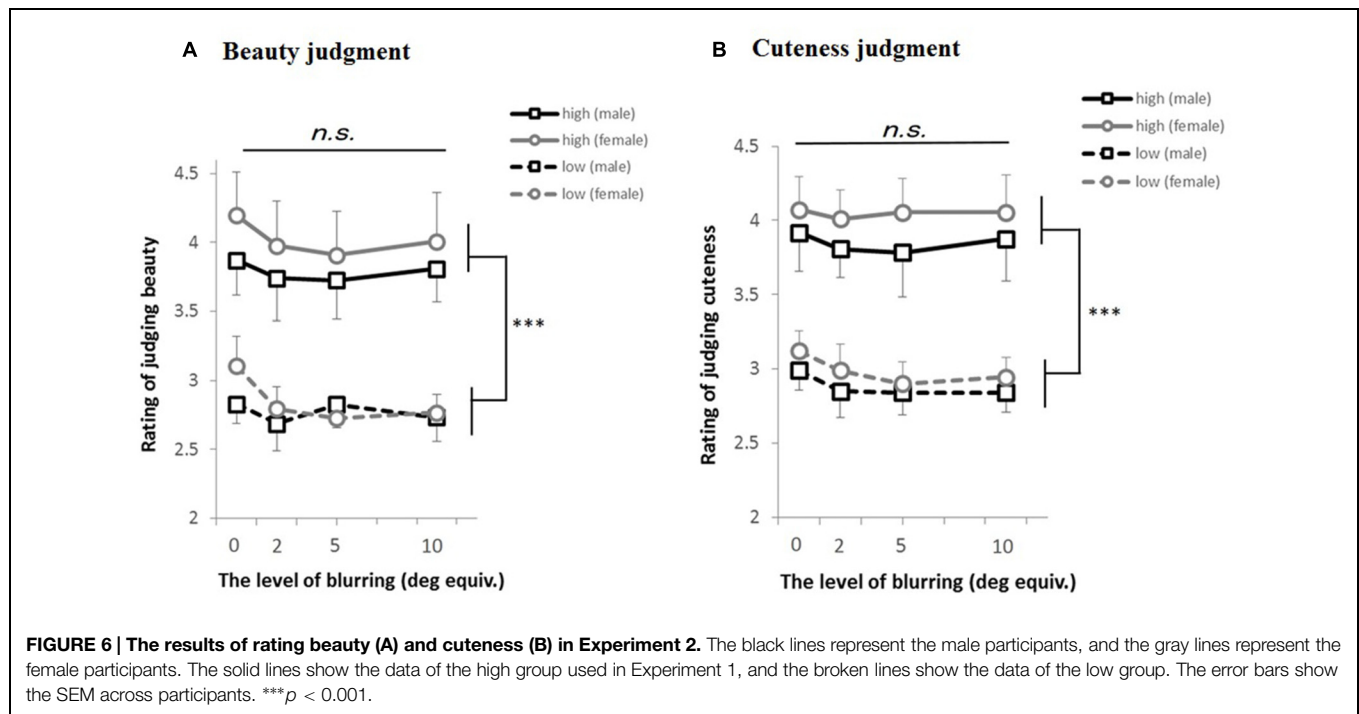
FIGURE 6 | The results of rating beauty (A) and cuteness (B) in Experiment 2. The black lines represent the male participants, and the gray lines represent the female participants. The solid lines show the data of the high group used in Experiment 1, and the broken lines show the data of the low group. The error bars show the SEM across participants. ***$p < 0.001$.

Beauty was detectable in peripheral vision as well as in central vision, regardless of the participant's gender. Conversely, cuteness was more difficult to detect in peripheral vision than in central vision, which was more pronounced in the male participants. While the underlying mechanisms are still open for further investigation, we can also understand these differences from a functional perspective. Beauty consists of averageness and symmetry, which function as indices of one's state of health (Rhodes et al., 2001) and one's quality of genes (Thornhill and Moller, 1997) for mate selection. Therefore, it is ecologically adaptive to first find a beautiful face in peripheral vision and then direct attention to the person. In fact, it has been reported that greater attention is directed toward a more attractive face (Shimojo et al., 2003; Leder et al., 2010). On the other hand, cuteness is evolutionally related to caregiving behaviors (Lorenz, 1943). Therefore, receiving cuteness may make the receiver observe carefully and concentrate on the cute object. This assumption is supported by the finding that cuteness can improve the performance of certain tasks that need attention (Nittono et al., 2012). However, averted attention may cause careless behaviors of the caregiver. Therefore, central vision is essential for cuteness, while judging cuteness in peripheral vision may be less important.

Gender and Cultural Effects on Cuteness

Japanese people tend to confuse being beautiful (*utsukushii*) with being cute (*kawaii*; Daibo, 2007). In this study, the high-beauty and the high-cuteness groups consisted of almost the same facial images, which suggest that beauty and cuteness were combined by the Japanese participants to some extent. Conversely, judgment of cuteness was significantly affected by peripheral viewing, whereas that of beauty was not, thus indicating that the Japanese participants did not completely confuse the two aspects.

The effect of eccentricity on cuteness judgment was particularly observed in the male participants. The possible reasons for this gender difference are as follows.

First, we used only female faces, which might have led to the asymmetric results. However, there is no straightforward reason to assume that people are more sensitive to cuteness of the same gender given that the female participants performed better. If the males had performed better in beauty judgment, then we could have argued that detecting beauty of the opposite gender quickly is advantageous in terms of mate selection. There was actually a slight tendency in which the males were better at judging beauty at 10° eccentricity than the females (**Figure 3**), and this was statistically supported by the difference in accuracy rates between 5° and 10°. However, the result that the males performed worse than females in cuteness judgment rejects the mate-selection-based explanation that males somewhat confused beauty with cuteness.

Second, it has been reported that females are more sensitive in perceiving cuteness due to female hormones (Sprengelmeyer et al., 2009). Our results suggest that females may have wider field of view in regard to cuteness judgment even though it becomes more difficult to perceive cuteness in peripheral vision. This may reflect the level of female hormones or it may be related to the general tendency that females play a more important role in caregiving behaviors. However, further investigations are needed for these suggestions.

Conclusion

Our results showed that judging beauty is invariant in peripheral and central vision, while judging cuteness is

degraded in peripheral vision. In addition, it was more difficult for the male participants to judge cuteness in peripheral vision, thus suggesting that gender differences can have a certain effect when judging cuteness. Finally, lower resolution in peripheral vision should not be the main cause of the tendency for cuteness (as described earlier) and central vision might be essential for judging cuteness while judging beauty could be detected more widely in peripheral vision. These results might be related to the functional difference between beauty and cuteness.

References

Alley, T. R. (1981). Head shape and the perception of cuteness. *Dev. Psychol.* 17, 650–654. doi: 10.1037/0012-1649.17.5.650

Cunningham, M. R. (1986). Measuring the physical in physical attractiveness: quasi-experiments on the sociobiology of female facial beauty. *J. Personal. Soc. Psychol.* 50, 925–935. doi: 10.1037/0022-3514.50.5.925

Daibo, I. (2007). A scope of cultural consideration of facial communication in social context. *Jpn. J. Inter. Soc. Psychol.* 7, 1–10.

DeBruine, L. M. (2004). Facial resemblance increases the attractiveness of same-sex faces more than other-sex faces. *Proc. Biol. Sci.* 271, 2085–2090. doi: 10.1098/rspb.2004.2824

Franklin, R. G., and Adams, R. B. (2010). The two sides of beauty: laterality and the duality of facial attractiveness. *Brain Cogn.* 72, 300–305. doi: 10.1016/j.bandc.2009.10.002

Geldart, S. (2010). That woman looks pretty, but is she attractive? Female perceptions of facial beauty and the impact of cultural labels. *Eur. Rev. Appl. Psychol.* 60, 79–87. doi: 10.1016/j.erap.2009.12.003

Guo, K., Liu, C. H., and Roebuck, H. (2011). I know you are beautiful even without looking at you: discrimination of facial beauty in peripheral vision. *Perception* 40, 191–195. doi: 10.1068/p6849

Hildebrandt, K. A., and Fitzgerald, H. E. (1979). Facial feature determinants of perceived infant attractiveness. *Infant Behav. Dev.* 2, 329–339. doi: 10.1016/S0163-6383(79)80043-0

Karraker, K. H., and Stern, M. (1990). Infants' physical attractiveness and facial expression: effects on adults' perceptions. *Basic. Appl. Soc. Psych.* 11, 371–385. doi: 10.1207/s15324834basp1104_2

Keating, C. F., Randall, D. W., Kendrick, T., and Gutshall, K. A. (2003). Do babyfaced adults receive more help? The (cross-cultural) case of the lost resume. *J. Nonverbal Behav.* 27, 89–109. doi: 10.1023/A:1023962425692

Kranz, F., and Ishai, A. (2006). Face perception is modulated by sexual preference. *Curr. Biol.* 16, 63–68. doi: 10.1016/j.cub.2005.10.070

Kuraguchi, K., Taniguchi, K., and Ashida, H. (2015). The impact of baby schema on perceived attractiveness, beauty, and cuteness in female adults. *Springerplus* 4, 164. doi: 10.1186/s40064-015-0940-8

Leder, H., Tinio, P. P. L., Fuchs, I. M., and Bohrn, I. (2010). When attractiveness demands longer looks: the effect of situation and gender. *Q. J. Exp. Psychol.* 63, 1858–1871. doi: 10.1080/17470211003605142

Lorenz, K. (1943). *Die Angeborenen Formen Moglicher Erfahrung* [Innate Forms of Potential Experience]. *Z. Tierpsychol.* 5, 235–409.

Nittono, H., Fukushima, M., Yano, A., and Moriya, H. (2012). The power of kawaii: viewing cute images promotes a careful behavior and narrows attentional focus. *PLoS ONE* 7:e46362. doi: 10.1371/journal.pone.0046362

Peskin, M., and Newell, F. N. (2004). Familiarity breeds attraction: effects of exposure on the attractiveness of typical and distinctive faces. *Perception* 33, 147–157. doi: 10.1068/p5028

Pfluger, L. S., Oberzaucher, E., Katina, S., Holzleitner, I. J., and Grammer, K. (2012). Cues to fertility: perceived attractiveness and facial shape predict reproductive success. *Evol. Hum. Behav.* 33, 708–714. doi: 10.1016/j.evolhumbehav.2012.05.005

Reis, H. T., Wilson, I. M., Monestere, C., Bernstein, S., Clark, K., Seidl, E., et al. (1990). What is smiling is beautiful and good. *Eur. J. Soc. Psych.* 20, 259–267. doi: 10.1002/ejsp.2420200307

Rhodes, G. (2006). The evolutionary psychology of facial beauty. *Annu. Rev. Psych.* 57, 199–226. doi: 10.1146/annurev.psych.57.102904.190208

Rhodes, G., Zebrowitz, L. A., Clark, A., Kalick, S. M., Hightower, A., and Mckay, R. (2001). Do facial averageness and symmetry signal health? *Evol. Human Behav.* 22, 31–46. doi: 10.1016/S1090-5138(00)00060-X

Senior, C. (2003). Beauty in the brain of beholder. *Neuron* 38, 525–528. doi: 10.1016/S0896-6273(03)00293-9

Shimojo, S., Simon, C., Shimojo, E., and Scheier, C. (2003). Gaze bias both reflects and influences preference. *Nat. Neurosci.* 6, 276–281. doi: 10.1038/nn1150

Sprengelmeyer, R., Perrett, D. I., Fagan, E. C., Cornwell, R. E., Lobmaier, J. S., Sprengelmeyer, A., et al. (2009). The cutest little baby face: a hormonal link to sensitivity to cuteness in infant faces. *Psych. Sci.* 20, 149–154. doi: 10.1111/j.1467-9280.2009.02272.x

Thornhill, R., and Moller, A. P. (1997). Developmental stability, disease and medicine. *Biol. Rev.* 72, 497–548. doi: 10.1017/S0006323197005082

Zajonc, R. B. (1968). Attitudinal effects of mere exposures. *J. Personal. Soc. Psych.* 9, 1–27. doi: 10.1037/h0025848

The relationship between level of autistic traits and local bias in the context of the McGurk effect

Yuta Ujiie [1,2]*, Tomohisa Asai [3] and Akio Wakabayashi [4]

[1] Information Processing and Computer Sciences, Graduate School of Advanced Integration Science, Chiba University, Chiba, Japan, [2] Japan Society for the Promotion of Science, Tokyo, Japan, [3] NTT Communication Science Laboratories, NTT Corporation, Kanagawa, Japan, [4] Faculty of Letters, Chiba University, Chiba, Japan

The McGurk effect is a well-known illustration that demonstrates the influence of visual information on hearing in the context of speech perception. Some studies have reported that individuals with autism spectrum disorder (ASD) display abnormal processing of audio-visual speech integration, while other studies showed contradictory results. Based on the dimensional model of ASD, we administered two analog studies to examine the link between level of autistic traits, as assessed by the Autism Spectrum Quotient (AQ), and the McGurk effect among a sample of university students. In the first experiment, we found that autistic traits correlated negatively with fused (McGurk) responses. Then, we manipulated presentation types of visual stimuli to examine whether the local bias toward visual speech cues modulated individual differences in the McGurk effect. The presentation included four types of visual images, comprising no image, mouth only, mouth and eyes, and full face. The results revealed that global facial information facilitates the influence of visual speech cues on McGurk stimuli. Moreover, individual differences between groups with low and high levels of autistic traits appeared when the full-face visual speech cue with an incongruent voice condition was presented. These results suggest that individual differences in the McGurk effect might be due to a weak ability to process global facial information in individuals with high levels of autistic traits.

Keywords: autism spectrum disorder, Autism Spectrum Quotient, the McGurk effect, local bias, individual differences

Introduction

Autism spectrum disorder (ASD) has been largely defined in terms of difficulties in social interaction and communication, patterns of repetitive behavior, and narrow interests (American Psychiatric Association, 1994, 2013). In earlier ASD research, the dysfunction of processing information relevant to social interaction was the main focus of investigations. It has been revealed that individuals with ASD show different patterns in face perception (e.g., Deruelle et al., 2004) and emotion recognition (e.g., Baron-Cohen et al., 2001a) compared to individuals with typical development (TD). In addition to dysfunction in processing visual stimuli, recent studies have shown that individuals with ASD exhibit atypical processing in audio-visual speech perception (Massaro and Bosseler, 2006; Smith and Bennetto, 2007), which indicates a limited ability to integrate visual and auditory information. This dysfunction is considered to lead to communication impairment in ASD because speech perception is one of the core functions of face-to-face communication.

In face-to-face communication with others, we realize what another person is saying through the processing of audio-visual speech information. An early study of audio-visual speech perception provided strong evidence that visual information improves the auditory speech percept (Sumby and Pollack, 1954). A classic example that demonstrates the interaction between hearing and vision in speech perception is the McGurk effect (McGurk and MacDonald, 1976). This effect may be experienced when the visual shape produced during speech of a phoneme (e.g., /ga/) is dubbed with a sound recording of a different phoneme (e.g., /ba/), which often causes a third, intermediate phoneme (e.g., /da/) to be perceived. Similarly, for the monosyllabic combination of the visual /ka/ and auditory /pa/, participants often reported hearing /ta/.

Abnormal processing of audio-visual speech integration in individuals with ASD has been reported in a number of studies (De Gelder et al., 1991; Williams et al., 2004; Iarocci et al., 2010; Taylor et al., 2010; Saalasti et al., 2011, 2012; Woynaroski et al., 2013; Stevenson et al., 2014). Iarocci et al. (2010) reported that children with ASD showed less visual influence and more auditory influence during bimodal speech perception than controls did, due to poor lip-reading ability, and this finding was supported by Williams et al. (2004). On the other hand, De Gelder et al. (1991) reported that children with ASD are less influenced by the auditory percepts from visual speech cues, although they did not differ in lip-reading ability from children with TD. Similar results have been reported in children with ASD (Stevenson et al., 2014) and in adults with ASD (Saalasti et al., 2011, 2012). Although, the results are mixed, it has often been reported that individuals with ASD exhibit a weak degree of visual influence on perceiving a voice during audio-visual speech perception.

Another study suggested that the reduced McGurk effect in individuals with ASD meant that there was a delay, rather than a deficit, in the development of audio-visual integration (Taylor et al., 2010). Taylor et al. (2010) showed that younger children with ASD exhibit delayed visual accuracy and audio-visual integration (the McGurk effect) compared to children with TD, but appeared to catch up with their TD peers in the older age ranges. In line with this, Keane et al. (2010) revealed no individual differences in the McGurk effect between adults with and without ASD. This was inconsistent with the results in some previous studies (Saalasti et al., 2012; Stevenson et al., 2014).

Neuropsychological data provide us with some advantages to identify whether individuals with ASD show weaker visual influence on perceiving a voice. In ASD, several studies have reported anatomical and functional abnormalities in the superior temporal sulcus (STS) (see, for a review, Zilbovicius et al., 2006; Redcay, 2008). The STS is critical for integration of auditory and visual speech information, and influences the likelihood of the McGurk effect occurring (Calvert et al., 2000; Nath and Beauchamp, 2011). Redcay (2008) argued that impairments in STS function might lead to abnormalities in speech perception in individuals with ASD. In the sample of individuals with TD, a functional magnetic resonance imaging (fMRI) study revealed a significant positive correlation between the likelihood of perceiving the McGurk effect and the amplitude of the

response in the left STS (Nath and Beauchamp, 2012). This means that individuals with a weak response in the left STS showed fewer instances of the McGurk effect when they observed audio-visual-incongruent stimuli. Another study using near-infrared spectroscopy (NIRS) reported a significantly negative correlation between level of autistic traits and regional cerebral blood volume in the left STS during face-to-face conversation among adults with TD (Suda et al., 2011). These studies led us to hypothesize that individuals with ASD (or a high level of autistic traits) might show fewer instances of the McGurk effect due to a weak response in the left STS.

The reason why previous results are mixed might be due to the heterogeneity of the clinical population. For instance, abnormalities in sensory inputs (hyper- or hypo-sensitivity), which is one of the core symptoms of ASD, have been found in more than 90% of individuals with ASD in at least one sensory domain (Tomchek and Dunn, 2007; Crane et al., 2009), but in which sensory domain the abnormality appears varies (e.g., visual: Simmons et al., 2009; auditory: Haesen et al., 2011; O'Connor, 2012; tactile: Foss-Feig et al., 2012). With regard to the McGurk effect, one study showed that a fusion response was correlated with the degree of auditory processing difficulty, as assessed by the Sensory Profile (Dunn and Westman, 1997) among individuals with ASD (Woynaroski et al., 2013). Saalasti et al. (2012) examined the distribution of the likelihood of the McGurk response occurring and showed that the difference between the control group and the clinical group with ASD was significant. Mixed results in previous studies might be due to the heterogeneity in the profile of hyper- or hypo-sensitivity, which is difficult to control in a clinical group.

An analog design is one approach used to study ASD symptoms among individuals with TD, by examining the relationship between level of autistic traits and performance on a cognitive or perceptual task. This approach is based on the dimensional model of ASD, which assumes that autistic traits are distributed on a continuum over clinical and general populations (Frith, 1991; Baron-Cohen, 1995). In order to assess the degree of autistic traits in any individual adult with a normal intelligence quotient, Baron-Cohen et al. (2001b) developed the Autism Spectrum Quotient (AQ). The AQ is a self-report questionnaire and is useful as a screening scale to not only distinguish between clinical and control groups but also measure the distribution of autistic traits within the general population. The validity and reliability of this screening scale have been confirmed in various countries (UK: Baron-Cohen et al., 2001b; the Netherlands: Hoekstra et al., 2008; Australia: Lau et al., 2013; and Japan: Wakabayashi et al., 2006b). Moreover, our pilot study (Ujiie and Wakabayashi, 2015) found that the overlap between level of autistic traits and the degree of hyper- or hypo-sensitivity, which was assessed by the Glasgow Sensory Questionnaire (Robertson and Simmons, 2012), was small among a general population. Because an analog design allows for control of the heterogeneity in the profile of hyper- or hypo-sensitivity, we adopted this design to examine the relationship between level of autistic traits and McGurk effects among a population with TD who were free from problems with sensory inputs.

The purpose of this study was to use an analog design based on the dimensional model of ASD to investigate the relationship between individual differences in the McGurk effect and autistic traits in the general population. First, we investigated whether autistic traits were correlated with a weaker visual influence on speech perception under a without-noise condition (Experiment 1). As the McGurk stimuli, we used the combination of auditory /pa/ and visual /ka/ stimuli, because this combination was likely to be perceived as a stronger illusion than other combinations (e.g., auditory /ba/ and visual /ga/ stimuli) in a Japanese sample (Sekiyama, 1994). There were three possible responses to the McGurk stimuli, which were as follows: audio response (/pa/ response), fused response (/ta/ response), and visual response (/ka/ response). In this study, we defined the rate of fused response, which was the frequency of the McGurk effect occurring, as the degree of visually captured percept when hearing and viewing the McGurk stimuli. We defined the rate of /pa/ response, which was the correct response to the audio-visual-incongruent stimuli, as the strength of visual influence on perceiving a voice. Thus, we hypothesized that the degree of autistic traits would correlate negatively with the rate of fused response and correlate positively with the rate of audio response in the context of the McGurk stimuli. In Experiment 2, we focused on the local bias toward visual speech cues in individuals with ASD, and investigated whether this bias underlies the link between level of autistic traits and the McGurk effect.

Individuals with ASD have been shown to tend to prefer local over global information when presented multiple information sources (Happe and Frith, 2006). Such cognitive specificity in individuals with ASD is called the local bias (Frith, 1991; Happe and Frith, 2006), and this bias has been mainly reported in relation to visuospatial tasks. For instance, individuals with ASD show better performance than individuals with TD on the Embedded Figures Test (EFT; Brosnan et al., 2012), in which one is required to find a local (embedded) target within a global context constructed of multiple figures. A similar result has been found in the Navon-type Global–Local Naming Task (Reed et al., 2011) and various perceptual and cognitive tasks (Happe and Frith, 2006).

Furthermore, the local bias has been found in face processing tasks (Joseph and Tanaka, 2003; Deruelle et al., 2004; Kätsyri et al., 2008). Deruelle et al. (2004) investigated whether children with ASD preferred to use local (high spatial frequency) rather than global (low spatial frequency) information during a face matching task. The results showed that children with ASD showed better performance when using local information than when they used global information. Kätsyri et al. (2008) found a similar result in the recognition of dynamic facial emotions among adults with ASD. In addition, some studies suggested a preference of gaze toward the mouth region when individuals with ASD perceived face stimuli (Klin et al., 2002; Joseph and Tanaka, 2003). Joseph and Tanaka (2003) revealed that individuals with ASD use more information from around the mouth region of face stimuli in face recognition tasks. Klin et al. (2002) showed that individuals with ASD tend to gaze at the mouth region during the viewing of conversation. These results indicate that individuals with ASD might prefer to use local information, particularly that from the mouth region, during

face processing. Individuals with high AQ scores who exhibit a local bias (Reed et al., 2011) might be less likely to experience the McGurk effect. To clarify this, we should examine whether face processing influences the occurrence of the McGurk effect, in general.

Whether face processing is needed for audio-visual speech perception has been discussed in the past. According to Bruce and Young's (1986) model, face perception has three important functions, comprising recognition of facial identity, facial expression, and facial speech. In line with this suggestion, some studies have demonstrated a relationship between face processing and speech perception. De Gelder et al. (1991) showed that face identification correlates positively with the influence of lip-reading on audio-visual-incongruent stimuli. Rosenblum et al. (2002) showed that face processing and speech perception share the same dynamic information. Some studies, however, suggested that a mouth-only presentation influenced voice perception and produced the McGurk effect as well as a whole-face presentation did (Rosenblum et al., 2000; Hietanen et al., 2001). Other studies showed the role of extraoral facial information in audio-visual speech perception (Thomas and Jordan, 2004; Jordan and Thomas, 2011). Jordan and Thomas (2011) revealed that occluded oral areas disrupted performance but that observers could use lip-reading and observe visual speech influences from extraoral areas. Thomas and Jordan (2004) showed that an extraoral movement during visual speech was effective to perceive visual speech cues and influenced audio-visual speech perception.

The role of holistic face processing on speech perception has been investigated by using methods to examine the effect of holistic processing on face perception. For instance, some studies investigated the face inversion effect, which is a phenomenon that causes difficulty in holistic face processing by showing an inverted face, in the context of audio-visual speech perception (Jordan and Bevan, 1997; Rosenblum et al., 2000; Eskelund et al., 2015). One study showed a robust effect (Eskelund et al., 2015) while another study found only a partial effect (Jordan and Bevan, 1997), which means that the face inversion effect depends on the stimulus. Rosenblum et al. (2000) examined the role of holistic facial information in audio-visual speech perception, compared to full-face Thatcher-type speech stimuli and inverted mouth-alone speech stimuli. In their study, the full-face Thatcher-type stimuli were created by combining an upright face with an inverted mouth visual speech cue. They reported that only the full-face Thatcher-type stimuli disrupted voice perception for the audio-visual-incongruent condition (the combination of auditory /va/ and visual /ba/). This result, which is called the McThatcher effect, was replicated in Eskelund et al. (2015). Hietanen et al. (2001) investigated the effect of facial configuration context on the McGurk effect. They manipulated the location of facial features in visual stimuli, using either a natural or scrambled location. In their results, only an asymmetrically scrambled face disrupted the likelihood of the McGurk effect, but this effect depended on the stimulus. They concluded that facial configuration information can be used in audio-visual speech perception, although this information is not necessary. These studies indicate that processing of global

(holistic) visual speech cues might influence the occurrence of the McGurk effect.

In summary, based on the dimensional model of ASD, we administered two experiments to examine the relationship between level of autistic traits and McGurk effects in university students. In Experiment 1, we investigated the correlation between level of autistic traits and individual differences in audio-visual speech perception. We hypothesized that individuals with high levels of autistic traits would show a reduced likelihood of the McGurk effect occurring than would individuals with low levels of autistic traits. In Experiment 2, we examined whether the local bias toward visual speech cues modulates individual differences in the McGurk effect, by manipulating presentation types of visual stimuli (parts of the face or the full face). With regard to the likelihood of the McGurk effect occurring, we hypothesized that the visual influence on voice perception would be greater in the full-face presentation condition than in the partial-face presentation condition. In addition, we hypothesized that individual differences in the McGurk effect would appear when the full-face visual speech cue with an incongruent voice condition was presented, because of the local bias toward visual speech cues in individuals with high levels of autistic traits. The outcomes from these experiments will allow us to understand the effect of face processing on speech perception, and how audio-visual speech integration in individuals with ASD functions from an analog perspective.

Experiment 1

In Experiment 1, we investigated the correlation between AQ scores and level of accuracy for perceiving audio-visual stimuli and auditory stimuli, and assessed the likelihood of the McGurk effect (the rate of /ta/ response) occurring among non-ASD university students. For audio-visual-incongruent stimuli, we hypothesized that, because of the weak visual influence of perceiving a voice, the AQ scores would correlate negatively with the likelihood of the McGurk effect occurring, and positively with the rate of the /pa/ response being reported.

Methods
Participants
Participants were 46 university students (12 males and 34 females) who were recruited from an introductory psychology class at Chiba University. The mean age of the participants was 19.4 years ($SD = 3.56$). All participants were native speakers of Japanese and reported normal hearing and vision. They provided written informed consent in the class, and took part voluntarily in this experiment. After the experiment, they received an oral debriefing.

Stimuli
Japanese version of the AQ
The AQ was normalized for use in the Japanese population by Wakabayashi et al. (2006b). The AQ contains 50 items for assessing the following five domains: social skill, attention switching, attention to detail, communication, and imagination. Participants rate each item on a 4-point response scale from

"agree" to "disagree." Each item is scored 0 or 1 point according to the scoring manner described in previous studies (Baron-Cohen et al., 2001b; Wakabayashi et al., 2006b), so that total scores on the AQ range from 0 to 50.

Audio-visual task
The audio-visual stimuli were created from simultaneous audio and video recordings of six Japanese speakers' utterances (three female). The visual stimuli were speakers' faces recorded using a digital video camera (GZ-EX370, JVC KENWOOD). The audio stimuli were the utterances (/pa/, /ta/, or /ka/) collected using a dynamic microphone (MD42, SENNHEISER). The video clip (720×480 pixels, 29.97 frames/s) and the speech sound (digitized at 48,000 Hz, with 16-bit quantization resolution) were combined and synchronized using Adobe Premiere Pro CS6. The mean duration of the audio-visual stimuli was 1.2 s.

There were three stimulus conditions, comprising audio-only (e.g., auditory /pa/), audio-visual congruent (e.g., auditory /pa/, visual /pa/), and audio-visual incongruent (e.g., auditory /pa/, visual /ka/). Each condition included 18 trials per block.

In the audio-only condition, the audio stimuli (/pa/, /ta/, or /ka/) were presented without the visual stimuli. In the audio-visual-congruent condition, all three combinations of the audio and visual stimuli were presented. In the audio-visual-incongruent condition, the combination of the auditory /ka/ stimulus dubbed with visual /pa/ was excluded, because the percept (e.g., /pka/) caused by this combination stimulus is not a Japanese native syllable. Therefore, the voice (/pa/) and video (/ka/) combined stimuli were presented three times per block to make the same number of audio-visual-congruent trials.

Apparatus
The experiment was conducted using Hot Soup Processor Version 3.3 (Onion software). The video signals were presented on a 19-inch cathode ray tube (CRT) monitor (E193FPp, Dell), and the speech sound was presented through a headphone (MDR-Z500, Sony) at approximately a 65 dB sound pressure level, adjusted using a mixing console (MW8CX, Yamaha).

Procedure
Participants were seated at a distance of approximately 50 cm from the CRT monitor, wearing the headphone. Participants were instructed to report what they heard (/pa/, /ta/, or /ka/) by a key press. In each trial, a fixation point was displayed for 1000 ms at the center of the CRT monitor, followed by either the congruent or the incongruent stimulus. Then, a blank display was presented until participants responded.

The first block included 18 congruent stimuli and 18 incongruent stimuli. The second block included 18 auditory stimuli. All participants completed both blocks after undergoing six practice trials each. The order of trials was randomized for each block. After all of the tasks were finished, participants completed the questionnaire.

Data Analysis
Statistical analysis was conducted using R version 2.15.2 for Windows (R Foundation for Statistical Computing, Vienna,

Austria). To examine the effect of stimuli conditions, we analyzed the mean accuracies for each condition using a One-Way analysis of variance (ANOVA), with conditions as a within-participants factor. The likelihood of the McGurk effect occurring was analyzed in the audio-visual-incongruent condition using a chi-square test. The relationship between task performance and AQ scores was analyzed using Pearson correlation coefficients. In addition, group differences between a high-AQ group and a low-AQ group were analyzed using independent samples t-tests for each condition.

Results

Table 1 shows mean accuracies for the audio-visual-congruent condition and the audio-only condition, and the mean response rate for the audio-visual-incongruent condition. A One-Way ANOVA with conditions as a within-participants factor revealed a main effect of conditions, $F_{(2, 90)} = 197.215$, $p < 0.01$, partial $\eta^2 = 0.81$. Multiple comparisons (Holm method) showed that accuracies for correctly perceiving the voice in the audio-only condition ($M = 97.3\%$) and the audio-visual-congruent condition ($M = 98.1\%$) were higher than in the audio-visual-incongruent condition ($M = 34.9\%$; $p < 0.05$). However, the accuracy in the audio-only condition did not differ from that in the audio-visual-congruent condition. In the audio-visual-incongruent condition, the rate of the /ta/ response ($M = 61.1\%$) was higher than the rate of the /pa/ response ($M = 34.9\%$; $\chi^2 = 7.66$, $p < 0.01$) and the /ka/ response ($M = 4.0\%$; $\chi^2 = 20.33$, $p < 0.01$), which confirmed the occurrence of the McGurk effect.

The AQ scores ranged from 10 to 37 ($M = 20.8$, $SD = 5.42$). The distribution of AQ scores in this sample was slightly higher than that reported in the original publication of the AQ (Baron-Cohen et al., 2001b). To examine the relationship between task performances and the AQ scores, we calculated Pearson correlation coefficients (**Table 1**). No significant correlation was observed for the audio-visual-congruent condition or the audio-only condition. For the audio-visual-incongruent condition, the AQ was significantly positively correlated with the /pa/ response,

$r_{(46)} = 0.29$, $p < 0.05$, and significantly negatively correlated with the /ta/ response, $r_{(46)} = -0.32$, $p < 0.05$. These correlations suggest that individuals with low AQ scores show a more visually captured response and less audio response than individuals with high AQ scores do.

Next, we examined group differences between individuals with high AQ scores and those with low AQ scores in each condition. From among the participants, we picked eight with scores of 15 or under (mean AQ – 1 SD), and another eight with scores of 26 or over (mean AQ + 1 SD). We regarded the former as the low-AQ group (4 males and 4 females, mean AQ = 13.5) and the latter as the high-AQ group (3 males and 5 females, mean AQ = 29.3). A between-groups t-test showed a significant difference in the AQ scores, $t_{(14)} = 11.26$, $p < 0.01$, $r = 0.95$. Similarly, we conducted independent samples t-tests for each condition. No significant difference was found in the audio-only condition, $t_{(14)} = 1.17$, ns, $r = 0.29$, or in the audio-visual-congruent condition, $t_{(14)} = 0.43$, ns, $r = 0.12$ (see Supplementary Material). For the audio-visual-incongruent condition (see **Figure 1**), the rate of the /ta/ response was higher in the low-AQ group ($M = 65.3\%$) than in the high-AQ group ($M = 43.1\%$). This difference was marginally significant, $t_{(14)} = 1.79$, $p < 0.10$, $r = 0.43$; however, the rate of the /pa/ response was not significantly different, $t_{(14)} = 1.62$, $p = 0.12$, $r = 0.40$. These results indicate that individuals with high AQ scores show weaker visually captured responses than individuals with low AQ scores do, although accuracies for perceiving voice and audio-visual speech did not differ.

Discussion

In this experiment, we investigated the relationship between audio-visual speech integration and the level of autistic traits in healthy students. We found that the level of autistic traits correlated negatively with the rate of fused response and positively with the rate of audio response in the audio-visual-incongruent condition. Moreover, the results revealed that individuals with high AQ scores showed a weaker fused response

TABLE 1 | Mean and standard deviations (SD) of response rate in all conditions.

	Mean	SD	Correlations with the AQ scores
AUDIO-ONLY CONDITION			
Correct response (/pa/, /ta/, /ka/)	0.97	0.03	0.00
AUDIO-VISUAL-CONGRUENT CONDITION			
Correct response (/pa/, /ta/, /ka/)	0.99	0.03	−0.10
AUDIO-ONLY-INCONGRUENT CONDITION			
Audio response (/pa/)	0.35	0.29	0.29*
Fused response (/ta/)	0.61	0.28	−0.30*
Visual response (/ka/)	0.04	0.08	−0.19

Also displayed are correlations between the AQ scores and response rates for experiment 1. N = 46. AQ, Autism-spectrum Quotient. Correct response (/pa/, /ta/, /ka/): Mean correct responses for all stimuli in the audio-visual-congruent condition or the audio-only condition. *p < 0.05, **p < 0.01.
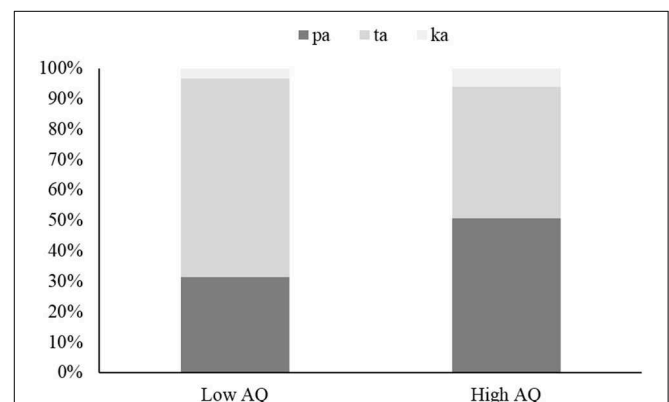

FIGURE 1 | The response rate for each audio-visual-incongruent stimulus in the low-AQ group and the high-AQ group. Possible responses to the stimuli were audio response (/pa/ response), fused response (/ta/ response), and visual response (/ka/ response).

than individuals with low AQ scores did, although there was no significant difference in the audio response rate. On the other hand, neither significant correlations nor group differences were found in the audio-visual-congruent condition and audio-only condition. These results indicate that individuals with higher levels of autistic traits tended to show a weaker visual influence on perceiving a voice when processing audio-visual-incongruent speech information.

Several studies reported that individuals with ASD showed a weaker visual influence only when McGurk stimuli are presented (e.g., De Gelder et al., 1991). This study replicated those results in a sample of university students. As we hypothesized, our results indicate that the weakness of visual influence on audio-visual speech perception exists along the distribution of AQ in the general population. This finding might support the dimensional model of ASD, because individuals with high AQ scores in this study and individuals with ASD in previous studies (e.g., De Gelder et al., 1991) showed a similar tendency when processing audio-visual-incongruent speech.

However, in this experimental task, it was not clear what factor led to a weaker visual influence on audio-visual speech perception. One possibility is that the local bias toward visual speech cues reflected individual differences in the McGurk effect. There was a local bias effect of cognitive specificity on individuals with ASD, meaning that there is a bias toward processing local information in preference to global information (Frith, 1991; Happe and Frith, 2006). In addition to the results for visuo-spatial tasks (Reed et al., 2011; Brosnan et al., 2012), recent studies have reported that individuals with ASD have a preference for feature-based processing of face stimuli (Joseph and Tanaka, 2003; Deruelle et al., 2004), and for focusing on the local (mouth) region during the viewing of conversation videos (Klin et al., 2002). Some studies have suggested that the influence of visual speech cues in processing audio-visual-incongruent stimuli is related to the processing of faces (De Gelder et al., 1991), especially in the global (holistic) facial context (Rosenblum et al., 2000). Thus, if global facial context enhances the influence of visual speech cues on perceiving a voice, individual differences in the McGurk effect between individuals with high AQ scores and those with low AQ scores might be due to a weak ability to process global facial information in McGurk stimuli. To confirm this, we conducted Experiment 2.

Experiment 2

In Experiment 2, we manipulated presentation types of visual stimuli to examine whether the local bias affected individual differences in the McGurk effect. We set two stimulus conditions, i.e., the audio-visual-congruent condition and the audio-visual-incongruent condition. For the audio-visual-incongruent condition, we defined the rate of fused responses as the frequency of visually captured percept, while we defined the rate of /pa/ responses as the strength of visual influence to perceiving voice, as in Experiment 1.

In addition, we created the following four types of visual stimuli: no image (audio-only), mouth-only, eyes and mouth, and full face. Only the full-face stimuli included global

facial information of visual speech cues. For the audio-visual-incongruent condition, we hypothesized that audio response would be observed less frequently in the full-face presentation than in the other stimuli conditions if the processing of global visual speech cues is related to the degree of visual influence on perceiving a voice. Moreover, we also hypothesized that the differences between individuals with high AQ scores and those with low AQ scores would diminish (or become small) when a voice and an incongruent visual speech cue without global visual information, such as only the mouth region, was presented.

Methods
Participants
Another 50 healthy students (12 males and 38 females), who were recruited from an introductory psychology class at Chiba University, participated in the experiment. The mean age of the participants was 19.4 years ($SD = 3.41$). All participants were native speakers of Japanese and reported normal hearing and normal (or corrected) vision. They provided written informed consent in the class and took part in the study voluntarily. After the experiment, they received an oral debriefing.

Stimuli
We used the same stimuli as in Experiment 1, comprising six (3 females) Japanese speakers' utterances of three syllables (/pa/, /ta/, or /ka/). There were two audio-visual stimulus conditions, i.e., the audio-visual congruent and audio-visual incongruent. The audio-visual stimuli consisted of a congruent auditory /pa/–visual /pa/, a congruent auditory /ta/–visual /ta/, a congruent auditory /ka/–visual /ka/, and an incongruent auditory /pa/–visual /ka/.

The four types of presentations of visual stimuli—no image (audio-only), mouth-only, eyes and mouth, and full face— (examples of the visual stimuli are shown in **Figure 2**) were created for each condition by using Adobe Premiere Pro CS6 to crop eye regions and the mouth region from visual images. The eye region included the region from the inner corner of the eyes to the outer corner. The mouth region included a range of motion of the upper lip and lower lip. This task consisted of 72 congruent stimuli and 24 incongruent stimuli per block.

Following the experimental tasks, we used the Japanese version of the AQ (Baron-Cohen et al., 2001b; Wakabayashi et al., 2006b) to measure the level of autistic traits in the participants.

Procedure
The experiment was carried out individually, using the same apparatus as in Experiment 1. Participants were seated at a distance of approximately 50 cm from the 19-in CRT monitor, wearing the headphone. They were instructed to report what they heard (/pa/, /ta/, or /ka/) by pressing buttons on a keyboard. In each trial, a fixation point was displayed for 1000 ms at the center of the CRT monitor. After that, either the congruent or the incongruent stimulus was presented, followed by a blank display presented until participants responded. All participants completed the two blocks of the main session after undergoing the 10-trial practice session. The order of trials was randomized
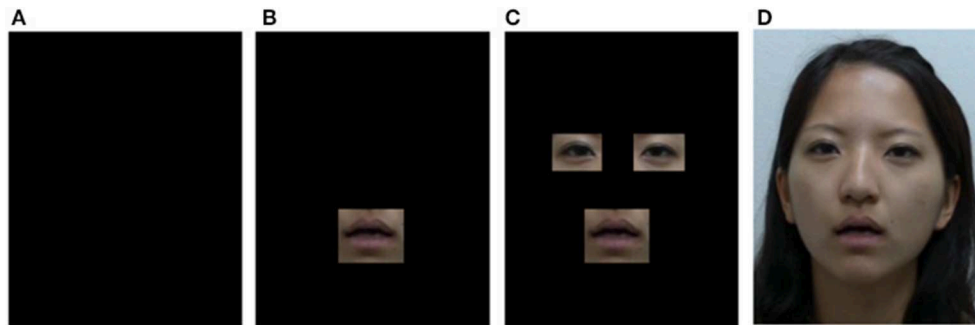
FIGURE 2 | Examples of the four types of visual stimuli used in Experiment 2. (A) No image (audio only). **(B)** Mouth-only presentation. **(C)** Eyes and mouth presentation. **(D)** Full-face image presentation. All of these images were presented with a congruent or incongruent voice in the experiment.

for each block. After the tasks were finished, participants completed the questionnaire.

Data Analysis

Statistical analysis was conducted using R version 2.15.2 for Windows (R Foundation for Statistical Computing, Vienna, Austria). In order to examine the effect of visual presentation type and stimulus condition, rates of correct (audio) responses were analyzed using a Two-Way ANOVA with visual presentation types and stimulus conditions as within-participant factors. As in Experiment 1, the relationship between the AQ scores and task performance was analyzed using Pearson correlation coefficients for each stimulus condition. In addition, group differences between high- and low-AQ groups were analyzed using a mixed ANOVA with visual presentation as a within-participant factor and groups as a between-participants factor for rates of correct responses in the audio-visual-congruent condition and of audio responses in the audio-visual-incongruent condition.

Results

Figure 3 summarizes mean accuracies for the congruent and incongruent stimuli conditions in all types of visual presentation. A Two-Way ANOVA with visual presentation types and stimulus conditions as within-participant factors revealed main effects of stimulus conditions, $F_{(1, 49)} = 201.10$, $p < 0.01$, partial $\eta^2 = 0.80$, and visual presentation, $F_{(3, 149)} = 110.72$, $p < 0.01$, partial $\eta^2 = 0.69$, and a significant one-way interaction, $F_{(1, 49)} = 132.10$, $p < 0.01$, partial $\eta^2 = 0.73$. Multiple comparisons (Holm method) showed that the accuracy for no image, which presented only the auditory stimulus, was lower than for the other types of presentation in the audio-visual-congruent condition ($p < 0.05$). On the other hand, in the audio-visual-incongruent condition, the rate of audio (correct) responses for no image was higher than for the other types of presentation ($p < 0.05$), and the audio response for the full-face presentation was lower than that for either the mouth or mouth and eyes presentation ($p < 0.05$). These results suggested that any type of visual speech cue improved the perception accuracy for audio-visual-congruent stimuli. In addition, as we expected, the influence of a visual speech cue on perceiving a voice was strongest for the presentation of full-face speech with the incongruent voice.
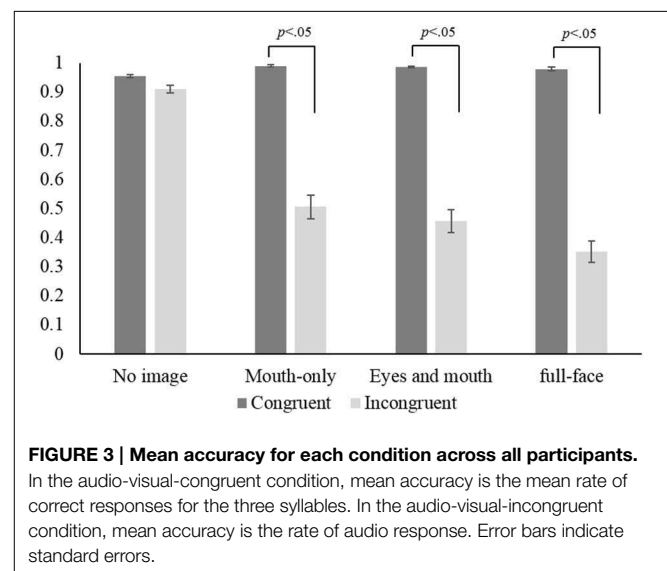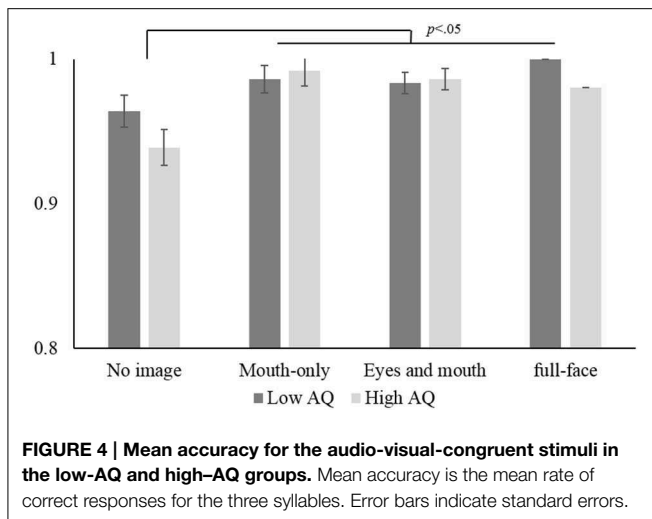

FIGURE 3 | Mean accuracy for each condition across all participants. In the audio-visual-congruent condition, mean accuracy is the mean rate of correct responses for the three syllables. In the audio-visual-incongruent condition, mean accuracy is the rate of audio response. Error bars indicate standard errors.

Next, we examined the relationship between the AQ scores and the effect of visual presentation on speech perception. The scores on the AQ ranged from 10 to 43 with a mean of 21.2 ($SD = 6.07$). In order to examine the relationship between task performance and AQ scores, we calculated Pearson correlation coefficients (see Supplementary Material). No significant correlation was observed in the audio-visual-congruent condition. In the audio-visual-incongruent condition, AQ scores were significantly positively correlated with the audio (/pa/) response, $r_{(50)} = 0.31$, $p < 0.05$, and negatively correlated with the fused (/ta/) response, $r_{(50)} = -0.31$, $p < 0.05$, but only for the full-face presentation. These correlations replicated the results in Experiment 1, which indicates that individuals with high AQ scores showed less of a visually captured response than individuals with low AQ scores did, although this only occurred in the full-face incongruent speech condition.

Then, we examined group differences between the high-AQ and low-AQ groups in each condition. From among the participants, we picked 10 with scores of 27 or over as the former (2 males and 8 females, mean AQ score = 30.1), and another 10

FIGURE 4 | Mean accuracy for the audio-visual-congruent stimuli in the low-AQ and high–AQ groups. Mean accuracy is the mean rate of correct responses for the three syllables. Error bars indicate standard errors.

FIGURE 5 | Mean audio response to the audio-visual-incongruent stimuli in the low-AQ and high–AQ groups. Error bars indicate standard errors.

with scores of 16 or under as the latter (5 males and 5 females, mean AQ score = 13.7). A between-groups t-test showed a significant difference in the AQ scores, $t_{(18)} = 10.28$, $p < 0.01$, $r = 0.92$. In the congruent condition (see **Figure 4**), a mixed ANOVA revealed that the main effect of visual presentation was significant, $F_{(3, 54)} = 11.42$, $p < 0.01$, partial $\eta^2 = 0.36$, but that the main effect of groups and the one-way interaction were not.

For audio (correct) responses in the incongruent condition (see **Figure 5**), although no significant main effect of groups was found, there was a significant main effect of visual presentation, $F_{(3, 54)} = 63.83$, $p < 0.01$, partial $\eta^2 = 0.75$, and a significant one-way interaction, $F_{(3, 54)} = 3.16$, $p < 0.05$, partial $\eta^2 = 0.15$. This interaction revealed that the simple main effect of groups was significant only in the full-face presentation condition, $F_{(1, 18)} = 5.37$, $p < 0.01$, partial $\eta^2 = 0.23$, so that the individual differences between the high- and low-AQ groups appeared only in the full-face presentation condition. We also found that the effect size of visual presentation was slightly smaller in the high-AQ group, $F_{(3, 54)} = 19.88$, $p < 0.01$, partial $\eta^2 = 0.72$, than in the low-AQ group, $F_{(3, 54)} = 47.11$, $p < 0.01$, partial $\eta^2 = 0.52$. Similar results were found for fused responses in the incongruent condition, i.e., a main effect of visual presentation, $F_{(3, 54)} = 67.49$, $p < 0.01$, partial $\eta^2 = 0.76$, and a significant one-way interaction, $F_{(3, 54)} = 3.14$, $p < 0.05$, partial $\eta^2 = 0.15$. These results indicate that the effect of global facial information was greater in the low-AQ group than in the high-AQ group, although this effect was found in both groups.

Discussion

In Experiment 2, we aimed to investigate the effect of global facial information on audio-visual speech perception, and its relationship with level of autistic traits. With regard to the former purpose, we hypothesized that audio responses would be observed less frequently for the full-face image of a visual speech cue than for the only-mouth image of a visual speech cue in the incongruent condition. As we expected, our results revealed that the rate of audio responses was lower for the full-face image of a visual speech cue than for the other three types of
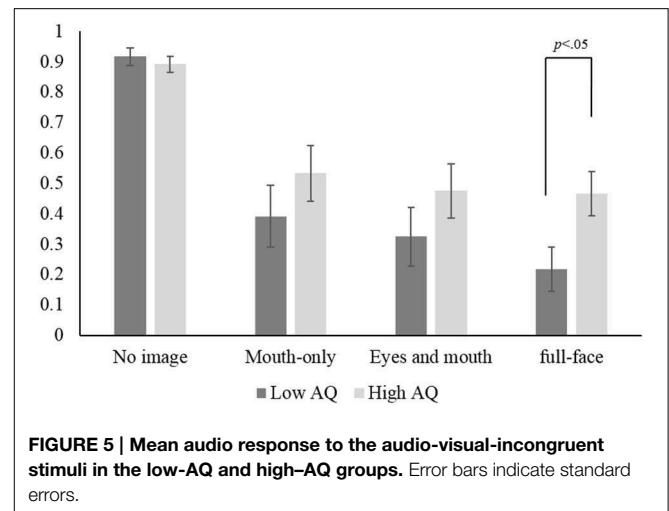
visual speech cue in the audio-visual-incongruent condition. This indicates that global facial information enhances the influence of a visual cue on perceiving a voice. Unlike previous results (e.g., Rosenblum et al., 2000; Hietanen et al., 2001), the difference in visual influence between the full-face presentation and the only-mouth presentation was robust in our study. Our result directly supports the assumption that the processing of global facial information (extraoral region) might be used for audio-visual speech integration (e.g., Thomas and Jordan, 2004; Eskelund et al., 2015).

With regard to the relationship with autistic traits, our results showed that the individual differences between the high-and low-AQ groups appeared only when a full-face image of a visual speech cue with an incongruent voice was presented. Such a group difference was not found in accuracies for the audio-visual-congruent stimuli. Furthermore, the effect of global facial information in the McGurk effect was small in the high-AQ group. This indicates that the local bias on face processing might play a role in audio-visual speech perception, as well as in the recognition of facial identity (Joseph and Tanaka, 2003; Deruelle et al., 2004) and of facial expression (Kätsyri et al., 2008). These results suggest that the visual influence on perceiving voice was weaker in individuals with high AQ scores than in those with low AQ scores because of the weakness of processing global facial information in the McGurk effect.

General Discussion

Implications for the Dimensional Model of Autism Spectrum Disorder

In two experiments, we examined the link between level of autistic traits and individual differences in audio-visual speech perception. The results demonstrated that level of autistic traits did not correlate with the accuracy for perceiving audio-visual-congruent speech, regardless of the visual speech presentation condition. Moreover, we did not find a correlation between level of autistic traits and the accuracy for perceiving auditory speech,

although individual differences in the audio-visual-incongruent condition, in which the McGurk effect was observed, were related to the degree of autistic traits in the general population. In the audio-visual-incongruent condition, individuals with high AQ scores showed fewer occurrences of the McGurk effect than individuals with low AQ scores did. These results indicate that autistic traits only correlated with the strength of visual influence on perceiving a voice in the audio-visual-incongruent condition.

Our findings have important implications for the dimensional model of ASD, especially for analog studies investigating symptoms of ASD in the general population. With regard to the influence of visual speech cues on perceiving a voice, our results are consistent with those of several previous studies on ASD (De Gelder et al., 1991; Williams et al., 2004; Saalasti et al., 2011; Stevenson et al., 2014). Such atypical processing in individuals with high AQ scores has been reported in the context of perceptual learning (Reed et al., 2011), perspective-taking (Brunye et al., 2012), and lexical effects on speech perception (Stewart and Ota, 2008). As with these studies, our results also support the dimensional model of ASD.

We considered that previous results were mixed due to the heterogeneity of a clinical ASD population in the profile of hyper- and hypo-sensitivity (Woynaroski et al., 2013). To eliminate this factor, we adopted an analog design and used a sample of individuals with TD, who were free from problems with sensory inputs. As we expected, we found significant relationships between level of autistic traits and individual differences in audio-visual speech perception. Some studies showed that it was possible to control the influence of other factors, such as the Big Five personality traits (Wakabayashi et al., 2006a), schizotypal personality (Wakabayashi et al., 2012), and degree of hyper- or hypo-sensitivity (Ujiie and Wakabayashi, 2015). These indicate that an analog design might be an effective approach in the investigation of ASD symptoms to control factors other than the degree of autistic traits.

The Role of Global Facial Information in the Occurrence of the McGurk Effect

Our results showed a more robust effect of global facial information in the occurrence of the McGurk effect, as compared to previous studies (e.g., Rosenblum et al., 2000; Hietanen et al., 2001). This indicates that extraoral region of visual speech cues might be used for audio-visual speech integration (Thomas and Jordan, 2004). Our results, however, could not reveal whether global facial information is critical for audio-visual speech perception. One previous study (Jordan and Thomas, 2011) stated that the mouth region of a visual speech cue is important for audio-visual speech perception, which is something we also found in this study, although the extraoral region could also be used. In this study, global facial information did not have a strong effect on audio-visual-congruent speech perception. This means that the accuracy when hearing a voice increased when any type of visual speech cue was presented with a congruent voice, compared to when only a voice was presented. Moreover, in the incongruent condition, the influence of a visual speech sound appeared even when a voice with an incongruent visual speech cue of only a mouth was exhibited. These findings indicate that

information provided by the mouth region is more critical for audio-visual speech perception, than is global facial information.

An issue in this study is that we did not consider the unnaturalness of the stimulation presentation. In this study, a mouth image included a range of motion of the upper lip and lower lip. This image of a visual speech cue was either presented on black background or presented along with other facial parts (eyes or full face). Jordan and Thomas (2011) pointed out that a display that does not obscure all of the face except for the mouth was unnatural. Therefore, rather than global face processing, this unnaturalness might have been what influenced audio-visual speech perception. Another issue is that we used only one combination of visual and audio syllables in the incongruent condition. Previous studies have shown that the effect of global face processing varied with the stimulus, such as a different talker or a different combination of syllables (Jordan and Bevan, 1997; Rosenblum et al., 2000). Nevertheless, the number of talkers in our stimuli (six talkers) was relatively larger than that used in previous studies (Jordan and Bevan, 1997; Rosenblum et al., 2000), as was our sample size.

The Relationship between Level of Autistic Traits and Local Bias in the McGurk Effect

With regard to the local bias exhibited by individuals with ASD, our results suggested a link between level of autistic traits and a weak ability to process global facial information in McGurk stimuli. In our results, the effect of global facial information in the McGurk stimuli was found to be smaller in individuals with low AQ scores than in individuals with high AQ scores, who show less likelihood of the McGurk effect occurring. This could be interpreted as indicating that individuals with high AQ scores show a local bias toward a visual speech cue and that their weak ability to process global facial information leads to individual differences in the McGurk effect. On the other hand, it is possible that other factors might have influenced our results, such as the atypical processing of global motion (Koldewyn et al., 2010), of visual attention (Zhao et al., 2013), or of gaze behavior (Klin et al., 2002).

Previous results provide us with some advantage to understand the influence of gaze behavior in our study. It has been shown that individuals with ASD exhibit atypical gaze behavior toward faces when they observe face stimuli (see, for a review, Senju and Johnson, 2009). Klin et al. (2002) indicated that individuals with ASD tend to fixate more on the mouth region when a dynamic face is presented, while individuals without ASD tend to fixate more on the region of the eyes. However, Saalasti et al. (2012) reported that no differences in gaze behavior between adults with ASD and controls that could have accounted for the individual differences in the McGurk effect. Moreover, Paré et al. (2003) showed that gaze fixations within the talker's face, which meant that gaze was fixed on the talker's mouth or on the talker's eyes, did not influence the likelihood of the McGurk effect occurring in adults with TD. Thus, it could be considered that, even if gaze behavior during trials differed between the high-AQ and low-AQ groups in Experiment 2, such

differences would not have substantially influenced the results of this study.

As another limitation in this study, it was unclear whether the individual differences in lip-reading were related to individual differences in the McGurk effect, because we did not use visual-only stimuli in this experiment. Some studies have reported that individuals with ASD experience a deficit in perceiving audio-visual speech because of their poor ability to lip-read (Williams et al., 2004; Woynaroski et al., 2013). Therefore, if individuals with high AQ scores have difficulties in lip-reading, individual differences in the McGurk effect might be caused by poor lip-reading ability, rather than by a local bias toward a visual speech cue. Nevertheless, the results of Experiment 2 showed a significant main effect of visual presentation in the congruent condition for both the high-AQ and low-AQ groups. In other words, when any type of congruent visual speech cue was exhibited, improved accuracy for perceiving a voice was found regardless of level of AQ. If the high-AQ group in this study had difficulties in lip-reading, such improvement would not have been found in that group. In order to clarify the role of a local bias toward a visual speech cue during audio-visual speech perception, these factors should be investigated directly in further studies.

Conclusion

In conclusion, level of the autistic traits in the general population was found to correlate negatively with visually influenced percepts with the McGurk stimuli. This is the first report of such a correlation. Moreover, individuals with high levels of autistic traits showed a weak ability to process global facial information during the McGurk stimuli.

Acknowledgments

Supplementary Material

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpsyg.2015.00891

References

American Psychiatric Association. (1994). *Diagnostic and Statistical Manual of Mental Disorders, 4th Edn.* Washington, DC: American Psychiatric Association.

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders, 5th Edn.* Washington, DC: American Psychiatric Association.

Baron-Cohen, S. (1995). *Mindblindness: An Essay on Autism and Theory of Mind.* Boston, MA: MIT Press/Bradford Books.

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., and Plumb, I. (2001a). The 'Reading the Mind in the eyes' test revised version: a study with normal adults, and adults with Asperger Syndrome or High-Functioning Autism. *J. Child Psychol. Psychiatry* 42, 241–252. doi: 10.1111/1469-7610.00715

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., and Clubley, E. (2001b). The Autism-Spectrum Quotient (AQ): evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *J. Autism Dev. Disord.* 31, 5–17. doi: 10.1023/A:1005653411471

Brosnan, M. J., Gwilliam, L. R., and Walker, I. (2012). Brief report: the relationship between visual acuity, the embedded figures test and systemizing in autism spectrum disorders. *J. Autism Dev. Disord.* 42, 2491–2497. doi: 10.1007/s10803-012-1505-0

Bruce, V., and Young, A. (1986). Understanding face recognition. *Br. J. Psychol.* 77, 305–327. doi: 10.1111/j.2044-8295.1986.tb02199.x

Brunye, T. T., Ditman, T., Giles, G. E., Mahoney, C. R., Kessler, K., and Taylor, H. A. (2012). Gender and autistic personality traits predict perspective-taking ability in typical adults. *J. Pers. Individ. Dif.* 52, 84–88. doi: 10.1016/j.paid.2011.09.004

Calvert, G. A., Campbell, R., and Brammer, M. J. (2000). Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. *Curr. Biol.* 10, 649–657. doi: 10.1016/S0960-9822(00)00513-3

Crane, L., Goddard, L., and Pring, L. (2009). Sensory processing in adults with autism spectrum disorders. *Autism* 13, 215–228. doi: 10.1177/1362361309103794

De Gelder, B., Vroomen, J., and van der Heide, L. (1991). Face recognition and lip-reading in autism. *Eur. J. Cogn. Psychol.* 3, 69–86. doi: 10.1080/09541449108406220

Deruelle, C., Rondan, C., Gepner, B., and Tardif, C. (2004). Spatial frequency and face processing in children with autism and Asperger syndrome. *J. Autism Dev. Disord.* 34, 199–210. doi: 10.1023/B:JADD.0000022610.09668.4c

Dunn, W., and Westman, K. (1997). The sensory profile: the performance of a national sample of children without disabilities. *Am. J. Occup. Ther.* 51, 25–34. doi: 10.5014/ajot.51.1.25

Eskelund, K., MacDonald, E. N., and Andersen, T. S. (2015). Face configuration affects speech perception: evidence from a McGurk mismatch negativity study. *Neuropsychologia* 66, 48–54. doi: 10.1016/j.neuropsychologia.2014.10.021

Foss-Feig, J. H., Heacock, J. L., and Cascio, C. J. (2012). Tactile responsiveness patterns and their association with core feature in autism spectrum disorders. *Res. Autism Spectr. Disord.* 6, 337–344. doi: 10.1016/j.rasd.2011.06.007

Frith, U. (1991). *Autism and Asperger's Syndrome.* Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511526770

Haesen, B., Boets, B., and Wagemans, J. (2011). A review of behavioral and electrophysiological studies on auditory processing and speech perception in autism spectrum disorders. *Res. Autism Spectr. Disord.* 5, 701–714. doi: 10.1016/j.rasd.2010.11.006

Happe, F., and Frith, U. (2006). The weak coherence account: detail-focused cognitive style in autism spectrum disorders. *J. Autism Dev. Disord.* 36, 5–25. doi: 10.1007/s10803-005-0039-0

Hietanen, J. K., Manninen, P., Sams, M., and Surakka, V. (2001). Does audiovisual speech perception use information about facial configuration? *Eur. J. Cogn. Psychol.* 13, 395–407. doi: 10.1080/09541440126006

Hoekstra, R. A., Bartels, M., Cath, D. C., and Boomsma, D. I. (2008). Factor structure, reliability and criterion validity of the Autism-Spectrum Quotient (AQ): a study in Dutch population and patient groups. *J. Autism Dev. Disord.* 38, 1555–1566. doi: 10.1007/s10803-008-0538-x

Iarocci, G., Rombough, A., Yager, J., Weeks, D., and Chua, R. (2010). Visual influences on speech perception in children with autism. *Autism* 14, 305–320. doi: 10.1177/1362361309353615

Jordan, T. R., and Bevan, K. (1997). Seeing and hearing rotated faces: influences official orientation on visual and audiovisual speech recognition. *J. Exp. Psychol. Hum. Percept. Perform.* 23, 388–403. doi: 10.1037/0096-1523.23.2.388

Jordan, T. R., and Thomas, S. M. (2011). When half a face is as good as a whole: effects of simple substantial occlusion on visual and audiovisual speech

perception. *Atten. Percept. Psychophys.* 73, 2270–2285. doi: 10.3758/s13414-011-0152-4

Joseph, R. M., and Tanaka, J. (2003). Holistic and part-based face recognition in children with autism. *J. Child Psychol. Psychiatry* 44, 529–542. doi: 10.1111/1469-7610.00142

Kätsyri, J., Saalasti, S., Tiippana, K., von Wendt, L., and Sams, M. (2008). Impaired recognition of facial emotions from low-spatial frequencies in Asperger syndrome. *Neuropsychologia* 46, 1888–1897. doi: 10.1016/j.neuropsychologia.2008.01.005

Keane, B. P., Rosenthal, O., Chun, N. H., and Shams, L. (2010). Audiovisual integration in high functioning adults with autism. *Res. Autism Spectr. Disord.* 4, 276–289. doi: 10.1016/j.rasd.2009.09.015

Klin, A., Jones, W., Schultz, R., Volkmar, R., and Cohen, D. (2002). Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism. *Arch. Gen. Psychiatry* 59, 809–816. doi: 10.1001/archpsyc.59.9.809

Koldewyn, K., Whitney, D., and Rivera, S. M. (2010). The psychophysics of visual motion and global form processing in autism. *Brain* 133, 599–610. doi: 10.1093/brain/awp272

Lau, W. Y. P., Kelly, A. B., and Peterson, C. C. (2013). Further evidence on the factorial structure of the autism spectrum quotient (AQ) for adults with and without a clinical diagnosis of autism. *J. Autism Dev. Disord.* 43, 2807–2815. doi: 10.1007/s10803-013-1827-6

Massaro, D. W., and Bosseler, A. (2006). Read my lips: the importance of the face in a computer-animated tutor for vocabulary learning by children with autism. *Autism* 10, 495–510. doi: 10.1177/1362361306066599

McGurk, H., and MacDonald, J. W. (1976). Hearing lips and seeing voices. *Nature* 264, 746–748. doi: 10.1038/264746a0

Nath, A. R., and Beauchamp, M. S. (2011). Dynamic changes in superior temporal sulcus connectivity during perception of noisy audiovisual speech. *J. Neurosci.* 31, 1704–1714. doi: 10.1523/JNEUROSCI.4853-10.2011

Nath, A. R., and Beauchamp, M. S. (2012). A neural basis for interindividual differences in the McGurk effect, a multisensory speech illusion. *Neuroimage* 59, 781–787. doi: 10.1016/j.neuroimage.2011.07.024

O'Connor, K. (2012). Auditory processing in autism spectrum disorder: a review. *Neurosci. Biobehav. Rev.* 36, 836–854. doi: 10.1016/j.neubiorev.2011.11.008

Paré, M., Richler, R., ten Hove, M., and Munhall, K. G. (2003). Gaze behavior in audiovisual speech perception: the influence of ocular fixations on the McGurk effect. *Percept. Psychophys.* 65, 553–567. doi: 10.3758/BF03194582

Redcay, E. (2008). The superior temporal sulcus performs a common function for social and speech perception: implications for the emergence of autism. *Neurosci. Biobehav. Rev.* 32, 123–142. doi: 10.1016/j.neubiorev.2007.06.004

Reed, P., Lowe, C., and Everett, R. (2011). Perceptual learning and perceptual search are altered in male university students with higher Autism Quotient scores. *Pers. Individ. Dif.* 51, 732–736. doi: 10.1016/j.paid.2011.06.016

Robertson, A. E., and Simmons, D. R. (2012). The relationship between sensory sensitivity and autistic traits in the general population. *J. Autism Dev. Disord.* 43, 775–784. doi: 10.1007/s10803-012-1608-7

Rosenblum, L. D., Yakel, D. A., Baseer, N., Panchal, A., Nodarse, B. C., and Niehus, R. P. (2002). Visual speech information for face recognition. *Percept. Psychophys.* 64, 220–229. doi: 10.3758/BF03195788

Rosenblum, L. D., Yakel, D. A., and Green, K. G. (2000). Face and mouth inversion affects on visual and audiovisual speech perception. *J. Exp. Psychol. Hum. Percept. Perform.* 26, 806–819. doi: 10.1037/0096-1523.26.2.806

Saalasti, S., Kätsyri, J., Tiippana, K., Laine-Hernandez, M., von Wendt, L., and Sams, M. (2012). Audiovisual speech perception and eye gaze behavior of adults with Asperger syndrome. *J. Autism Dev. Disord.* 42, 1606–1615. doi: 10.1007/s10803-011-1400-0

Saalasti, S., Tiippana, K., Kätsyri, J., and Sams, M. (2011). The effect of visual spatial attention on audiovisual speech perception in adults with Asperger syndrome. *Exp. Brain Res.* 213, 283–290. doi: 10.1007/s00221-011-2751-7

Sekiyama, K. (1994). Difference in auditory-visual speech perception between Japanese and Americans: McGurk effect as a function of incompatibility. *J. Acoust. Soc. Jpn.* 15, 143–158. doi: 10.1250/ast.15.143

Senju, A., and Johnson, M. H. (2009). Atypical eye contact in autism: models, mechanisms and development. *Neurosci. Biobehav. Rev.* 33, 1204–1214. doi: 10.1016/j.neubiorev.2009.06.001

Simmons, D. R., Robertson, A. E., McKay, L. S., Toal, E., McAleer, P., and Pollick, F. E. (2009). Vision in autism spectrum disorders. *Vis. Res.* 49, 2705–2739. doi: 10.1016/j.visres.2009.08.005

Smith, E. G., and Bennetto, L. (2007). Audiovisual speech integration and lipreading in autism. *J. Child Psychol. Psychiatry* 48, 813–821. doi: 10.1111/j.1469-7610.2007.01766.x

Stevenson, R. A., Siemann, J. K., Woynaroski, T. G., Schneider, B. C., Eberly, H. E., Camarata, S. M., et al. (2014). Brief report: arrested development of audiovisual speech perception in autism spectrum disorders. *J. Autism Dev. Disord.* 44, 1470–1477. doi: 10.1007/s10803-013-1992-7

Stewart, M. E., and Ota, M. (2008). Lexical effects on speech perception in individuals with autistic traits. *Cognition* 109, 157–162. doi: 10.1016/j.cognition.2008.07.010

Suda, M., Takei, Y., Aoyama, Y., Narita, K., Sakurai, N., Fukuda, M., et al. (2011). Autistic traits and brain activation during face-to-face conversations in typically developed adults. *PLoS ONE* 6:e20021. doi: 10.1371/journal.pone.0020021

Sumby, W. H., and Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *J. Acoust. Soc. Am.* 26, 212–215. doi: 10.1121/1.1907309

Taylor, N., Isaac, C., and Milne, E. (2010). A comparison of the development of audiovisual integration in children with autism spectrum disorders and typically developing children. *J. Autism Dev. Disord.* 40, 1403–1411. doi: 10.1007/s10803-010-1000-4

Thomas, S. M., and Jordan, T. R. (2004). Contributions of oral and extraoral facial movement to visual and audiovisual speech perception. *J. Exp. Psychol. Hum. Percept. Perform.* 30, 873–888. doi: 10.1037/0096-1523.30.5.873

Tomchek, S. D., and Dunn, W. (2007). Sensory processing in children with and without autism: a comparative study using the short sensory profile. *Am. J. Occup. Ther.* 61, 190–200. doi: 10.5014/ajot.61.2.190

Ujiie, Y., and Wakabayashi, A. (2015). Psychometric properties and overlap of the GSQ and AQ among Japanese university students. *Int. J. Psychol. Stud.* 7, 195–205. doi: 10.5539/ijps.v7n2p195

Wakabayashi, A., Baron-Cohen, S., and Ashwin, C. (2012). Do the traits of autism-spectrum overlap with those of schizophrenia or obsessive-compulsive disorder in the general population? *Res. Autism Spectr. Disord.* 6, 717–725. doi: 10.1016/j.rasd.2011.09.008

Wakabayashi, A., Baron-Cohen, S., and Wheelwright, S. (2006a). Are autistic traits an independent personality dimension? A study of the Autism-Spectrum Quotient (AQ) and the NEO-PI-R. *J. Pers. Individ. Dif.* 41, 873–883. doi: 10.1016/j.paid.2006.04.003

Wakabayashi, A., Baron-Cohen, S., Wheelwright, S., and Tojo, Y. (2006b). The Autism-Spectrum Quotient (AQ) in Japan: a crosscultural comparison. *J. Autism Dev. Disord.* 36, 263–270. doi: 10.1007/s10803-005-0061-2

Williams, J. H. G., Massaro, D. W., Peel, N. J., Bosseler, A., and Suddendorf, T. (2004). Visual–auditory integration during speech imitation in autism. *Res. Dev. Disabil.* 25, 559–575. doi: 10.1016/j.ridd.2004.01.008

Woynaroski, T. G., Kwakye, L. D., Foss-Feig, J. H., Stevenson, R. A., Stone, W. L., and Wallace, M. T. (2013). Multisensory speech perception in children with autism spectrum disorders. *J. Autism Dev. Disord.* 43, 2891–2902. doi: 10.1007/s10803-013-1836-5

Zhao, S., Uono, S., Yoshimura, S., Kubota, Y., and Toichi, M. (2013). Can gaze-cueing be helpful for detecting sound in autism spectrum disorder? *Res. Autism Spectr. Disord.* 7, 1250–1256. doi: 10.1016/j.rasd.2013.07.001

Zilbovicius, M., Meresse, I., Chabane, N., Brunelle, F., Samson, Y., and Boddaert, N. (2006). Autism, the superior temporal sulcus and social perception. *Trends Neurosci.* 29, 359–366. doi: 10.1016/j.tins.2006.06.004

Attentional bias in competitive situations: winner does not take all

Zhongqiang Sun, Tian Bai, Wenjun Yu, Jifan Zhou, Meng Zhang and Mowei Shen*

Department of Psychology and Behavioral Sciences, Zhejiang University, Hangzhou, China

Compared to previous studies of competition with participants' direct involvement, the current study for the first time investigated the influence of competitive outcomes on attentional bias from a perspective of an onlooker. Two simple games were employed: the Rock-Paper-Scissors game (Experiment 1) in which the outcome is based on luck, and Arm-wrestling (Experiment 2), in which the outcome is based on the competitors' strength. After observing one of these games, participants were asked to judge a stimulus presented on either the winner's or loser's side of a screen. Both experiments yielded the same results, indicating that the onlookers made much quicker judgments on stimuli presented on the loser's side than the winner's side. This suggests the existence of an attention bias for loser-related information once a competition has ended. Our findings provide a new lens through which the influence of competition results on human cognitive processing can be understood.

Keywords: competition, winner, loser, attentional bias, evolutionary psychology

Some battles you win, some battles you lose.

– The Romance of the Three Kingdoms

Competition is a ubiquitous and age-old behavior pattern and can range from a rivalry between two contestants to a war among several tribes. As the opening quotation suggests, competition is cruel because victory and defeat always come along with it. With regard to the influence of competition on the surrounding, the outcome of the rivalry may be the most important aspect. For instance, in social context, victory or defeat in war could potentially determine the survival of a tribe; while in dyadic context, an individual's win or lose in a competition could also affect the way of being treated by other people. There's an old saying that winner takes all. Is it also true in the social-cognitive processes? How would the asymmetric competing outcome direct the third-party onlooker's early stage processing on winner-/loser-related information? The answer still remains unclear, and is what we concerned in the current research.

Psychological research has valued the study of competition for decades. Compared to the research from other multiple disciplines, including sociology (e.g., Axelrod, 1997; Podolny, 2010), organizational behavior (Malhotra, 2010), education (Conti et al., 2001), and even biology and ecology (Earley et al., 2013), psychological studies pay more attention to information processing and behavioral patterns during the competitive interaction, as well as the influence of competition on subsequent interactions with others in the social group. Most studies have focused on the interaction during the competitive process, in which competition has been found to affect the actions of the moment (Ruys and Aarts, 2010) and the judgment and evaluation of others (Vonk, 1998), and to even distort cognitive representations (Xiao et al., 2012). In addition,

compared to cooperation, competition is different in terms of both individual action patterns and neural activation (Decety et al., 2004; Georgiou et al., 2007).

As a shared experience, the competitive scenario in society consists of three components: the competitor, the competitive process, and the effects of the outcome on others in the social group. The first two components are determined by the competitors themselves, and the latter one is determined by the third-party onlookers. These two types of people may experience very different cognitive processes, so a full picture of the competition event requires an integration of these processes. However, previous research has focused on the former two components, resulting in an insufficient understanding of the process from the perspective of the third-party onlookers.

The existing findings related to the third-party perspective, however, limited, mainly concern the high-level conscious processing of logical reasoning and moral judgment, especially with regard to third-party punishment, which concerns how non-stakeholders punish the offender during the competition (e.g., Fehr and Gächter, 2002; Fehr and Fischbacher, 2004).

Although such mechanisms underlying high-order processes have been examined by empirical studies, in terms of evolutionary theories, adaptive psychological mechanisms are presumed to exist at all levels of cognition, including both the aforementioned high-order processes and the relatively automatic early stage forms, such as attention and perception. Early stage cognition is of equal importance to the high-order process because those underlying mechanisms are the cornerstones shaping adaptive high-order cognition (see Kurzban et al., 2001; Maner et al., 2007); however, this area has been left relatively unexplored.

Therefore, the current research aimed to fill this gap in the literature by placing an emphasis on how the competitive outcomes influence the third-party onlookers in terms of the distribution of early stage attentional resources. In particular, by displaying a competitive interaction to an individual, we investigated the shifting of attention immediately after the outcome being announced. As we know, attention is the door to human cognition, and all unequal distributions of our cognitive resources originate from attention. It is also a key to enable us to understand this world by selecting relevant information out of irrelevant noise and processing the important parts of the information we receive (Carrasco, 2011). In this specific competitive situation, the outcomes were always asymmetric (i.e., not a draw). The side (winner or loser) with higher subjective value should capture more visual attention, and the information relevant to the winner and loser would then be processed differently from the perspective of the third-party onlooker.

We hypothesized that the loser, as a kind of negative stimulus, would capture attention first. It is evolutionarily adaptive for negative information to be more influential than positive (Baumeister et al., 2001) because negative things may threaten one's survival. This advantage in terms of processing negative stimuli has been extensively demonstrated from multiple aspects. For instance, compared to positive stimuli, negative stimuli capture attention earlier (Eastwood et al., 2001; Fox et al., 2002; Koster et al., 2004; Soares et al., 2009), are memorized more

solidly (Taylor, 1991), and are constructed with more cognitive interpretations (Abele, 1985).

In social interactions particularly, cooperating with each other is an effective way to help individuals increasing their fitness. It can be risky, though, since the chance of survival also depends largely on how the individuals choose their partner. If a loser who is an incapable partner is chosen, the strength of the group will be heavily discounted, which may further hinder the achievement of group success. In this sense, the strategy of cooperating without considering the capability and history of one's partners is not optimal in the long run. Instead, a more egoistic strategy would enable an individual to detect and subsequently avoid a loser.

Similar mechanisms have been suggested in cheater detection studies (Cosmides and Tooby, 1989; Tooby and Cosmides, 2005). Research shows an enhanced memory for untrustworthy faces rather than for trustworthy faces, revealing that untrustworthy faces were of high ecological value and relatively salient (e.g., Mealey et al., 1996; Oda, 1997; Yamagishi et al., 2003; Bayliss and Tipper, 2006). This mechanism of human bias in information processing may exist not only for untrustworthy individuals in social exchange, but also for a range of other harmful stimuli (Bell and Buchner, 2012). Given the reviewed empirical studies, we predicted an attentional bias toward loser-related information.

To investigate the influence of a competition situation, it is true that a naturalistic context or paradigm would be ideal, but the results would be affected by too many uncontrolled factors simultaneously. For instance, the winner and loser are likely to show different expressions and behaviors at the conclusion of an agonistic encounter (Lippold et al., 2008; Matsumoto and Hwang, 2012), which would affect the onlooker's attention distribution to a large extent. Fortunately, it is possible to control for these factors in the context of a laboratory experiment. Our paradigm was to display to participants two kinds of competitive games on a computer screen to represent the competition situation. In this way, we excluded the personal features of the competitors and isolated the winning/losing information, enabling control of detection of onlooker's rapid switching of attention. In Experiment 1, the Rock-Paper-Scissors game (RPS) was presented as the competitive situation. As a popular and simple game, it is widely used to study competition-related issues (e.g., Sinervo and Lively, 1996; Semmann et al., 2003; Wang et al., 2014). This game has a strong advantage in that three candidate actions are mutually restricted, and no action holds absolute predominance: Rock defeats Scissors, Scissors defeat Paper, and Paper defeats Rock. To be specific, each gesture could be either the winner or the loser in different pair condition, which could be regarded as a counterbalance procedure. By synthesizing all three pairs conditions in our analysis, the influence of visual difference of various stimuli could be minimized. However, the outcomes of the RPS game are considered to be based on luck to a large extent. Given that most match results are based on competitors' different capabilities, we hence adopted in Experiment 2 another popular game, arm-wrestling, which requires actual strength. Rapid reaction and judgment are necessary to examine automatic early stage mechanisms. Both games employed in the current study are advantageous for this consideration.

Experiment 1

Methods

Participants

Fourteen participants (seven females, 18–26 years-old) were paid to participate in the experiment. All had no history of neurological problems and had normal or corrected-to-normal vision. The participants provided written and informed consent before the experiments, and the procedures were in compliance with the Code of Ethics of the World Medical Association (Declaration of Helsinki), as well as approved by the Research Ethics Board of Zhejiang University.

The sample size in the current study was determined by a power analysis based on predicted effect size, using G*power 3 (Faul et al., 2007, 2009). According to the effect size ($\eta_p^2 = 0.22$) obtained from the pilot experiment, the analysis suggested a sample size of 14. This sample size was adopted in all the following experiments.

Stimuli

Three pictures of gestures were adopted from the RPS game (see **Figure 1A**). In order to eliminate the influence of luminance difference, the gesture pictures were monochromatized to black (0, 0, 0, RGB). Stimuli were presented on a gray background (80, 80, 80) CRT monitor of a 17-inch computer (100 Hz refresh rate). Each gesture occupied a $3° \times 4°$ rectangular area, centered $5°$ to the left or right of a central fixation cross. The direction of each gesture horizontally pointed to either left or right in different experimental conditions. Two or three dots were set as the test item.

Design and Procedure

Participants were seated in an electrically shielded and sound-attenuated recording chamber at a distance of 70 cm from the CRT monitor. Participants were asked to keep their eyes centrally fixated.

The procedure of Experiment 1 is shown in **Figure 1B**. We designed a dual-task paradigm, which was revised from Posner's cueing paradigm (Posner, 1980). Each trial began with a fixation cross presented randomly for a duration of 500 to 1000 ms; then, a gesture array was displayed for 1000 ms, consisting of two same or different gestures. When two different gestures were displayed, the winner and loser were determined (i.e., Rock defeats Scissors, Scissors defeats Paper, and Paper defeats Rock); otherwise, a draw was declared. Then, a 100-ms blank interval was inserted, followed by a 2000-ms test item. The test item was located either at the same position as the winner (Test-in-Winner condition) in 50% of the non-draw trials or the same position as the loser (Test-in-Loser condition) in the rest of the non-draw trials. The position of the winning gesture was balanced between left and right. If it was a draw, the test item was randomly located in either the left or the right visual field. The participant was first required to indicate whether the test item contained two dots or three dots by pressing one of two keys, with accuracy rather than response speed being stressed. Then, after a 500-ms blank interval, a secondary task required the participant to recall whether the Paired Gesture was

a draw and respond by pressing one of another two keys. This secondary task was presented to keep the participants involved when seeing the gesture array. The interval between trials was randomly set from 1000 to 1500 ms (see detailed videos on the website: http://www.psych.zju.edu.cn/english/redir.php?catalog id=15773).

Each participant completed 48 trials for each of the two test-item positions (Test-in-Winner and Test-in-Loser), which were evenly distributed among the three possible winner-loser situations (Rock-Scissors, Scissors-Paper, and Paper-Rock). They completed another 48 trials for the draw condition, resulting in a total of 144 randomly presented trials. The whole experiment was divided into three blocks with a 2-min break between blocks. Before the formal experiment, there were at least 20 practice trials to ensure that the participants understood the instructions.
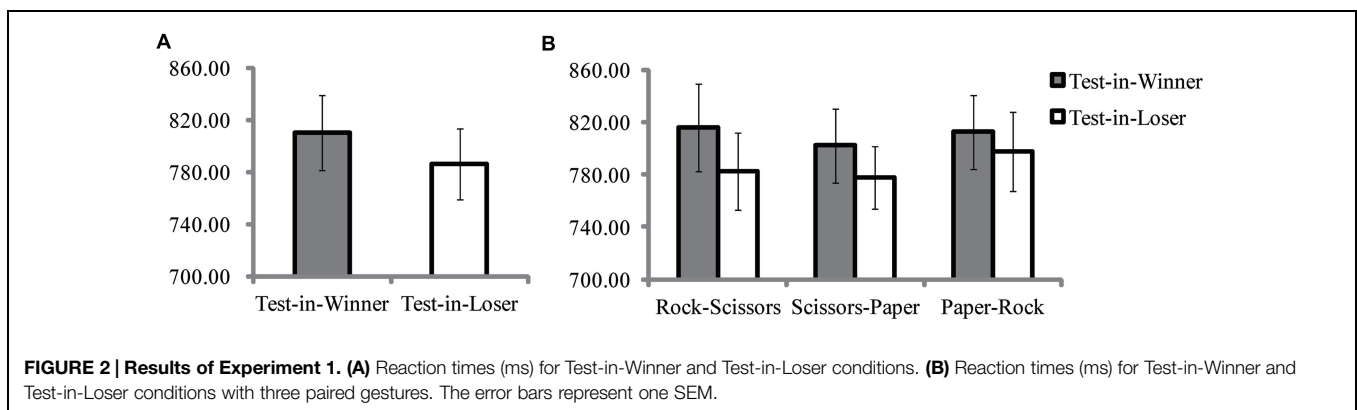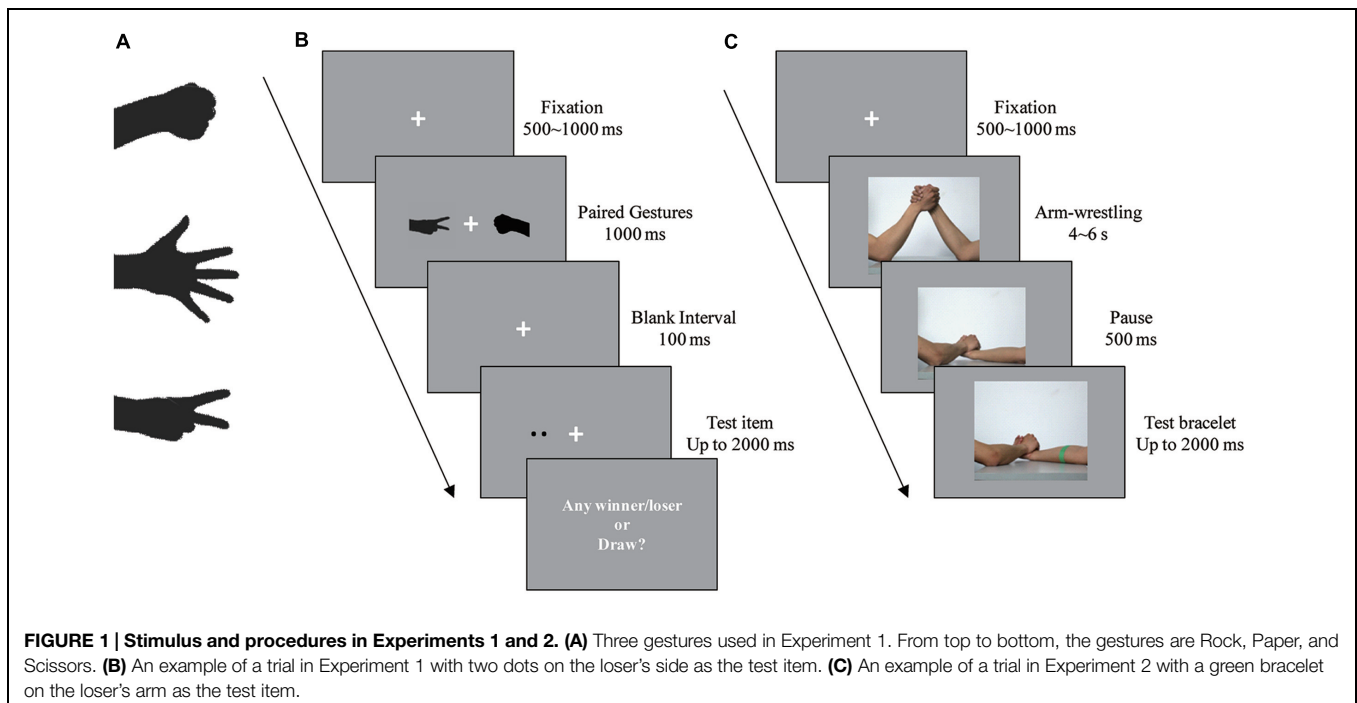
Results

Trials with inaccurate responses were excluded from the reaction time (RT) analyses (7.19% of all trials), as well as the outliers with RTs more than 2 SD above or below the mean (4.81% of all trials).

To exclude the potential influence of unilateral advantage, we compared the RTs in the draw condition between the situations when the dot was displayed in the left and right visual fields, and no significant difference was found (left, mean $\pm$ SD, 794.92 $\pm$ 119.39 ms; right, 805.93 $\pm$ 115.56), $t(13) = -0.81$, $p > 0.250$. Given that draws were not of interest to us, we will not discuss draw outcomes in the following sections.

We conducted a two-way analysis of variance (ANOVA) for dual-task RT and accuracy, with test-item position (Test-in-Winner and Test-in-Loser) and winner-loser situations (Rock-Scissors, Scissors-Paper, and Paper-Rock) as independent variables for non-draw data.

Interestingly, for RT, a significant main effect for test-item position was found, $F(1,13) = 7.03$, $p = 0.020$, $\eta_p^2 = 0.35$, while none was found for winner-loser situation (see **Figure 2A**), $F(2,26) = 1.08$, $p > 0.250$, $\eta_p^2 = 0.08$. *Post hoc* contrast analyses revealed a somewhat faster response speed for items on the loser side [784.83 $\pm$ 101.25, 95% Confidence Interval or 95% CI (727.53, 842.13)] compared to the winner side [811.34 $\pm$ 108.61, 95% CI (749.05, 873.63)]. Moreover, no interaction was found between test-item position and winner-loser situation, $F(2,26) = 0.93$, $p > 0.250$, $\eta_p^2 = 0.07$, implying that performance with all three winner-loser situations shared almost the same tendency in terms of results (see **Figure 2B**)[1]. No main effect for accuracy was found for either test-item position [Test-in-Winner, 93.06 $\pm$ 3.89%, 95% CI (90.81%, 95.30%)]; [Test-in-Loser, 92.56 $\pm$ 3.68%, 95% CI (90.44%, 94.68%)], $F(1,13) = 0.28$, $p > 0.250$, $\eta_p^2 = 0.02$, or winner-loser situation, $F(2,26) = 0.08$, $p > 0.250$, $\eta_p^2 = 0.01$, nor was there interaction between the variables, $F(2,26) = 1.23$, $p > 0.250$, $\eta_p^2 = 0.09$. The accuracy results strongly confirmed that the salient RT difference was not due to a speed-accuracy trade-off.

---

[1]We also converted all RT data (Experiments 1 and 2) by logarithmic transformation (Ln), and re-analyzed those transformed data in both two experiments. The patterns of the results after logarithmic transformation were the same as the original statistical analysse.

FIGURE 1 | Stimulus and procedures in Experiments 1 and 2. (A) Three gestures used in Experiment 1. From top to bottom, the gestures are Rock, Paper, and Scissors. **(B)** An example of a trial in Experiment 1 with two dots on the loser's side as the test item. **(C)** An example of a trial in Experiment 2 with a green bracelet on the loser's arm as the test item.

FIGURE 2 | Results of Experiment 1. (A) Reaction times (ms) for Test-in-Winner and Test-in-Loser conditions. **(B)** Reaction times (ms) for Test-in-Winner and Test-in-Loser conditions with three paired gestures. The error bars represent one SEM.

Experiment 2

In Experiment 1, we found that the attention of the participants, as third-party onlookers, was captured by the information on the loser's side. One might argue that the gestures appearing in the RPS game are randomly chosen by the competitors and presented to the onlookers, and thus the outcomes are based on luck to a large extent. As we know, most competition outcomes are not the result of luck but are directly relevant to a competitor's actual ability, such as strength, or power; that is to say, the strongest wins. Therefore, it is of great importance to verify the above results in a situation that depends on a competitor's strength. Accordingly, in Experiment 2, we adopted another traditional type of match: arm-wrestling. During the match, two competitors hold their left or right hands together and try their best to press and move the other's hand. The one who presses the other's hand onto the table first is determined to be the winner. Unlike the RPS game, a draw cannot be

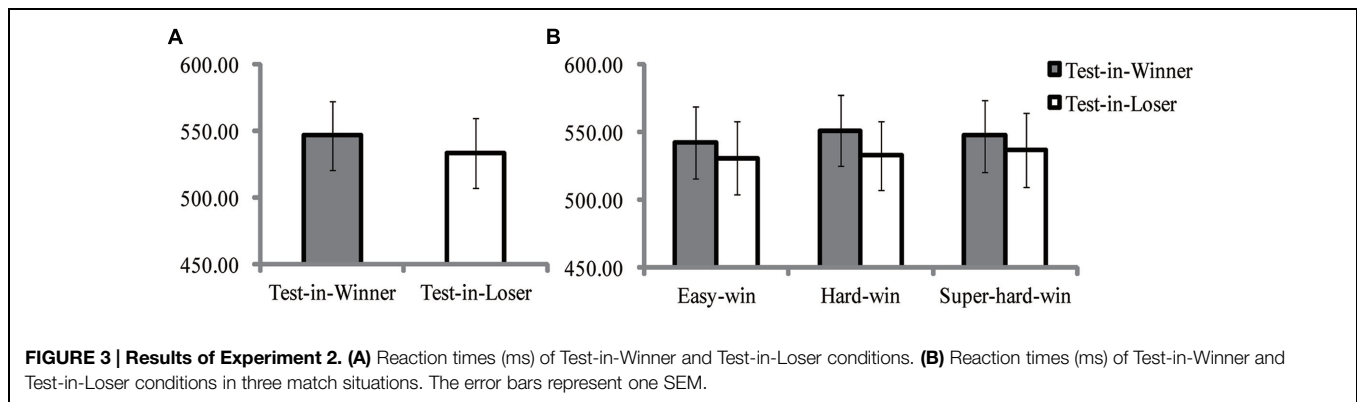declared because the winner and loser are always decided for each round.

Methods
Participants
A new group of 14 participants (five females, 19–26 years-old) was paid to participate in the experiment. Other aspects were the same as those in Experiment 1.

Stimuli
The arm-wrestling match between two volunteers (S & W) was recorded by a camera without showing any identifying information such as the face and clothes. Since the match itself is quite simple, participants may become accustomed to the video information after several trials and respond before the result comes out when they see the first part of the video. To prevent this occurrence, three different situations were adopted to play the match out when either S or W was the winner. The

FIGURE 3 | Results of Experiment 2. (A) Reaction times (ms) of Test-in-Winner and Test-in-Loser conditions. **(B)** Reaction times (ms) of Test-in-Winner and Test-in-Loser conditions in three match situations. The error bars represent one SEM.

following are three situations that could apply when S is the winner:

Easy-win: after a 3–4 s stalemate, S wins; total match duration of 4 s;

Hard-win: S plays more strongly than W at the beginning, and after a 1–2 s stalemate, S wins; total match duration of 5 s;

Super-hard-win: W plays more strongly than S at the beginning, and after a 1–2 s stalemate, S fights back to win; total match duration of 6 s.

The three same situations were applied when W was the winner. Those six videos were also processed to create another six mirrored versions by exchanging the position of the two volunteers in order to balance the position of the winner. All 12 videos were presented in a mixed order during the experiment. Each video occupied a 20° × 20° rectangular area on a gray background (80, 80, 80) of a 17-inch CRT monitor (100 Hz refresh rate).

In addition, we froze the last frame of the video in which the winner/loser had just been declared for use as the test picture, and we attached an extra red or green bracelet to the arm of one volunteer as the test item.

Design and Procedure

The procedure for Experiment 2 is shown in **Figure 1C**. After presentation of a fixation cross lasting 500–1000 ms, one of the videos was shown in the center of screen. Once the winner/loser was declared, the video paused for 500 ms, followed by a 2000-ms presentation of the test picture. In the test picture, the red or green bracelet was located on the winner's arm in 50% of the trials (Test-in-Winner condition) or the loser's in the rest of the trials (Test-in-Loser condition). The participant was required to judge the color of the bracelet, with accuracy rather than response speed being stressed. The interval between trials was randomly set from 1000 to 1500 ms.

Each participant completed 96 trials for each of the two conditions, with a total of 192 randomly presented trials. These trials were evenly distributed among the 12 aforementioned videos. The whole experiment was divided into four blocks with a 2-min break between blocks. Before the formal experiment, there were at least 20 trials for practice to ensure that the participants understood the instructions.

Results

Trials either with inaccurate responses (2.49% of all trials) or with RTs more than 2 SD above or below the mean (3.42% of all trials) were excluded from the RT analyses.

Similar to Experiment 1, two-way ANOVAs were conducted for both RT and accuracy, with test-item position (Test-in-Winner and Test-in-Loser) and match situation (Easy-win, Hard-win, and Super-hard-win) as independent variables.

The results for RT almost replicated those in Experiment 1. A significant main effect was only found for test-item position (see **Figure 3A**), $F(1,13) = 10.27$, $p = 0.007$, $\eta_p^2 = 0.44$. *Post hoc* contrast showed a relatively shorter RT in the Test-in-Loser condition [533.58 ± 98.28, 95% CI (477.78, 591.28)] than in Test-in-Winner [546.87 ± 96.45, 95% CI (491.74, 603.39)]. Nor significant main effect for match situation, $F(2,26) = 1.13$, $p > 0.250$, $\eta_p^2 = 0.08$, nor interaction was found (see **Figure 3B**), $F(2,26) = 0.49$, $p > 0.250$, $\eta_p^2 = 0.04$.

For accuracy, the only significant main effect was found for test item position, $F(1,13) = 6.01$, $p = 0.030$, $\eta_p^2 = 0.32$, while the main effect for match situation was not significant, $F(2,26) = 1.68$, $p = 0.210$, $\eta_p^2 = 0.11$. There was also no interaction between test-item position and match situation, $F(2,26) = 0.05$, $p > 0.250$, $\eta_p^2 = 0.004$. *Post hoc* analysis revealed a slightly higher accuracy when the test item was on the loser side [$98.09 \pm 1.51\%$, 95% CI (97.22, 98.96%)] rather than on the winner side [$96.97 \pm 2.39\%$, 95% CI (95.60, 98.35%)]. The accuracy results again excluded the potential influence of a speed-accuracy trade-off.

Discussion

The results indicated that the third-party onlookers made quicker judgments for stimuli presented on the loser's side compared to those on the winner's side, implying the existence of an attentional bias toward the loser. Two competitive games were included, based either on the competitors' ability or on random chance. In Experiment 1, in which the competitor throws rock, paper, or scissors randomly, the onlookers responded to stimuli presented on the loser's side much more quickly, though these stimuli were not directly relevant to the loser. In Experiment 2, we presented an arm-wrestling game, a competition that

required strength on the part of the competitor, and we attached a colored bracelet to the target competitor's arm. The results suggested a faster response for loser-related information, which replicated the pattern in the RPS game. The findings here demonstrated that no matter whether the competition outcome was decided randomly or with real strength, the onlookers vigilantly attended to stimuli that were relevant to the loser.

This attentional bias toward the loser in the competition was thus verified for the first time from the perspective of a third-party onlooker. The unequal information (i.e., winner and loser) that is generated from social interaction behaviors such as competition leads to a bias in third-party onlookers' early information processing. This fits with the theory that humans vigilantly attend to negative information, which is known as negativity bias. This bias focuses on the adaptive implications of negative-positive asymmetrical processes with the result that negative events are more salient and dominant in many situations (Taylor, 1991; Cacioppo and Berntson, 1994; Rozin and Royzman, 2001). If a negative stimulus is overlooked, people could lose some portion of their own resources, or even worse, pay the price of losing their life and decreasing the possibility of perpetuating their genes (Baumeister et al., 2001). In competitive situations, the loser represents this negative stimulus. Hence, it is reasonable to process this negative stimulus more quickly and accurately than neutral or positive stimuli, as this may result in an increased chance of survival.

Quickly detecting losers is not only a more egoistic strategy for individuals, but it is also a stable strategy for group survival. This can be analyzed using evolutionary game theory (Smith, 1982). If a strategy adopted by a population guides interactions and persists in the group for a long time because it produces more fitness benefit and outweighs any alternative strategy, it is known as an evolutionary stable strategy (ESS; Smith, 1982). In our study, it could be interpreted as a necessary condition to reliably detect the loser in a human interaction. For example, if an individual cannot reliably detect a loser, their unconditional collaborating with the loser will increase the fitness of any loser they meet in the population. When cooperating with a loser, however, his/her low probability of success will lead to an unrewarding cooperation, as well as a net fitness cost. As a result, a population of unconditional collaborators could be invaded and finally outcompeted due to their using this behavioral strategy with a lower probability of success, when compared to those who avoid losers and seeks winners with whom to cooperate. In this case, conditional cooperation, which requires the ability to detect losers, is an ESS.

Moving beyond previous studies in which participants were involved in competition as a contestant, the current research was instead conducted from the perspective of a third-party onlooker. Additionally, the complex competitive behaviors of humans were represented here by two simple and classical games, the RPSs and

Arm-wrestling, that could be manipulated easily in behavioral studies, thus providing a novel opportunity to investigate the current issue.

Furthermore, there might exist some interesting issues coming along with the current finding. Apart from attention, does the asymmetric competing outcome also affect human's other cognitive processing such as perception and memory? For instance, a mnemonic advantage was already found on cheater-related information (Bell and Buchner, 2012). Analogically, is it possible for loser to induce a similar mnemonic bias toward itself? Further studies need to examine the specific mechanism causing this attentional bias and extend its application. For instance, previous studies found that attention bias modification procedure could reduce attentional bias for threat, thereby diminishing anxiety symptom (e.g., MacLeod et al., 2002; Heeren et al., 2015; Linetzky et al., 2015). Therefore, it could be possible that experimental training inducing an attention bias toward gain-related material will modify competitive information vulnerability, which may decease the attentional bias toward loser-related information. Meanwhile, it is also intriguing to explore that whether this attention bias is innate from one's birth or acquired from social interaction experience later in life. Appropriate adjustment on current paradigm might benefit to find out its answer in children of different ages.

Conclusion

As the first and common doorway of cognition, attention helps us determine which information takes priority to be encoded. Further processing, such as logical reasoning or decision making, can only be accessed once the information has been attended to. The findings from these two experiments suggest an attentional bias toward loser-related information in a competitive situation. The current research advances the social study concerning competition, and develops the extent of studying the influence of this social interaction on our cognitive function, to an early stage processing level.

Acknowledgments

References

Abele, A. (1985). Thinking about thinking: causal, evaluative and finalistic cognitions about social situations. *Eur. J. Soc. Psychol.* 15, 315–332. doi: 10.1002/ejsp.2420150306

Axelrod, R. M. (1997). *The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration.* Princeton, NJ: Princeton University Press.

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., and Vohs, K. D. (2001). Bad is stronger than good. *Rev. Gen. Psychol.* 5, 323–370. doi: 10.1037/1089-2680.5.4.323

Bayliss, A. P., and Tipper, S. P. (2006). Predictive gaze cues and personality judgments should eye trust you? *Psychol. Sci.* 17, 514–520. doi: 10.1111/j.1467-9280.2006.01737.x

Bell, R., and Buchner, A. (2012). How adaptive is memory for cheaters? *Curr. Dir. Psychol. Sci.* 21, 403–408. doi: 10.1177/0963721412458525

Cacioppo, J. T., and Berntson, G. G. (1994). Relationship between attitudes and evaluative space: a critical review, with emphasis on the separability of positive and negative substrates. *Psychol. Bull.* 115, 401–423. doi: 10.1037/0033-2909.115.3.401

Carrasco, M. (2011). Visual attention: the past 25 years. *Vision Res.* 51, 1484–1525. doi: 10.1016/j.visres.2011.04.012

Conti, R., Ann, M., and Picariello, M. L. (2001). The impact of competition on intrinsic motivation and creativity: considering gender, gender segregation and gender role orientation. *Pers. Individ. Differ.* 31, 1273–1289. doi: 10.1016/S0191-8869(00)00217-8

Cosmides, L., and Tooby, J. (1989). Evolutionary psychology and the generation of culture, part II: case study: a computational theory of social exchange. *Ethol. Sociobiol.* 10, 51–97. doi: 10.1016/0162-3095(89)90013-7

Decety, J., Jackson, P. L., Sommerville, J. A., Chaminade, T., and Meltzoff, A. N. (2004). The neural bases of cooperation and competition: an fMRI investigation. *Neuroimage* 23, 744–751. doi: 10.1016/j.neuroimage.2004.05.025

Earley, R. L., Lu, C., Lee, I., Wong, S. C., and Hsu, Y. (2013). Winner and loser effects are modulated by hormonal states. *Front. Zool.* 10:6. doi: 10.1186/1742-9994-10-6

Eastwood, J. D., Smilek, D., and Merikle, P. M. (2001). Differential attentional guidance by unattended faces expressing positive and negative emotion. *Percept. Psychophys.* 63, 1004–1013. doi: 10.3758/BF03194519

Faul, F., Erdfelder, E., Buchner, A., and Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. *Behav. Res. Methods* 41, 1149–1160. doi: 10.3758/BRM.41.4.1149

Faul, F., Erdfelder, E., Lang, A. G., and Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* 39, 175–191. doi: 10.3758/BF03193146

Fehr, E., and Fischbacher, U. (2004). Third-party punishment and social norms. *Evol. Hum. Behav.* 25, 63–87. doi: 10.1016/S1090-5138(04)00005-4

Fehr, E., and Gächter, S. (2002). Altruistic punishment in humans. *Nature* 415, 137–140. doi: 10.1038/415137a

Fox, E., Russo, R., and Dutton, K. (2002). Attentional bias for threat: evidence for delayed disengagement from emotional faces. *Cogn. Emot.* 16, 355–379. doi: 10.1080/02699930143000527

Georgiou, I., Becchio, C., Glover, S., and Castiello, U. (2007). Different action patterns for cooperative and competitive behaviour. *Cognition* 102, 415–433. doi: 10.1016/j.cognition.2006.01.008

Heeren, A., Mogoase, C., McNally, R. J., Schmitz, A., and Philippot, P. (2015). Does attention bias modification improve attentional control? A double-blind randomized experiment with individuals with social anxiety disorder. *J. Anxiety Disord.* 29, 35–42. doi: 10.1016/j.janxdis.2014.10.007

Koster, E. H., Crombez, G., Van Damme, S., Verschuere, B., and De Houwer, J. (2004). Does imminent threat capture and hold attention? *Emotion* 4, 312–317. doi: 10.1037/1528-3542.4.3.312

Kurzban, R., Tooby, J., and Cosmides, L. (2001). Can race be erased? Coalitional computation and social categorization. *Proc. Natl. Acad. Sci. U.S.A.* 98, 15387–15392. doi: 10.1073/pnas.251541498

Linetzky, M., Pergamin-Hight, L., Pine, D. S., and Bar-Haim, Y. (2015). Quantitative evaluation of the clinical efficacy of attention bias modification treatment for anxiety disorders. *Depress. Anxiety* 32, 383–391. doi: 10.1002/da.22344

Lippold, S., Fitzsimmons, L. P., Foote, J. R., Ratcliffe, L. M., and Mennill, D. J. (2008). Post-contest behaviour in black-capped chickadees (*Poecile atricapillus*): loser displays, not victory displays, follow asymmetrical countersinging exchanges. *Acta Ethol.* 11, 67–72. doi: 10.1007/s10211-008-0043-4

MacLeod, C., Rutherford, E., Campbell, L., Ebsworthy, G., and Holker, L. (2002). Selective attention and emotional vulnerability: assessing the causal basis of their association through the experimental manipulation of attentional bias. *J. Abnorm. Psychol.* 111, 107–123. doi: 10.1037/0021-843X.111.1.107

Malhotra, D. (2010). The desire to win: the effects of competitive arousal on motivation and behavior. *Organ. Behav. Hum. Decis. Process.* 111, 139–146. doi: 10.1016/j.obhdp.2009.11.005

Maner, J. K., Gailliot, M. T., and DeWall, C. N. (2007). Adaptive attentional attunement: evidence for mating-related perceptual bias. *Evol. Hum. Behav.* 28, 28–36. doi: 10.1016/j.evolhumbehav.2006.05.006

Matsumoto, D., and Hwang, H. S. (2012). Evidence for a nonverbal expression of triumph. *Evol. Hum. Behav.* 33, 520–529. doi: 10.1016/j.evolhumbehav.2012.01.005

Mealey, L., Daood, C., and Krage, M. (1996). Enhanced memory for faces of cheaters. *Ethol. Sociobiol.* 17, 119–128. doi: 10.1016/0162-3095(95)00131-X

Oda, R. (1997). Biased face recognition in the prisoner's dilemma game. *Evol. Hum. Behav.* 18, 309–315. doi: 10.1016/S1090-5138(97)00014-7

Podolny, J. M. (2010). *Status Signals: A Sociological Study of Market Competition*. Princeton, NJ: Princeton University Press.

Posner, M. I. (1980). Orienting of attention. *Q. J. Exp. Psychol.* 32, 3–25. doi: 10.1080/00335558008248231

Rozin, P., and Royzman, E. B. (2001). Negativity bias, negativity dominance, and contagion. *Pers. Soc. Psychol. Rev.* 5, 296–320. doi: 10.1207/S15327957PSPR0504_2

Ruys, K., and Aarts, H. (2010). When competition merges people's behavior: interdependency activates shared action representations. *J. Exp. Soc. Psychol.* 46, 1130–1133. doi: 10.1016/j.jesp.2010.05.016

Semmann, D., Krambeck, H. J., and Milinski, M. (2003). Volunteering leads to rock–paper–scissors dynamics in a public goods game. *Nature* 425, 390–393. doi: 10.1038/nature01986

Sinervo, B., and Lively, C. M. (1996). The rock-paper-scissors game and the evolution of alternative male strategies. *Nature* 380, 240–243. doi: 10.1038/380240a0

Smith, J. M. (1982). *Evolution and the Theory of Games*. Cambridge: Cambridge university press. doi: 10.1017/CBO9780511806292

Soares, S. C., Esteves, F., and Flykt, A. (2009). Fear, but not fear-relevance, modulates reaction times in visual search with animal distractors. *J. Anxiety Disord.* 23, 136–144. doi: 10.1016/j.janxdis.2008.05.002

Taylor, S. E. (1991). Asymmetrical effects of positive and negative events: the mobilization-minimization hypothesis. *Psychol. Bull.* 110, 67–85. doi: 10.1037/0033-2909.110.1.67

Tooby, J., and Cosmides, L. (2005). "Conceptual foundations of evolu–tionary psychology," in *The Handbook of Evolutionary Psychology*, ed. D. M. Buss (Hoboken, NJ: John Wiley & Sons), 5–67.

Vonk, R. (1998). Effects of cooperative and competitive outcome dependency on attention and impression preferences. *J. Exp. Soc. Psychol.* 34, 265–288. doi: 10.1006/jesp.1998.1350

Wang, Z., Xu, B., and Zhou, H. H. (2014). Social cycling and conditional responses in the Rock-Paper-Scissors game. *Sci. Rep.* 4, 5830. doi: 10.1038/srep05830

Xiao, Y. J., Van Bavel, J. J., Van Bavel, J. J., and Cunningham, W. A. (2012). See your friends close and your enemies closer: social identity and identity threat shape the representation of physical distance. *Pers. Soc. Psychol. Bull.* 38, 959–972. doi: 10.1177/0146167212442228

Yamagishi, T., Tanida, S., Mashima, R., Shimoma, E., and Kanazawa, S. (2003). You can judge a book by its cover: evidence that cheaters may look different from cooperators. *Evol. Hum. Behav.* 24, 290–301. doi: 10.1016/S1090-5138(03)00035-7

Pupillometric evidence for the locus coeruleus-noradrenaline system facilitating attentional processing of action-triggered visual stimuli

Ken Kihara[1]*, Tatsuto Takeuchi[2], Sanae Yoshimoto[2], Hirohito M. Kondo[3,4] and Jun I. Kawahara[5]

[1] Graduate School of Science and Engineering, Kagoshima University, Kagoshima, Japan, [2] Department of Psychology, Japan Women's University, Kanagawa, Japan, [3] NTT Communication Science Laboratories, NTT Corporation, Kanagawa, Japan, [4] United Graduate School of Child Development, Osaka University, Osaka, Japan, [5] Department of Psychology, Hokkaido University, Hokkaido, Japan

It has been argued that attentional processing of visual stimuli is facilitated by a voluntary action that triggers the stimulus onset. However, the relationship between action-induced facilitation of attention and the neural substrates has not been well established. The present study investigated whether the locus coeruleus-noradrenaline (LC-NA) system is involved in this facilitation effect. A rapid serial visual presentation paradigm was used to assess the dynamics of transient attention in humans. Participants were instructed to change a digit stream to a letter stream by pressing a button and specifying successive targets of four letters. Pupil dilation was measured as an index of LC-NA function. Accuracy of target identification was better when the temporal delay between participants' key press and target onset was 800 ms than when targets appeared just after the key press or when targets appeared without key press. Accuracy of target identification was positively correlated with both the peak amplitude of pupil dilation and the pupil size at the time of the key press. These results indicate that target identification in the visual task is closely linked to pupil dilation. We conclude that the LC-NA system plays an important role in the facilitation of transient attention driven by voluntary action.

Keywords: temporal attention, voluntary action, locus coeruleus-noradrenaline system, pupillometry, rapid serial visual presentation

Introduction

We interact with the environment around us on a daily basis in order to achieve various goals. Moreover, the visual events that appear in this environment are often triggered by our voluntary actions. The triggering of a visual stimulus voluntarily may modulate our ability to allocate attention toward a particular point in time, or temporal attention, thereby contributing to perceptual enhancement (Hommel et al., 2001; Stock and Stock, 2004). The attentional blink paradigm is useful in the examination of limits of temporal attention for focusing on a series of visual events (Raymond et al., 1992). Typically in an attentional blink paradigm, two targets are embedded in a rapid serial visual presentation (RSVP) stream of non-targets and viewers must identify both targets. When these targets are separated in a stream by less than 500 ms, identification of the first target impairs processing of the second target (i.e.,

the attentional blink deficit). The attentional blink deficit is generally considered to result from a failure in temporal orienting of attention to the second target after processing the first target (see Martens and Wyble, 2010, for a recent review). In the context of the attentional blink, transient attention which develops rapidly after onset of a visual target, reaches a peak around 100 ms, and decays quickly thereafter (Weichselgartner and Sperling, 1987; Müller and Rabbitt, 1989; Nakayama and Mackeben, 1989). It is considered to be involved in the limitation of the temporal attention (e.g., Olivers and Meeter, 2008). Notably, Kihara and Kawahara (2011) demonstrated that when participants voluntarily triggered the appearance of the first target, the identification accuracy of both first and second target increased. This finding suggests that transient attention driven by the onset of the first target is facilitated by voluntary triggering of the first target, enhancing the first-target processing itself. This permits a rapid orienting of temporal attention from the first to the second target, thus reducing the subsequent attentional blink deficit. However, the neurobiological mechanisms responsible for this attentional facilitation of visual processing triggered by voluntary actions have remained unclear.

Here we postulate that the locus coeruleus-noradrenaline (LC-NA) system is responsible for this attentional facilitation. It is known that the LC-NA system plays an important role in the enhancement of transient attention (for reviews, see Berridge and Waterhouse, 2003; Aston-Jones and Cohen, 2005; Corbetta et al., 2008; Sara and Bouret, 2012). The LC neuronal responses in monkeys are selectively elicited within 100 ms by successful detection of the onset of visual targets; such phasic responses do not occur with missed targets and distractors (Aston-Jones et al., 1994; Clayton et al., 2004). The NA levels in the parietal cortex are induced by LC phasic responses (Foote and Morrison, 1987), and they result in the facilitation of transient attention. In fact, it has been claimed that the LC-NA system is involved in the attentional blink deficit which is modulated by transient attention (Nieuwenhuis et al., 2005).

Previous findings lead to the expectation that attentional processes are facilitated by a voluntary action prior to a visual stimulus, and that the LC-NA system is involved in these processes. However, there is no empirical evidence for a relationship between the voluntary triggering of the visual stimuli and activation of the LC-NA system. To address this, the present study measured pupil diameters in order to identify whether attention processes are affected by the LC-NA system because pupil dilation reflects a phasic response of LC neurons (Koss, 1986; Aston-Jones and Cohen, 2005; Samuels and Szabadi, 2008; Murphy et al., 2011; Laeng et al., 2012; Eldar et al., 2013). The firing rate of LC neurons in monkeys is highly correlated with changes in pupil diameter (Rajkowski et al., 1993). Furthermore, a number of previous studies have demonstrated that the LC-NA system contributes to performance of attentional tasks, as indexed by pupil measurement (for a review, see Laeng et al., 2012). Accordingly, it is reasonable to assume that involvement of the LC-NA system in attentional facilitation, as reflected in pupil dilation, can be observed

during a voluntary action that is directed toward a future visual stimulus.

The purpose of the present study was to investigate whether the LC-NA system is involved in the facilitation of transient attention of visual stimuli triggered by voluntary actions, using the pupillary response as a dependent variable. As mentioned earlier, previous studies indicate that transient attention is critical for reporting visual targets presented among an RSVP of to-be-ignored distractors (Olivers and Meeter, 2008; Martens and Wyble, 2010). Weichselgartner and Sperling (1987) used a simple task in which a set of four consecutive targets appeared in an RSVP stream to investigate the nature of transient attention. In this task, via pressing labeled keys, participants reported the identities of four successive targets embedded in a rapid stream of visual non-targets. A typical result of this type of task is that the first target is correctly reported more often than other targets. It is known that up to four items can be maintained within short-term memory (Miller, 1956; Cowan, 2001). If no distractors were presented after the targets, four successive targets could be reported correctly (Olivers et al., 2007). Thus, the differences in the accuracy of reporting targets reflect transient attention, not memory limitations. In addition, previous pupillometric studies have shown the relationship between the activation of the LC-NA system and the identification of a single or multiple target(s) embedded in an RSVP stream (Privitera et al., 2010; Zylberberg et al., 2012). Accordingly, by examining pupillary changes during the RSVP task, we can evaluate the contribution of the LC-NA system to the facilitation of transient attention induced by self-triggered stimulus targets.

In the present study, an RSVP stream consisted of distractors and targets, which respectively, formed two successive strings of items. In the current task, the RSVP stream opened with a variable length string of distractor digits and this was followed by a string of target letters. The experimental conditions required participants to voluntarily trigger, via a key press, the onset of the string of target letters. Finally, we measured changes in participants' pupillary responses to both the key press and target stimuli. Specifically, we examined the relationship between pupillary change and accuracy of target identification by manipulating the temporal delays between the key press and target onset (Kihara and Kawahara, 2012). We hypothesized that if the LC-NA system contributes to the facilitation of the transient attention induced by self-triggered stimulus targets, then we should find a positive correlation between the accuracy of target identification and maximum pupil diameter size, observed after the voluntary key press. Note that although pupil dilation also occurs if targets need to be memorized for later recall (Peavler, 1974), similar pupil dilation would be observed if the number of to-be-reported targets is the same (Kahneman and Beatty, 1966; Beatty, 1982). A previous study on different patterns of pupillary change suggested that the temporal delays between key press and target onset affect attentional state (Kihara and Kawahara, 2012). Thus, it is possible to separate attention-related pupil dilation from memory-related dilation.

Materials and Methods

Participants

Thirty four adults (18 males and 16 females, mean age 22.8 years, range 20–42 years) participated in this experiment. Data from six participants were excluded due to excessive artifacts in their pupil recordings, leaving the data from 28 participants for subsequent analysis. All had self-reported normal or corrected-to-normal vision. Written informed consent was obtained from all participants. This experiment was in accordance with the Declaration of Helsinki and approved by the Committee of Ethics, Chukyo University.

Apparatus

The experiment was conducted in a darkened room. Stimuli were presented on a 17-inch computer monitor driven at a 60-Hz refresh rate and controlled by MATLAB with the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997). Viewing distance was 57 cm, and head position was maintained by a chin rest. Pupil diameter was recorded using a ViewPoint Eye Tracker (Arrington Research, Inc. Scottsdale, AZ, USA) with a sampling rate of 220 Hz. A video camera and infrared light-emitting diodes were positioned in front of the right eye. The eye tracker was calibrated to each participant at the start of each block of experimental trials. Artifacts and eye blinks were detected by the eye tracking software, and trials in which eye blinks occurred during the time window from the start of the RSVP to the onset of the fourth target were discarded as failing to obtain pupil data.

Stimuli and Procedures

Figure 1 illustrates the sequence of events on a single trial. The RSVP stream consisted of two parts; the first was a stream of 200 (or fewer) digits and the second was a stream of 20 capital letters. The digits were randomly chosen from 0 to 9, with a constraint that the same digit was not presented successively. The capital letters were randomly chosen from A to Z, excluding the letters I, O, and Q. Identical letters were never presented in a trial. Each item subtended a visual angle of approximately $1° \times 1°$. We used the presentation rate of 50 ms per item because previous studies (Martens et al., 2006; Bowman and Wyble, 2007) have successfully shown minute temporal dynamics in a regular two-target attentional blink procedure at this rate of presentation and our pilot data with this presentation rate demonstrated observable transient attentional effects (Kihara and Kawahara, 2011). The items were colored in dark gray (1.2 cd/m^2) against a black background (0.3 cd/m^2).

Letter onset time was also varied as a within-participants letter-onset factor. This variable reflects the temporal delay between the voluntary action and onset time of the letter stream. This variable allows observation of the temporal course of the impact of the voluntary action upon LC-NA system as indexed in changes in a participant's pupillary response, which was eventually reflected in target reporting scores (Kihara and Kawahara, 2011). According to the preliminary findings, the letter onset variable can be effectively realized with eight letter-onset conditions: seven self conditions and one automatic condition. Each of the eight conditions was presented for
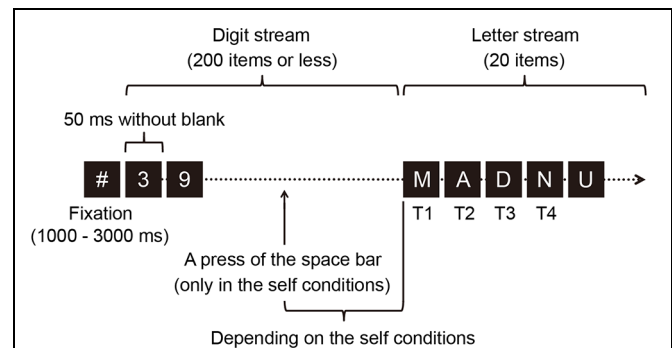

FIGURE 1 | A schematic illustration of the event sequence in a trial. Each item in a rapid serial visual presentation (RSVP) stream was presented for 50 ms with no inter-stimulus interval. Each RSVP stream consisted of a digit stream and a letter stream. The procedure involved eight letter-onset conditions. Seven were self conditions in which a participant voluntarily pressed a key to determine the onset time of the switch from the digit sequence to the letter sequence. The voluntary action in these conditions involved a participant voluntarily initiating this switch by pressing the space bar at any arbitrary time (within 10 s). Pre-determined temporal delays occurred between the voluntary key press and the onset of the first capital letter (0–50, 100–150, 200–250, 300–350, 400–450, 600–650, or 800–850 ms). The other condition was an automatic condition, in which the onset time of the letter sequence was automatically determined. In both conditions, participants reported the first, second, third, and fourth targets (T1, T2, T3, and T4, respectively) at the end of the target stream.

one block of forty trials. The order of the block presentation was randomized across participants, except that the automatic condition was not presented as the first block. In the self conditions, pre-determined temporal delays were set between the voluntary key press and the onset of the first capital letter (0–50, 100–150, 200–250, 300–350, 400–450, 600–650, or 800–850 ms); respectively, these correspond to 0, 2, 4, 6, 8, 12, or 16 items between the key press and the first letter. These conditions are labeled 0-, 100-, 200-, 300-, 400-, 600-, or 800- ms condition, respectively. In each trial in each self-condition block, the frame at which participants pressed the key was recorded. In a block of the automatic condition, the first letter appeared automatically at the next frame as that recorded during the immediately preceding self-condition block. This procedure enabled us to minimize the variance in the number of items preceding the first letter between the self- and the automatic-condition blocks for each participant (Kihara and Kawahara, 2012). The order of the trials was randomized.

At the beginning of each block, an instruction relating to key press was displayed on the screen. Each trial began with a hash mark (#) presented for 1000–3000 ms to assist in fixation, followed by a stream of digits. Under the self conditions, participants were asked to voluntarily press the space bar once during the digit stream, within 10 s of the start of the RSVP stream in order to change the stream from digits to letters. When participants pressed the space bar, the RSVP changed with a temporal delay that depended on the conditions. Under the 0-ms delay condition, the first item of the letter stream was presented within 50 ms of the key press, i.e., the first target letter, T1, appeared immediately after the key press that quickly followed

a digit item. When participants failed to press the key, the first letter appeared automatically (i.e., the 201st item was the first letter of the letter stream in this case). Under the automatic condition, participants were instructed to refrain from pressing the space bar because the first letter would appear automatically. Once all the items in a stream were presented, participants identified the first four letters by pressing the corresponding keys. Therefore, the first four letters were designed as targets (i.e., T1–T4). A warning message was presented when participants failed to refrain from/perform a response under the automatic/self conditions. Participants were allowed to report the four targets at their own pace.

Data Analyses

In this experiment, trials in which participants failed to refrain/perform a key press response (0.4% of the total trials) or trials on which the recording of pupil data failed (14.3% of the total trials) were excluded from subsequent analyses. On average, in the self conditions, participants pressed the space bar 1,315 ms (SEM = 110) after the onset of the RSVP stream. We retrospectively counted each letter item reported as one of the four targets regardless of reporting order. Note that the reports of the four targets that can be regarded as correct identifications are indicated by circles in **Figure 2**; therefore, reports of the fifth and later items are regarded as wrong identifications and plotted in this figure as without circles. For example, if a participant reported first, second, fifth, and sixth letters, the first two reports were coded as correct, whereas the latter two were coded as incorrect responses.

Pupillary responses were computed as the percentage increase in pupil area compared with baseline over 100 ms before the onset of the digit stream in each trial (i.e., during presentation of the hash mark). The pre-action period (from the start of an RSVP stream to the voluntary key press) was varied in each trial. In this study, there were not enough trials involving the pre-action period of more than 1000 ms. Thus, to obtain a reliable estimate of ideal pupillary responses relative to the pre-action period, we excluded pupillary data more than 1000 ms before the onset of the key press in the self conditions and the onset of the first letter in the automatic condition, which was based on the immediately preceding self-condition block.

Tukey HSD tests were used as *post hoc* comparisons (alpha-level = 0.05). Pearson correlation coefficients (*r*) were computed to estimate the linear correlation of the behavioral and pupillary data. The Smirnov–Grubbs' test was used for evaluating outliers.

Results

The baselines of the pupillary response were not significantly different among the eight letter-onset conditions (i.e., seven self conditions and one automatic condition), confirmed by a one-way repeated-measures analysis of variance (ANOVA), $F(7,189) = 0.90$, n.s., $\eta_p^2 = 0.03$. **Figure 2** shows the rates of responses reported (as percentages) for items as a function of temporal delay from the voluntary key press to the onset of each item in the eight letter-onset conditions. Because we

define targets as the first four letters in the RSVP sequence, the first four data points (with circles) in **Figure 2** indicate correct identification rates of four targets, whereas the fifth and later data points (without circles) indicate the rates of reports of the fifth (or later) items. A two-way repeated-measures ANOVA was conducted on reporting rates with the temporal position of the first four letters and the letter-onset as factors. This analysis yielded significant main effects of the temporal position of the first four letters, $F(3,81) = 63.29$, $p < 0.001$, $\eta_p^2 = 0.70$, and the letter-onset, $F(7,189) = 6.22$, $p < 0.001$, $\eta_p^2 = 0.19$. *Post hoc* comparisons revealed that the accuracy for the second target was significantly higher than for other targets. These results indicate that the peak of transient attention developed at around 100 ms after the onset of the first target (Weichselgartner and Sperling, 1987). *Post hoc* comparisons also showed that the mean accuracy for the four targets in the 600- and 800-ms conditions were significantly higher than those in the other letter-onset conditions, suggesting that longer temporal delay between the action and the letter-onset facilitated report of the first four letters. Importantly, the interaction between the letter-onset factor and the temporal position of the first four letters factor was also significant, $F(21,567) = 5.20$, $p < 0.001$, $\eta_p^2 = 0.16$. Follow up analyses of simple main effects did not show significant effects of the first, $F(7,189) = 1.15$, n.s., $\eta_p^2 = 0.04$, third, $F(7,189) = 1.76$, n.s., $\eta_p^2 = 0.06$, or fourth targets, $F(7,189) = 0.73$, n.s., $\eta_p^2 = 0.03$. However, the second target yielded a significant main effect, $F(7,189) = 3.99$, $p < 0.001$, $\eta_p^2 = 0.13$. *Post hoc* comparisons revealed that the accuracy of second-target identification was higher for the 600- and 800-ms conditions than the 100-ms condition; accuracy was also higher in the 800-ms condition than in either the 0-ms or automatic conditions. Thus, it appears that a facilitatory effect of transient attention occurred at least 800 ms after the voluntary action.

Figure 3 shows time courses of the observed pupillary dilation, where time is locked to the onset of the first target. Each function represents a grand mean of averages of individual participants. Under all the conditions except the automatic condition, cubic function-like changes were observed: the pupil diameters rapidly increased until about 450 ms after the key press, then leveled off or decreased slightly for a while, and then increased again. It is reasonable to assume that this pattern of pupil changes consists of two components that reflect different cognitive processes. The first component represents a sharper, transient rise associated with the voluntary triggering of the targets, whereas the second component indicates a gradual increase after target onset. The first component is obviously associated with the voluntary key press. On the other hand, the second component was also evident in the automatic condition. Thus, the second component does not depend on voluntary action, but rather on the involvement of the RSVP task in which participants have to memorize and report the visual targets at the end of the RSVP stream. It has been shown that the increase in pupil dilation after the onset of to-be-reported items reflects memory load (Kahneman and Beatty, 1966; Peavler, 1974; Beatty, 1982). Therefore, it is highly likely that the second component of pupil dilation in the present study reflects the memory load associated with the four memorized targets, which were to be reported after the
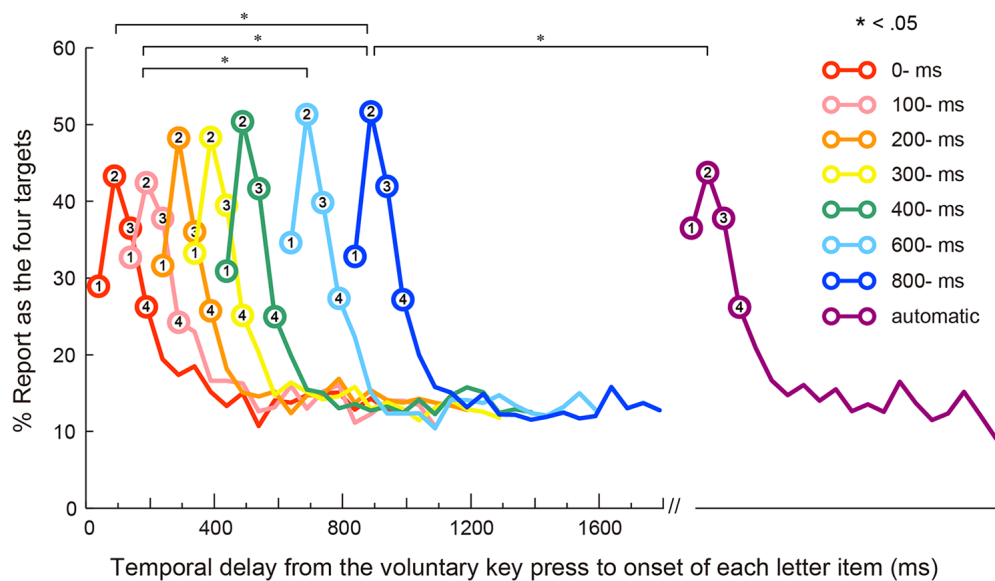
FIGURE 2 | Behavioral result (significant differences in the report rates of the second target are marked with asterisks: *p < 0.05). Mean reporting rates of the four targets as a function of temporal delay from the voluntary key press to the onset of each item in each condition. The first four data points (with circles) of each letter-onset condition represent percentages of target reporting. The number on each circle represents each target. The fifth and later data points (without circles) represent percentages of wrong item reporting. (Note: the letter stream was presented automatically without key press in the automatic condition). The differences between the letter-onset conditions are reported for only the second target.
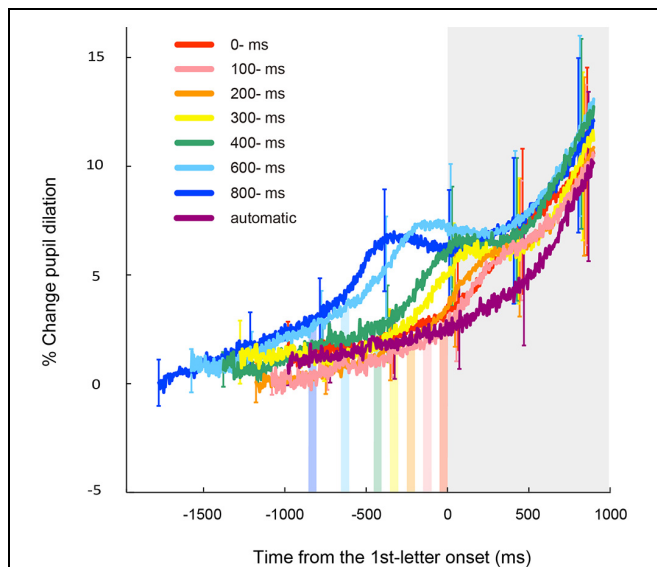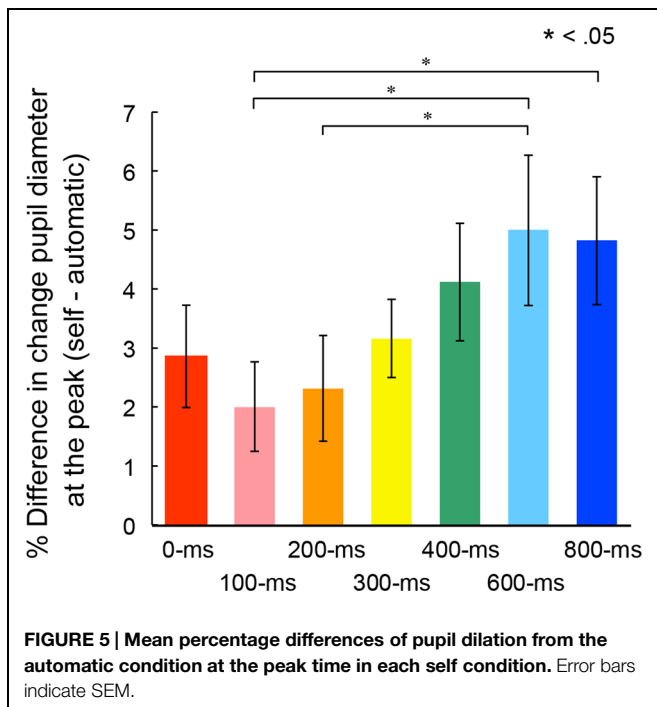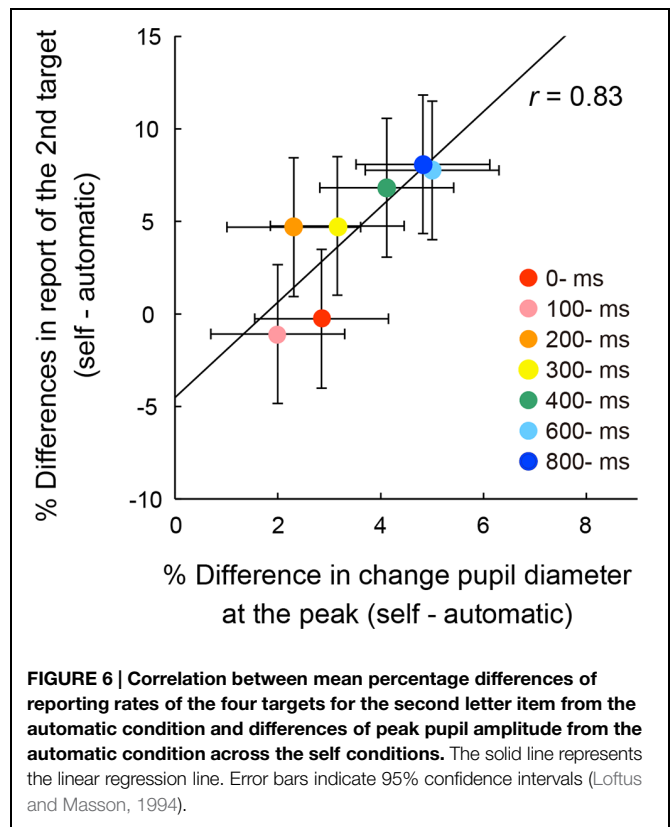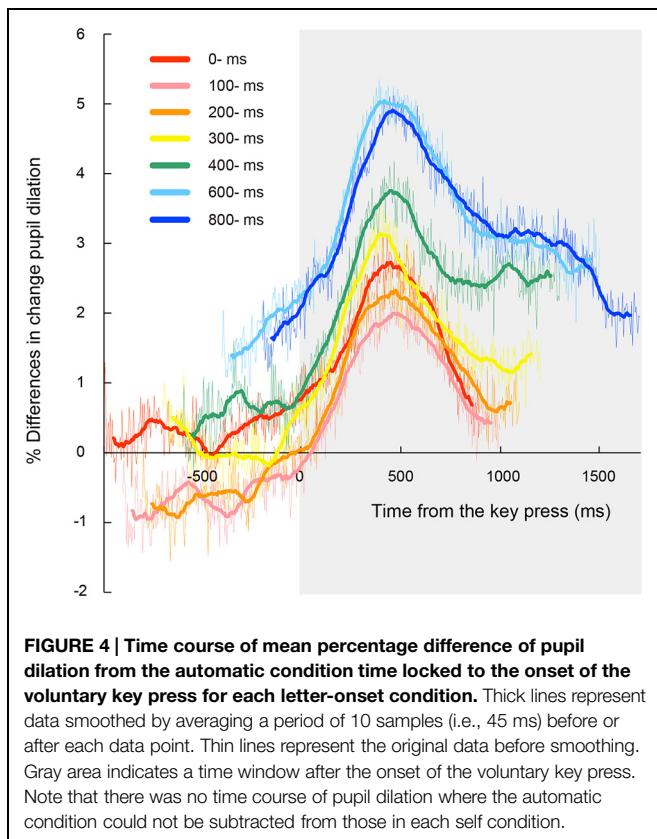

FIGURE 3 | Results of the time course of mean pupil dilation from the baseline time locked to the onset of the first target for each letter-onset condition. Each function represents a grand mean of averages of individual participants and is plotted in different colors. Vertical color bars indicate the time window (i.e., the frame duration of 50 ms) in which the voluntary key presses could be executed. Gray area indicates a time window after the onset of the first letter. Error bars indicate 95% confidence intervals.

letter stream. We isolated the pupillary responses related to the voluntary triggering by subtracting the pupillary responses in the automatic condition from those in self conditions, and

then, sorted the subtracted data based on the onset of the voluntary key press (thin lines in **Figure 4**). To clarify each peak, the time courses of pupil dilation were smoothed by averaging a period of 10 samples (i.e., 45 ms) before and after each data point (thick lines in **Figure 4**). The data of **Figure 4** clearly demonstrate that the voluntary action induced a transient increase in pupil dilation, which peaked at around 450 ms after the action (446, 459, 468, 419, 441, 414, and 455 ms for the 0-, 100-, 200-, 300-, 400-, 600-, and 800-ms delay conditions, respectively).

Figure 5 shows mean percentage differences of pupil dilation in the self conditions relative to dilation levels in the automatic condition for the peak time in each self condition. A one-way ANOVA conducted on the mean peak amplitudes revealed a main effect of the self conditions, $F(6,162) = 3.37$, $p < 0.01$, $\eta_p^2 = 0.11$. *Post hoc* comparisons revealed that peak amplitudes were significantly larger in the 600-ms condition than those in the 100- and 200-ms conditions. The peak amplitude in the 800-ms conditions was also larger than those in the 100-ms condition. These results suggest that the peak in transient pupillary dilation occurred when the targets appeared 600- or 800-ms after the voluntary action.

To examine the relationship between pupil response and action-triggered attentional facilitation, a correlation coefficient was computed between peak pupil dilation and performance of the second target reporting (**Figure 6**). The accuracy for second target responses was calculated by subtracting the report rate for the second targets in the automatic condition from those in each self condition [We used the report rate of the second target rather than the first, as the initial target could reflect inhibition

FIGURE 4 | Time course of mean percentage difference of pupil dilation from the automatic condition time locked to the onset of the voluntary key press for each letter-onset condition. Thick lines represent data smoothed by averaging a period of 10 samples (i.e., 45 ms) before or after each data point. Thin lines represent the original data before smoothing. Gray area indicates a time window after the onset of the voluntary key press. Note that there was no time course of pupil dilation where the automatic condition could not be subtracted from those in each self condition.

FIGURE 6 | Correlation between mean percentage differences of reporting rates of the four targets for the second letter item from the automatic condition and differences of peak pupil amplitude from the automatic condition across the self conditions. The solid line represents the linear regression line. Error bars indicate 95% confidence intervals (Loftus and Masson, 1994).

FIGURE 5 | Mean percentage differences of pupil dilation from the automatic condition at the peak time in each self condition. Error bars indicate SEM.
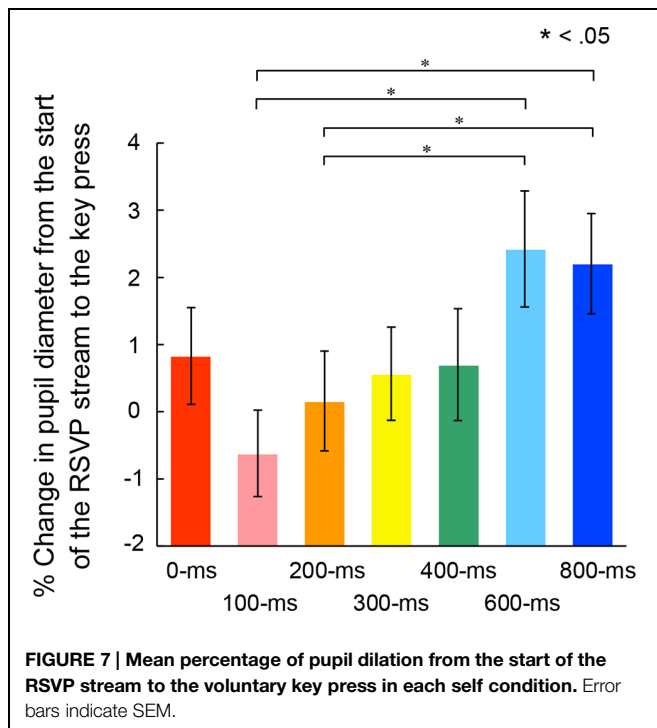
of leading non-targets (Kawahara and Enns, 2009)]. No outliers were identified by the Smirnov–Grubbs test. There was a strong and significant correlation between the peak values and the rates

of reporting for the second target ($r = 0.83$, $p < 0.05$, 95% CI [0.21, 0.98]). It is likely that the transient pupillary dilation was associated with the facilitation of transient attention elicited by the voluntary triggering of visual targets. The peak of pupillary dilation and the attentional facilitation increased as the temporal delay increased.

We examined the possibility that the ratio of pupillary dilations began to differ before the voluntary triggering of the targets among the self conditions. In this case, there should be significant differences of the enlargement ratio of pupil diameter at the voluntary key press. **Figure 7** shows the mean enlargement ratio up to the action. A one-way ANOVA indicated a significant main effect of self conditions, $F(6,162) = 5.21$, $p < 0.001$, $\eta_p^2 = 0.16$. *Post hoc* comparisons revealed that the ratio was significantly larger in the 600- and 800-ms conditions than those in the 100- and 200-ms conditions. In addition, there were significant correlations between the pupil enlargement ratio and the peak pupil amplitude, $r = 0.91$, $p < 0.05$, 95% CI [0.50, 0.99] (**Figure 8**, left panel) and between the pupil enlargement ratio before the key press and the differences of the rates of second target reporting, $r = 0.76$, $p < 0.05$, 95% CI [0.00, 0.96] (**Figure 8**, right panel). Thus, it is reasonable to interpret the observed differences in peak amplitudes of pupil dilation as due mainly to the enlargement of pupil size prior to the voluntary key press.

In summary, the results of behavioral and pupil data indicate that the transient increase in pupil diameter peaked at about 450 ms after a participant's voluntary action, and it was closely associated with the facilitatory effect of transient attention to
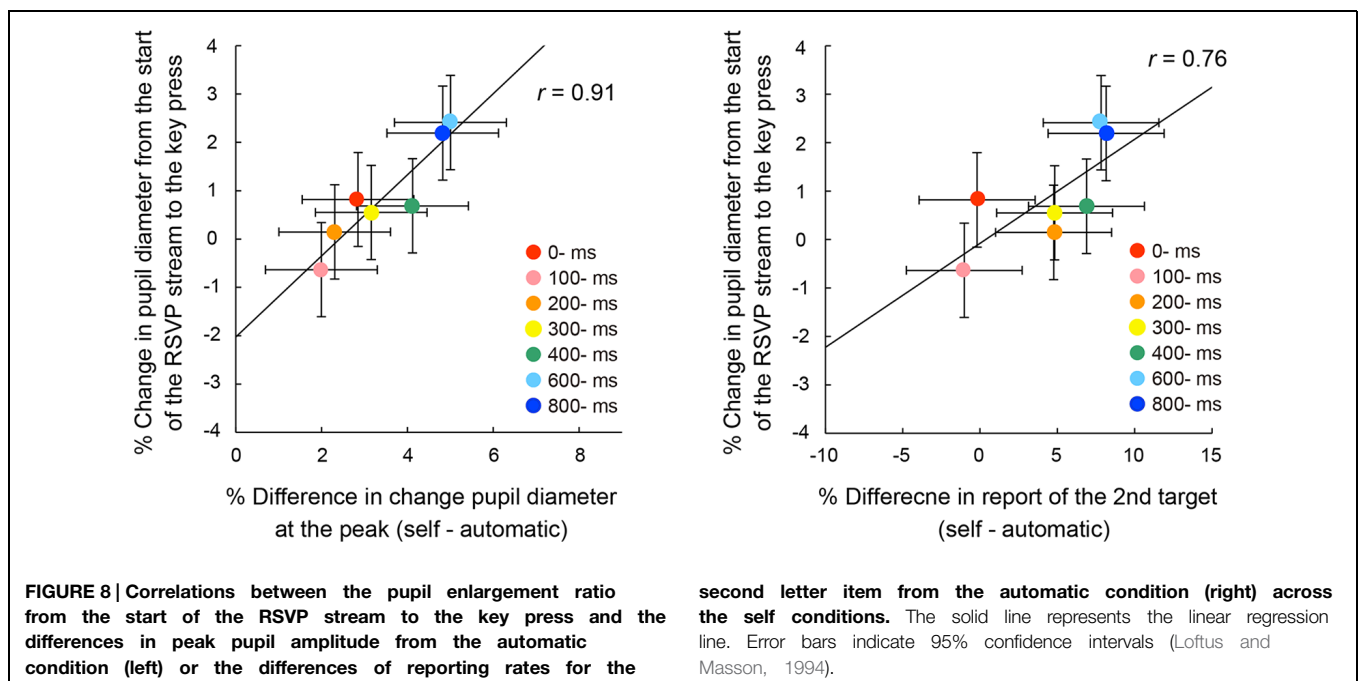
FIGURE 7 | Mean percentage of pupil dilation from the start of the RSVP stream to the voluntary key press in each self condition. Error bars indicate SEM.

visual targets triggered by the action. This transient pupil dilation depended on pupil size at the time of the voluntary action.

Discussion

The present results demonstrate that the LC-NA system mediates the action-induced facilitation of transient attention as indexed by pupil diameter. In this study, identification accuracy of the second target, embedded in an RSVP stream, was higher than the first, third, and fourth targets, thus indicating the occurrence of transient attention (Weichselgartner and Sperling, 1987). Importantly, identification accuracy of the second target improved when target onsets were triggered by a voluntary key press, suggesting the facilitation of transient attention. This facilitation is observed when the target letters were presented about 600–800 ms after the voluntary action. In addition, the reporting rates of targets were significantly correlated with the peak of pupil dilation at around 450 ms after the voluntary key press across the self conditions. Because pupil dilation reflects activation of the LC-NA system (e.g., Laeng et al., 2012), this finding implies that this system is involved in facilitation of the transient attention driven by action-triggered targets. Thus, we suggest that the LC-NA system leads to the facilitation of transient attention for targets triggered by the action.

We found that the pupil dilation began prior to the action. Previous studies have demonstrated similar results of pupil dilation before a motor response (e.g., Einhäuser et al., 2010; Privitera et al., 2010; Smallwood et al., 2011). This component of the transient pupillary response appears to be a precursor of future voluntary movements (Hupé et al., 2009). Our results also demonstrated that the size of the pupil at the action onset was affected by the temporal delay between the key press and the onset of subsequent target. In this task, participants had to maintain attention to the RSVP stream from the action to the onset of the targets. Accordingly, attentional load would be higher when the maintenance period was longer. It is likely that participants pressed the key in the 600- or 800-ms SOA conditions after sufficient attention was developed as to maintain attention until the appearance of the targets. Task-related decision processes

FIGURE 8 | Correlations between the pupil enlargement ratio from the start of the RSVP stream to the key press and the differences in peak pupil amplitude from the automatic condition (left) or the differences of reporting rates for the second letter item from the automatic condition (right) across the self conditions. The solid line represents the linear regression line. Error bars indicate 95% confidence intervals (Loftus and Masson, 1994).

and motor responses are associated with the activation of the LC-NA system (Aston-Jones and Cohen, 2005). Therefore, we believe that the LC neuronal responses are elicited prior to the voluntary action if participants are adequately prepared for the outcomes of these actions, and the facilitation of the transient attention depends on the activation of the LC-NA system related to the motor decisions.

Under the present testing conditions, both the identification accuracies of the second target and the transient increases in pupil diameter were highest when the temporal delay between the key press to the onset of the targets was about 600–800 ms. It is possible that the hidden peaks of these two variables could be observed if the temporal delay were more than 800 ms, although this possibility is questionable because both 600 and 800 ms conditions yielded very similar results. Of course it would be interesting to investigate what length of temporal delay between voluntary action and onset of visual stimulus is optimal, i.e., namely most effective for the facilitation of transient attention. However, this issue is outside of the scope of the present article, which attempts to clarify the relationship between the activation of the LC-NA system and the facilitation of transient attention induced by the voluntary action.

Our findings could possibly be generalized to daily activities where visual events are triggered by a voluntary action. As noted in the introduction, an attentional blink study has demonstrated that it is possible to facilitate attentional processing of voluntary-triggered stimuli embedded in an RSVP stream (Kihara and Kawahara, 2012). Other behavioral studies also suggest a similar facilitation effect. For example, the flash-lag effect, in which a flash is perceived to lag behind a moving object even if both are presented physically aligned (Nijhawan, 2002), has been shown to be reduced when the onset of the flash was triggered by a key press (López-Moliner and Linares, 2006). Temporal orienting of attention plays an important role in the flash-lag effect (Baldo and Klein, 1995; Murakami, 2001; Shioiri et al., 2010; Ichikawa and Masakura, 2013; see also Whitney, 2002, for a review). Interestingly, it has been suggested that the LC-NA system is involved in transient attentional modulation for the flash-lag effect (Bachmann, 2010), as well as for the attentional blink (Nieuwenhuis et al., 2005). Thus, the contribution of the LC-NA system to the facilitation of transient attention to action-triggered visual stimuli is not necessarily limited to RSVP tasks.

It is notable that target accuracy was highest for the second target item, regardless of the temporal delays between the voluntary action and target onset. This pattern of results is frequently observed when people must report multiple items under very short SOAs (Potter et al., 2002; Kawahara and

Enns, 2009). Although Weichselgartner and Sperling's (1987) study showed that the first critical item was reported most frequently, the apparent inconsistency between the present and previous studies may be due to the difference in RSVP rate. Weichselgartner and Sperling (1987) used a presentation rate of 10–12.5 items per second, whereas we adopted 20 items per second. The more rapid rates used in the current task allowed us, to demonstrate the latency of transient attention triggered by target onset. Weichselgartner and Sperling (1987) also reported that the distribution of reportability was bimodal, revealing the existence of both transient and sustained attention. In contrast, the present results yielded a unimodal distribution reflecting transient attention. We assume that the different distributions can be also explained by the difference in RSVP rate. In Weichselgartner and Sperling's (1987) results, accuracy was highest for the first critical item and next highest for the second item; given RSVP rates, this implies that the transient attention would continue for about 200 ms after the onset of the first target. In this case, during this time window of transient attention, four targets were presented in our task, which used a double-speed RSVP. Therefore, the first three or four targets could be identified as the to-be-identified four targets relatively easily, thus obscuring the second peak of attentional enhancement observed in Weichselgartner and Sperling (1987).

Conclusion

We investigated the relationship between the LC-NA system and the facilitation of transient attention to visual stimuli whose onset was triggered by a voluntary action by measuring changes in pupil diameter, which reflects the levels of NA released from LC. We found that the reporting rate of a second letter was closely associated with pupil dilation with a peak at around 450 ms after the voluntary action. The peak of pupil dilation depended on the pupil size at the time of the key press, suggesting that the activation of the LC-NA system related to the motor decision contributes to the facilitation of the transient attention. To our knowledge, this is first study to demonstrate that LC-NA system-mediated pupil dilation is related to the facilitation of transient attention driven by action-induced stimuli.

Acknowledgment

References

Aston-Jones, G., and Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. *Annu. Rev. Neurosci.* 28, 403–450. doi: 10.1146/annurev.neuro.28.061604.135709

Aston-Jones, G., Rajkowski, J., Kubiak, P., and Alexinsky, T. (1994). Locus coeruleus neurons in monkey are selectively activated by attended cues in a vigilance task. *J. Neurosci.* 14, 4467–4480.

Bachmann, T. (2010). "Priming and retouch in flash-lag and other phenomena of the streaming perceptual input," in *Space and Time in Perception and Action*, eds R. Nijhawan and B. Khurana (Cambridge: Cambridge University Press), 536–557.

Baldo, M. V., and Klein, S. A. (1995). Extrapolation or attention shift? *Nature* 378, 565–566. doi: 10.1038/378566a0

Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychol. Bull.* 91, 276–292. doi: 10.1037/0033-2909.91.2.276

Berridge, C. W., and Waterhouse, B. D. (2003). The locus coeruleus-noradrenergic system: modulation of behavioral state and state-dependent cognitive processes. *Brain Res. Rev.* 42, 33–84. doi: 10.1016/S0165-0173(03)00143-7

Bowman, H., and Wyble, B. (2007). The simultaneous type, serial token model of temporal attention and working memory. *Psychol. Rev.* 114, 38–70. doi: 10.1037/0033-295X.114.1.38

Brainard, D. H. (1997). The psychophysics toolbox. *Spat. Vis.* 10, 433–436. doi: 10.1163/156856897x00357

Clayton, E. C., Rajkowski, J., Cohen, J. D., and Aston-Jones, G. (2004). Phasic activation of monkey locus ceruleus neurons by simple decisions in a forced-choice task. *J. Neurosci.* 24, 9914–9920. doi: 10.1523/JNEUROSCI.2446-04.2004

Corbetta, M., Patel, G., and Shulman, G. L. (2008). The reorienting system of the human brain: from environment to theory of mind. *Neuron* 58, 306–324. doi: 10.1016/j.neuron.2008.04.017

Cowan, N. (2001). The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behav. Brain Sci.* 24, 87–114. doi: 10.1017/S0140525X01003922

Einhäuser, W., Koch, C., and Carter, O. L. (2010). Pupil dilation betrays the timing of decisions. *Front. Hum. Neurosci.* 4:18. doi: 10.3389/fnhum.2010.00018

Eldar, E., Cohen, J. D., and Niv, Y. (2013). The effects of neural gain on attention and learning. *Nat. Neurosci.* 16, 1146–1153. doi: 10.1038/nn.3428

Foote, S. L., and Morrison, J. H. (1987). Extrathalamic modulation of cortical function. *Annu. Rev. Neurosci.* 10, 67–95. doi: 10.1146/annurev.ne.10.030187.000435

Hommel, B., Müsseler, J., Aschersleben, G., and Prinz, W. (2001). The Theory of Event Coding (TEC): a framework for perception and action planning. *Behav. Brain Sci.* 24, 849–878. doi: 10.1017/S0140525X01000103

Hupé, J. M., Lamirel, C., and Lorenceau, J. (2009). Pupil dynamics during bistable motion perception. *J. Vis.* 9, 1–19. doi: 10.1167/9.7.10

Ichikawa, M., and Masakura, Y. (2013). Effects of consciousness and consistency in manual control of visual stimulus on reduction of the flash-lag effect for luminance change. *Front. Psychol.* 4:120. doi: 10.3389/fpsyg.2013.00120

Kahneman, D., and Beatty, J. (1966). Pupil diameter and load on memory. *Science* 154, 1583–1585. doi: 10.1126/science.154.3756.1583

Kawahara, J., and Enns, J. T. (2009). Selection difficulty and interitem competition are independent factors in rapid visual stream perception. *J. Exp. Psychol. Hum. Percept. Perform.* 35, 146–158. doi: 10.1037/a0013164

Kihara, K., and Kawahara, J. (2011). Voluntary production of visual items modulates transient attention twice. *J. Vis.* 11, 121. doi: 10.1167/11.11.121

Kihara, K., and Kawahara, J. I. (2012). Voluntary triggering of the first target attenuates the attentional blink. *Atten. Percept. Psychophys.* 74, 312–321. doi: 10.3758/s13414-011-0233-4

Koss, M. C. (1986). Pupillary dilation as an index of central nervous system alpha 2-adrenoceptor activation. *J. Pharmacol. Methods* 15, 1–19. doi: 10.1016/0160-5402(86)90002-1

Laeng, B., Sirois, S., and Gredeback, G. (2012). Pupillometry: a window to the preconscious? *Perspect. Psychol. Sci.* 7, 18–27. doi: 10.1177/1745691611427305

Loftus, G. R., and Masson, M. E. J. (1994). Using confidence intervals in within-subject designs. *Psychon. Bull. Rev.* 1, 476–490. doi: 10.3758/Bf03210951

López-Moliner, J., and Linares, D. (2006). The flash-lag effect is reduced when the flash is perceived as a sensory consequence of our action. *Vision Res.* 46, 2122–2129. doi: 10.1016/j.visres.2005.11.016

Martens, S., Munneke, J., Smid, H., and Johnson, A. (2006). Quick minds don't blink: electrophysiological correlates of individual differences in attentional selection. *J. Cogn. Neurosci.* 18, 1423–1438. doi: 10.1162/jocn.2006.18.9.1423

Martens, S., and Wyble, B. (2010). The attentional blink: past, present, and future of a blind spot in perceptual awareness. *Neurosci. Biobehav. Rev.* 34, 947–957. doi: 10.1016/j.neubiorev.2009.12.005

Miller, G. A. (1956). The magical number seven plus or minus two: some limits on our capacity for processing information. *Psychol. Rev.* 63, 81–97. doi: 10.1037/h0043158

Müller, H. J., and Rabbitt, P. M. (1989). Reflexive and voluntary orienting of visual attention: time course of activation and resistance to interruption. *J. Exp. Psychol. Hum. Percept. Perform.* 15, 315–330. doi: 10.1037/0096-1523.15.2.315

Murakami, I. (2001). The flash-lag effect as a spatiotemporal correlation structure. *J. Vis.* 1, 126–136. doi: 10.1167/1.2.6

Murphy, P. R., Robertson, I. H., Balsters, J. H., and O'connell R, G. (2011). Pupillometry and P3 index the locus coeruleus-noradrenergic arousal function in humans. *Psychophysiology* 48, 1532–1543. doi: 10.1111/j.1469-8986.2011.01226.x

Nakayama, K., and Mackeben, M. (1989). Sustained and transient components of focal visual attention. *Vision Res.* 29, 1631–1647. doi: 10.1016/0042-6989(89)90144-2

Nieuwenhuis, S., Gilzenrat, M. S., Holmes, B. D., and Cohen, J. D. (2005). The role of the locus coeruleus in mediating the attentional blink: a neurocomputational theory. *J. Exp. Psychol. Gen.* 134, 291–307. doi: 10.1037/0096-3445.134.3.291

Nijhawan, R. (2002). Neural delays, visual motion and the flash-lag effect. *Trends Cogn. Sci.* 6, 387–393. doi: 10.1016/S1364-6613(02)01963-0

Olivers, C. N. L., and Meeter, M. (2008). A boost and bounce theory of temporal attention. *Psychol. Rev.* 115, 836–863. doi: 10.1037/a0013395

Olivers, C. N. L., Van Der Stigchel, S., and Hulleman, J. (2007). Spreading the sparing: against a limited-capacity account of the attentional blink. *Psychol. Res.* 71, 126–139. doi: 10.1007/s00426-005-0029-z

Peavler, W. S. (1974). Pupil size, information overload, and performance differences. *Psychophysiology* 11, 559–566. doi: 10.1111/j.1469-8986.1974.tb01114.x

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat. Vis.* 10, 437–442. doi: 10.1163/156856897X00366

Potter, M. C., Staub, A., and O'connor, D. H. (2002). The time course of competition for attention: attention is initially labile. *J. Exp. Psychol. Hum. Percept. Perform.* 28, 1149–1162. doi: 10.1037//0096-1523.28.5.1149

Privitera, C. M., Renninger, L. W., Carney, T., Klein, S., and Aguilar, M. (2010). Pupil dilation during visual target detection. *J. Vis.* 10, 1–14. doi: 10.1167/10.10.3

Rajkowski, J., Kubiak, P., and Aston-Jones, G. (1993). Correlations between locus coeruleus (LC) neural activity, pupil diameter and behavior in monkey support a role of LC in attention. *Soc. Neurosci. Abstr.* 19, 974.

Raymond, J. E., Shapiro, K. L., and Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: an attentional blink? *J. Exp. Psychol. Hum. Percept. Perform.* 18, 849–860. doi: 10.1037//0096-1523.18.3.849

Samuels, E. R., and Szabadi, E. (2008). Functional neuroanatomy of the noradrenergic locus coeruleus: its roles in the regulation of arousal and autonomic function Part II: physiological and pharmacological manipulations and pathological alterations of locus coeruleus activity in humans. *Curr. Neuropharmacol.* 6, 254–285. doi: 10.2174/157015908785777193

Sara, S. J., and Bouret, S. (2012). Orienting and reorienting: the locus coeruleus mediates cognition through arousal. *Neuron* 76, 130–141. doi: 10.1016/j.neuron.2012.09.011

Shioiri, S., Yamamoto, K., Oshida, H., Matsubara, K., and Yaguchi, H. (2010). Measuring attention using flash-lag effect. *J. Vis.* 10, 1–13. doi: 10.1167/10.10.10

Smallwood, J., Brown, K. S., Tipper, C., Giesbrecht, B., Franklin, M. S., Mrazek, M. D., et al. (2011). Pupillometric evidence for the decoupling of attention from perceptual input during offline thought. *PLoS ONE* 6:e18298. doi: 10.1371/journal.pone.0018298

Stock, A., and Stock, C. (2004). A short history of ideo-motor action. *Psychol. Res.* 68, 176–188. doi: 10.1007/s00426-003-0154-5

Weichselgartner, E., and Sperling, G. (1987). Dynamics of automatic and controlled visual attention. *Science* 238, 778–780. doi: 10.1126/science.3672124

Whitney, D. (2002). The influence of visual motion on perceived position. *Trends Cogn. Sci.* 6, 211–216. doi: 10.1016/S1364-6613(02)01887-9

Zylberberg, A., Oliva, M., and Sigman, M. (2012). Pupil dilation: a fingerprint of temporal selection during the "attentional blink". *Front. Psychol.* 3:316. doi: 10.3389/fpsyg.2012.00316

The self in conflict: actors and agency in the mediated sequential Simon task

Michiel M. Spapé[1]*, Imtiaj Ahmed[1,2], Giulio Jacucci[2] and Niklas Ravaja[1,3,4]

[1] Helsinki Institute for Information Technology HIIT, Aalto University, Espoo, Finland, [2] Department of Computer Science, University of Helsinki, Helsinki, Finland, [3] Department of Social Research, University of Helsinki, Helsinki, Finland, [4] School of Business, Aalto University, Helsinki, Finland

Executive control refers to the ability to withstand interference in order to achieve task goals. The effect of conflict adaptation describes that after experiencing interference, subsequent conflict effects are weaker. However, changes in the source of conflict have been found to disrupt conflict adaptation. Previous studies indicated that this specificity is determined by the degree to which one source causes episodic retrieval of a previous source. A virtual reality version of the Simon task was employed to investigate whether changes in a visual representation of the self would similarly affect conflict adaptation. Participants engaged in a mediated Simon task via 3D "avatar" models that either mirrored the participants' movements, or were presented statically. A retrieval cue was implemented as the identity of the avatar: switching it from a male to a female avatar was expected to disrupt the conflict adaptation effect (CAE). The results show that only in static conditions did the CAE depend on the avatar identity, while in dynamic conditions, changes did not cause disruption. We also explored the effect of conflict and adaptation on the degree of movement made with the task-irrelevant hand and replicated the reaction time pattern. The findings add to earlier studies of source-specific conflict adaptation by showing that a visual representation of the self in action can provide a cue that determines episodic retrieval. Furthermore, the novel paradigm is made openly available to the scientific community and is described in its significance for studies of social cognition, cognitive psychology, and human–computer interaction.

Keywords: cognitive control, conflict adaptation, feature integration, mediated interaction, episodic retrieval

Introduction

Cognitive control refers to the ability to withstand temptation and avoid distraction in order to reach certain goals. This is true for definitions from both social and clinical studies – in which such goals are generally longer term, abstract and self-referencing (Baumeister et al., 2000) – and cognitive science – in which they tend to be short term ("in the next block"), very specific ("press a button as cued by the center of the stimulus, not its flankers") and referencing a specific task designed by the experimenter (here Eriksen and Eriksen, 1974). Despite these differences, cognitive control is commonly portrayed as a kind of limited resource that allows us to handle conflicts and interferences: should the resource run low, we may fail to act quickly or correctly.

This somewhat dualistic characterization of control is reflected in models formalizing conflict and control in terms of models featuring two routes. A stimulus can trigger, quickly or automatically, responses that are typical for our normal functioning: the urge is to deal with

this token stimulus as with any other of its kind. A secondary type of processing works its slow, willful way top–down from a goal level toward the more complex processing of the stimulus. For example, in the popular Stroop task (Stroop, 1935), in which we are asked to respond "green" if the word "red" is written in green, we are almost overwhelmed by the automatic reaction to repeat our well-rehearsed training and "read out loud" the word, rather than mind the coloring. Thus, the conflict is between competing responses of the two routes, while the executive control is supposed to suppress the incorrect response.

Despite the apparent simplicity of dual-route models, they do elegantly account for a more recently found effect called *conflict adaptation*. The effect has also been referred to as Gratton effect, or sequential conflict modulation effect, and refers to the observation that after experiencing one instance of conflict, subsequent conflict becomes easier. The effect seems to extend across diverse conflict tasks, including the Stroop task (Egner and Hirsch, 2005b; Spapé and Hommel, 2008), the Simon task (Simon and Rudell, 1967; Hommel et al., 2004) and the Eriksen Flanker Task (Eriksen and Eriksen, 1974; Gratton et al., 1992). To improve clarity, we shall refer to the conflict adaptation effect (CAE) independently from the specific paradigm in which it is encountered, formulating it as:

$$CAE = (cC - cI) - (iC - iI)$$

In which capital *C*s and *I*s denote currently compatible (congruent, non-conflicting) and incompatible (incongruent, conflicting) trials, whereas lower case *c*s and *i*s refer to preceding (often termed N-1) compatibility and incompatibility. The formula thus quantifies the effect as the reduction of conflict-effects as a function of preceding trials.

Dual-route models of executive control account for the CAE by suggesting that a conflicting trial – the word "red" in green – triggers the recruitment of attentional resources to cope with the response uncertainty (Botvinick et al., 2001). Depending on the preferred model, this would mean for our example either that task-relevant route (the color-response association) is facilitated, or that part of the irrelevant stimulus processing route (the word-response association) is suppressed. The result is more or less the same: if, on a subsequent trial, the word "green" is presented in red, the system should be able to cope with ease: both our enhanced color-route, or our attenuated verbal route leaves us well-prepared for correct action.

However, recent observations suggest dual-route models may not adequately account for localized, or *context dependent* conflict adaptation. For example, if attentional resources are generically recruited after experiencing conflict, one should predict smaller subsequent conflict effects, independent of the task – which is not always the case (Notebaert and Verguts, 2008). Furthermore, even within a task, changing a task-irrelevant feature between two Stroop (Spapé and Hommel, 2008) or Simon (Spapé et al, 2011) displays, critically reduces the CAE. Finally, the outcome of conflict in terms of reward has also been shown to affect the CAE (Van Steenbergen et al., 2009). It seems, then, that a unitary, limited resource type of executive control would fail to account for these observations.

Sequences of conflict, however, involve many more cognitive functions than just executive control. To understand what happens in any kind of task repetitions, it is necessary to take a more detailed look at the specific features involved in sequences of conflict. For one, it has been argued that if conflict changes (i.e., cI and iC sequences), some part of the stimulus or response *must be* different as well, whereas if the conflict does not change (in cC and iI), there is usually a proportion of trials in which the whole stimulus-response scenario is repeated. In other words, priming – rather than cognitive control – was pointed out to be at least partly responsible for the CAE pattern (Mayr et al., 2003).

Further aggravating the situation was the observation by Hommel et al. (2004) who showed that increased errors and reaction latencies observed in cI and iC sequences could be traced back to their constituent features *partly* repeating. Following in the footsteps of Kahneman et al. (1992), they provided evidence that if one scenario (e.g., an arrow left pointing to the left) is similar to a previous representation in that features are repeated (an arrow left pointing to the right), an episodic retrieval effect ensues. This is problematic for two reasons: (1) the repeated feature (the location of the arrow) thus prompts a no longer relevant and indeed conflicting response; and (2) the partial overlap itself may be problematic for the cognitive system (Treisman, 1996; Hommel et al., 2001).

It is thus possible that the workings of episodic retrieval, memory and a type of pattern recognition may account for both the CAE and the context dependency of the CAE. This "stronger" account suggests that the data can fully be accounted for by referring to the "lower-level" functions involved in priming (Mayr et al., 2003), episodic retrieval (Hommel et al., 2004) and contingency learning (Schmidt and Besner, 2008). Thus, there would be very little theoretical need to postulate the extra limited resource to sometimes come to our aid and cognitive control is reduced to an illusory epiphenomenon of free will.

Alternatively, a mechanism featuring episodic retrieval causing conflict adaptation could reconcile "pure control" with context dependency effects. As we have argued before (Spapé and Hommel, 2008, 2014), it is possible that the similarity of situations between two trials may not only retrieve the previous episodes in terms of their constituent features, but also in terms of control parameters. Thus, tasks involving an amount of similarity, because, e.g., a Simon stimulus gradually rotated into its new position, causing updated episodic memory (Spapé and Hommel, 2010) or a voice presenting an auditory Stroop stimulus is repeated (Spapé and Hommel, 2008), may result in conflict adaptation. Conversely, gradually rotating the Simon display to the wrong position or presenting a stimulus in a different tone of voice may interfere with retrieval of executive control (for a similar proposal, see Egner, 2014).

Present Study

The mapping of contingencies of conflict adaptation thus remains important while the debate concerning the status of conflict adaptation continues. The present study was somewhat inspired by the earlier cited observation of the context dependency of the CAE (Spapé and Hommel, 2008). In that study, the

words "high" and "low" were mixed with high and low tones, and participants were asked to judge the pitch of the tones and ignore the words. A type of Stroop effect was observed—participants found it difficult to not imitate the voice—as well as conflict adaptation—the Stroop effect was smaller after incompatible trials. The context dependency was in the voice: although it was entirely irrelevant to the task, changing the voice from one gender to the other caused interference with the CAE.

A visual version of this task was designed for the present study, with one critical change: the degree of ownership over the contextual change. Rather than changing something entirely irrelevant as in the original study, or changing the task itself (Notebaert and Verguts, 2008), we set out to change the degree to which the change was related to the person *involved in* the task. Participants were engaged in the task in two conditions: directly or mediated by a visual representation of themselves, which we will refer to as the "avatar." Similar to the original study, this avatar served as a contextual cue, and could either alternate or repeat between two genders. Although entirely irrelevant to the task, changes in avatar identity should, according to the episodic retrieval account of the CAE, affect the conflict-control pattern. That is, repeating the avatar should act as a cue, prompting retrieval of the preceding trial and possibly its conflict-related aspects. Changes in the identity of the avatar should, conversely, interfere with retrieval and thereby reduce the CAE.

However, to go beyond previous studies related to the context-dependency of the CAE, we investigated whether the relationship between the participant and their virtual identity would have an effect on conflict and control. By using a motion tracking device, we established a sense of agency over the avatar, projecting it as standing in front of the participant and mimicking the participants' gestures. Previous studies used similar techniques in order to manipulate the representation of the self toward the virtual identity (Lenggenhager et al., 2007). In the present experiment, we contrast this "dynamic" condition in which the avatar is displayed as co-acting the participant's gestures, with a "static," control condition in which the avatar did not move.

On the one hand, creating a sense of agency over the avatar by making it respond to the task necessarily increases the degree to which the avatar is task-relevant. Given that conflict-resolution has previously been found to work on task-relevant features (Egner and Hirsch, 2005a), a conflict-control point of view would predict changes in a task-related avatar's identity to be of greater impact than changes in a static, and therefore neutral and irrelevant, picture. On the other hand, however, the degree of agency over the avatar could create the impression that the avatar is "part of" the participant. Thus, a superficial change in the visual appearance of the self-related object should be negated by the sense that it acts as a pointer toward the distal representation: the participant him or herself.

The motion tracking device furthermore enabled us to go a step beyond the traditional reaction times (RTs). Recent studies used single-handed pointing movements (Buetti and Kerzel, 2008) and mouse pointer trajectories (Scherbaum et al., 2010) and analyzed movement trajectories in order to dissociate conflict mechanisms underlying the Simon effect. In these studies, the spatial location of a stimulus was found to cause a shift in movement trajectory toward the stimulus (Buetti and Kerzel, 2009). Here, we explored whether this continuous, "visuomotor" Simon effect (Wiegand and Wascher, 2005) could similarly be observed in a gesture-based, two-handed paradigm. Similar to these studies, we expected the visual location of the stimulus to evoke unintentional movement toward that location. However, in this two-handed study, such movement should occur in the *other* hand, even though it is irrelevant for executing the desired gesture. To our knowledge, there are as yet no studies directly testing the conflict dependency of the CAE on this type of movement trajectory measure, but we expected the pattern of the *irrelevant movement* (IM) to largely follow that of traditional RT.

Materials and Methods

Participants

We partly based the number of participants on similar episodic studies, such as Spapé and Hommel (2008), who observed a sizable effect size of identity switches on conflict control of $\eta_p^2 = 0.56$ with 14 subjects. However, given the unknown, additional factor of avatar animation, and the novel apparatus in use, we ultimately recruited 18 volunteers (seven female). They were 27.1 ± 3.2 years of age and took part in the study in exchange for cinema tickets. Before signing informed consent, they were informed of their rights in accordance with the Declaration of Helsinki. One (female) participant could not complete the study and was removed from further analysis.

Apparatus and Stimuli

The Xbox-360 Kinect (Microsoft, Redmond, WA, USA) is a motion sensing input device that uses a depth camera to track up to six persons and estimate full skeletal tracking information of two persons. Its sensor has a frame rate of 30 Hz, a field of view of $57° \times 43°$, and $27°$ of vertical tilt range, to obtain information for estimating the 3D spatial position of 20 joints for each body. In the study we used it for tracking the position of both hands relative to the torso. Furthermore, we calculated the participant's joint orientation. In the dynamic condition of the present study, the detected joint orientation was projected onto the avatar, giving it participant-avatar congruence in bodily motion.

Figure 1 shows the basic characteristics of the Simon task, which was displayed on a 95.17 cm × 57.10 cm virtual screen which itself was projected on a 254 cm × 142.875 cm Screenline real screen. All task related stimuli – the circles, stars, and fixation crosshair – were 28.55 cm × 28.55 cm. Left and right locations were defined as occurring at, respectively, 28.58 cm left and right from the center of the screen. The 3D character, referred to as the "avatar," was presented at a location below and slightly overlapping the central fixation, as to give the impression that it was standing in between the participant and the virtual screen. It was 25.32 cm × 105.51 cm in size (of which the lower ca. 30 cm not visible) and was of either male or female gender.

Procedure

After reading written instructions, participants witnessed a demonstration of the experiment involving one of the authors
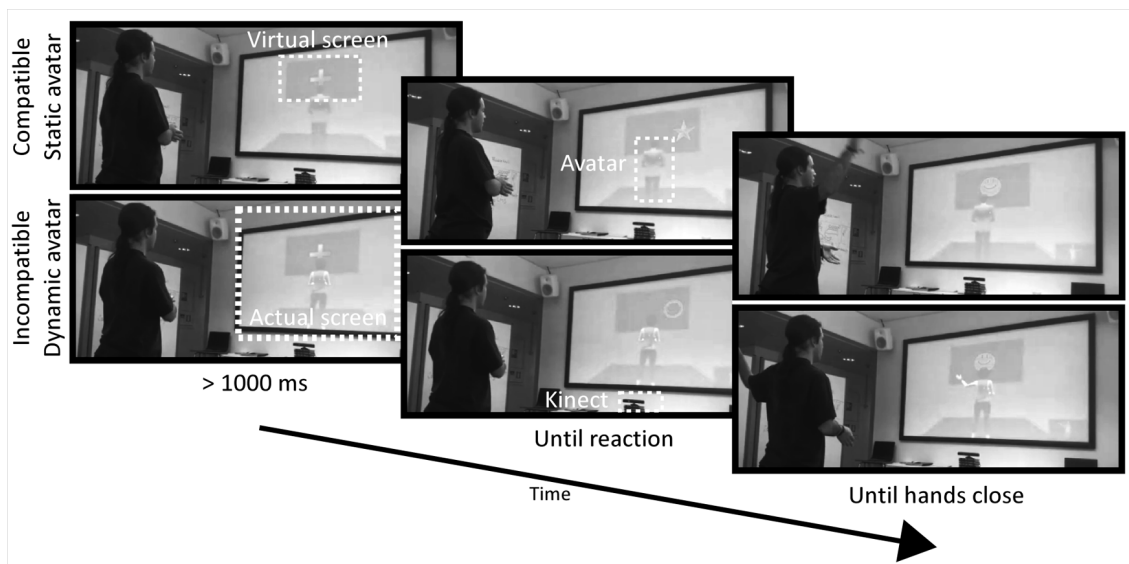
FIGURE 1 | Schematic display of the trial procedure for compatible and incompatible conditions. Participants were instructed to keep their hands together, until the target stimulus (a circle or star) was displayed, prompting a left or right handed action, respectively. Both positive (pictured) and negative feedback was displayed in the first 16 trials, whereas during the rest of the experiment, either negative feedback – following incorrect reactions – or blank virtual screens – following correct actions – were used. The Simon effect refers to the effect of incompatibility between stimulus- and response-location, as is the case in the lower middle panel.

undertaking 16 trials to show the task. Participants were then asked to stand at a distance between 2.5 and 3.5 m from the screen with the arms spread wide, while the instruments were calibrated. If participants had no further questions, they were asked to move their hands together to start the first trial of the experiment.

Every trial started with a fixation crosshair, displayed for ca. 1 s of stable identification of both participant's hands remaining near the center of their body. Then, a star or circle was presented to the left or right of the virtual screen. Participants were instructed to move their left arm left if a circle was shown and their right arm right if a star was shown, irrespective of the location of the stimulus. Movements were detected if the participant moved either hand 20 cm lateral to their shoulders, at which point the star or circle was removed from the screen. Only once the participant moved both their hands back together would the next trial begin. Avatars were presented throughout the experiment as either "static" or "dynamic," the latter case referring to the scenario that the movements of the participants were reflected in the movements of the avatar.

Design and Measurements

The general design of the experiment was based on 2 (locations, left vs. right) $\times$ 2 (shapes requiring left vs. right responses) $\times$ 2 (avatar identities) $\times$ 2 (animations) $\times$ 16 = 256 trials with one block of 128 trials for each type of animation, presented in counter-balanced order with equal numbers of compatible (location = response) and incompatible (location $\neq$ response) trials. The analysis was based on two four-way repeated measures ANOVAs with animation (static vs. dynamic), avatar repetition (vs. alternation), previous compatibility (vs. incompatibility), and current compatibility (vs. incompatibility) as factors. Within each block, a restricted random sampling procedure was used to generate at least 12 occurrences for each design cell.

Two measurements were tested independently: RT and incorrect movement (IM) velocity. The RT was measured as the difference between the onset of the target stimulus (i.e., the circle or star) and the time at which a displacement of either of the participant's hand was detected at least 20 cm relative to the corresponding shoulder. The IM was measured as the peak velocity of the average movement trajectory of the inactive hand prior to the final movement (occurring on average at 601 $\pm$ 25 ms after target onset). The movement of the correct hand was also recorded, but not analyzed, as it is confounded with RT (see **Figure 2**).

Results

The first eight trials as well as the first trial in each block were considered still part of training and removed from analysis. All trials with slow (RT > 1000 ms) or incorrect reactions were also removed, as well as the first trial directly after such scenarios, constituting 9.1 $\pm$ 6.3% of trials.

In repeated measures ANOVAs with animation of the avatar (static vs. dynamic), the repetition of the avatar (repeated vs. alternated), the previous compatibility (vs. incompatibility), and current compatibility (vs. incompatibility) on RT and IM, current compatibility significantly affected both RT, $F(1,15) = 194.64$, MSE = 785.36, $p < 0.001$, $\eta_p^2 = 0.92$, and IM, $F(1,15) = 26.01$, MSE = 26.52, $p < 0.001$, $\eta_p^2 = 0.62$. This suggested a robust Simon effect, with incompatible conditions being associated
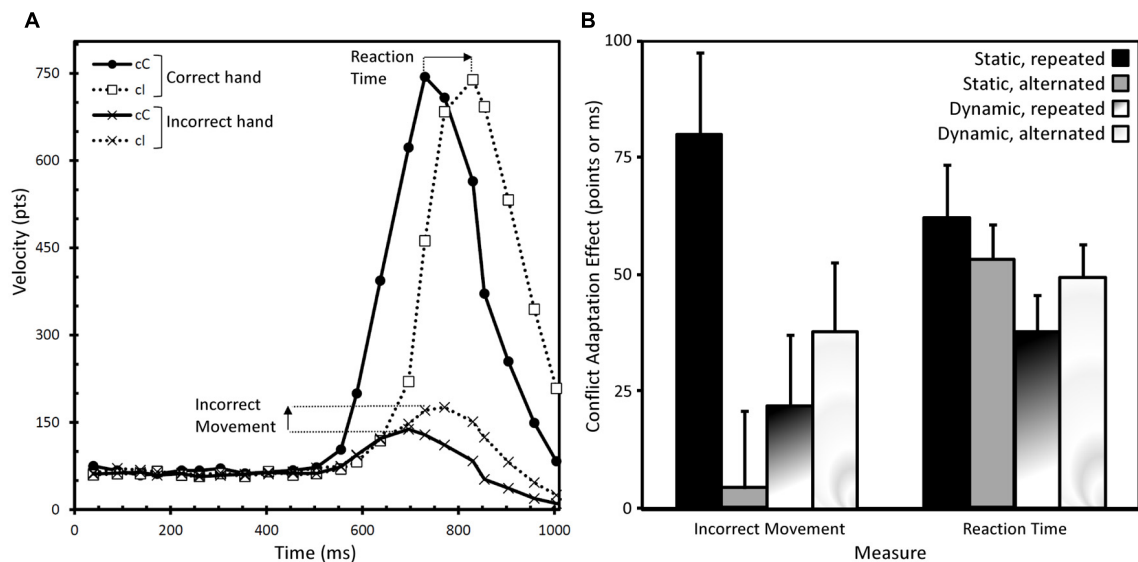
FIGURE 2 | Measurements and results. (A) Grand Averages of the compatible-after compatible (cC) and incompatible-after-compatible (cI) conditions for the correct and incorrect hand over time. The two arrows indicate the difference between the two dependent variables used in **(B)**. **(B)** Conflict adaptation effect (CAE) for each dependent variable and each combination of avatar animation and repetition. Vertical bars indicate one standard error.

with slower RTs (ca. 47 ms) and more IM than compatible ones. Previous compatibility also significantly affected RT, $F(1,16) = 29.31$, MSE = 158.26, $p < 0.001$, $\eta_p^2 = 0.65$, and IM $F(1,16) = 13.10$, MSE = 14.98, $p = 0.002$, $\eta_p^2 = 0.45$, with compatibility in the preceding trial resulting in faster RTs, but less IM.

Neither of the other main effects was significant for RT, ps > 0.59, and IM, ps > 0.20. In general, the IM measure showed a pattern similar to the RT, with interacting variables significantly affecting either both RT and IM, or neither. However, one effect was uniquely observed for one measure: compatibility significantly interacted with avatar identity, $F(1,17) = 4.60$, MSE = 10.70, $p = 0.048$, $\eta_p^2 = 0.22$, for IM only. This indicated that the compatibility effect was larger (C-I = 40.4 pts) after repeated than after alternated (23.3 pts) avatar identities.

Critically, a significant interaction effect between previous and current compatibility was observed for both measures, RT $F(1,15) = 80.31$, MSE = 545.71, $p < 0.001$, $\eta_p^2 = 0.83$; IM $F(1,15) = 13.02$, MSE = 16.88, $p = 0.002$, $\eta_p^2 = 0.45$. This showed a clear replication of a CAE, with the effect of incompatibility being reduced following incompatibility, for both RT (cC – cI = 73 ms, iC – iI = 22 ms) and IM (cC – cI = 49.8 pts, iC – iI = 13.9 pts). Finally, a significant four-way interaction suggested conflict adaptation to be dependent on both the repetition of the avatar, and its animation, RT $F(1,15) = 5.25$, MSE = 84.36, $p = 0.04$, $\eta_p^2 = 0.25$, and IM $F(1,15) = 10.37$, MSE = 8.60, $p = 0.005$, $\eta_p^2 = 0.39$.

To better understand the significant four-way interaction, we calculated the interaction term for each individual combination of avatar animation and avatar repetition. These CAE scores represent the decrease in the conflict effect as a function of preceding trial and are summarized in **Figure 2**. As can be seen

from the figure, a maximal CAE was observed in repeated, static conditions for both RT and IM, indicating a replication of a standard CAE or Gratton effect (Gratton et al., 1992; Botvinick et al., 2001). CAEs were lower during static, alternated trials, with the CAE in IM turning to insignificance (4.15 ± 16.58 pts), replicating previous observations of the context dependency of the CAE. However, this context dependency itself was modulated by the animation of the avatar as, with dynamic conditions, the alternated avatar identities no longer caused a disruption of the CAE.

Discussion

The results show that both the identity of the avatar, and its relation with the participant, affect cognitive performance. In general, participants suffered from a smaller conflict effect after conflict was repeated. Replicating previous studies suggesting conflict adaptation acts locally, or depends critically on irrelevant cues, the CAE was found to be disrupted if the identity of the avatar was changed. In other words, despite the avatar itself being entirely irrelevant to the task, a subtle change in its appearance reduced the CAE. This could be due to the change in cue disrupting recall of the preceding episode, disrupting feature integration and perhaps recall of control-related parameters.

One might imagine, as we sketched in the introduction, that perceiving the avatar as actively mimicking the participant's actions would make it necessarily related to the task, as opposed to, as in the static case, an accidental bystander. Consequently, a change in the mirror image could constitute a particularly disrupting, if not disturbing event: after all, such an imaginary change in self-perception is a classic motif in horror stories (Dietrich, 1992) and a symptom in psychiatry (Maack and

Mullen, 1983). Whether frightful or merely task-relevant, the predicted effect of avatar changes should from this perspective be larger in animated than in static conditions.

However, this prediction clearly did not hold. Conditions in which the avatar was displayed dynamically, with its movements mimicking those of the participant, showed no longer the disruptive effect of identity changes on the CAE. Indeed, if anything, the effect sometimes even seemed to *increase* after a change.

One way to account for this could be in terms of an integration process that makes the avatar similar to "a tool" as held by the participant. In the rubber-hand illusion, seeing an object being stroked and feeling the sensation on the real hand brings about the perception that the virtual object is part of oneself (Botvinick and Cohen, 1998). Here, a virtual persona is likewise presented in synchrony with the participant's actions. By acting consistently in concert with the subject, it is likely that a bi-directional association is formed (Hommel, 1996), between one's own intentions and the behavior carried out by the avatar. Such bidirectional association has recently been shown to elicit a certain unity between model and imitator, as shown by facilitated action execution if a model anticipates imitation rather than counter-imitation (Pfister et al., 2013).

Thus, if perceiving the dynamic avatar results in similar co-representation, the result could be that in the dynamic condition, the avatar is not necessarily an aspect of the task anymore, but an aspect of the agent. This, in turn, should have a critical effect on control in the degree to which the new and the old trial relate: the superficial identity of the avatar may have changed, but it should still point toward the same distal (Hommel, 2009) property. The repetition would then act as an episodic recall cue for the preceding trial, in which the same agent (i.e., the participant him- or herself) was present. In other words, different task-related, whether relevant or irrelevant, features may retrieve preceding, potentially partially overlapping trials, but changes in the avatar still relate to the self-same agent, who was always present in the preceding trial as well.

A competing explanation for the findings could be that the dynamically portrayed avatar made it more difficult to see changes affecting the identity of the avatar. However, this seems to run counter previous studies showing effects on conflict adaptation to remain even with stimulus displays featuring dynamic contextual cues (Spapé and Hommel, 2014). Alternatively, the animation itself was not critical in disturbing the context dependency of the CAE, but the fact that the animation was congruent with the participant's own movement. This form of agency could perhaps counteract the effect by inducing a type of "change-blindness" (Simons and Levin, 1997) to the changes in identity. In the end, however, this forward-interfering account seems presently difficult to distinguish from the earlier, retrieval-based one.

Finally, we would like to discuss some novel aspects of the platform and methodology used in the experiment, as with the publication of this article, we release it as open source, freely available (source[1]) to the academic community. The compressed archive contains source, binaries and a short documentation file (see README.txt inside archive). Notice that, apart from the dynamic and static conditions referred to in the present manuscript, the platform also allows pre-programmed avatar animations with an onset equal to the average RT of the participant. We decided not to use these animations for the present study, as we had no predictions for model-imitator incongruency at the time (but see Pfister et al., 2013), but we could well-imagine this option could be of potential interest to fellow researchers.

The first aspect to note, particularly of interest for studies of conflict control, could be in the use of motion tracking. Although the field remains dominated by simple RTs and 2–4 alternatives forced-choice paradigms, current theoretical models, neuroscience methods and motor control paradigms (Scherbaum et al., 2010; Spapé and Serrien, 2010; Serrien and Spapé, 2011) indicate that focusing on the far end-point of an action – the time at which a button is fully pressed – ignores valuable data. Although previous studies found compatibility affecting response force as well as RT (van der Lubbe et al., 2001), the present study goes further to show the time-course of response conflict in the irrelevant response modality. It is possible that the *other* hand provides a more optimal indicator of conflict than the correct hand, as it is presumably less affected by early control operations that may partially negate the final RT. Of course, previous studies have circumvented the issue by providing measures related to the activation of the irrelevant motor cortex (Valle-Inclán, 1996) and muscles (Hasbroucq et al., 1999). However, the presented IM measure has the advantage of being very directly related to irrelevant response tendency as well as being rather cost-effective in terms of expenses of consumer grade apparatus and the time involved for participants and researchers (no recording preparation or calibration requirements).

The second aspect of the study that merits further discussion is the virtualized design. The experiment in a wider setting may provide a relatively low-cost virtual reality platform for studies of cognition and social identity. Here, we showed effects of changing one's identity, implying that the setup can be a useful tool for the study of social and virtual identity. Social psychological effects, such as social facilitation (Zajonc, 1965) and conformity (Asch, 1951) can be easily tested without relying on confederates by adding extra avatars and operating them remotely (see Blascovich et al., 2002 for an overview of the benefits of immersive virtual environments). Tests of implicit stereotyping and embodied cognition could involve the adjustment of the shape of the avatar to enable identification with various cultural stereotypes. In sum, the study demonstrates that the present design (open source code[1]) may provide an interesting, new way for a variety of researchers and fields of study.

Finally the study blends the fields of executive control and conflict with the study of human–computer interaction (HCI). Given the growing diversity of input techniques and the heterogeneity of user interfaces, basic psychological studies can inform design by taking into account how different interaction

---

[1]www.cognitology.eu/SelfInConflict.html

techniques inflict conflict or provide control. User interfaces, such as employed in the study are increasingly becoming part of everyday consumer products such as game consoles (Harper and Mentis, 2013) and public displays (Kuikkaniemi et al., 2011). This has prompted research in HCI to reconsider embodied interaction with virtual representations (Wilson et al., 2012). The study also demonstrates self-representing avatars may positively contribute to interfaces designed for scenarios with common distraction and a high demand for attentional control. This should motivate further investigation of effects of avatars on various persuasion phenomena on a wide range of different application contexts.

References

Asch, S. E. (1951). "Effects of group pressure upon the modification and distortion of judgments," in *Groups, Leadership, and Men*, ed. H. Guetzkow (Pittsburgh: Carnegie Press), 177–190.

Baumeister, R. F., Muraven, M., and Tice, D. M. (2000). Ego depletion: a resource model of volition, self-regulation, and controlled processing. *Soc. Cogn.* 18, 130–150. doi: 10.1521/soco.2000.18.2.130

Blascovich, J., Loomis, J., Beall, A. C., Swinth, K. R., Hoyt, C. L., and Bailenson, J. N. (2002). Immersive virtual environment technology as a methodological tool for social psychology. *Psychol. Inq.* 13, 103–124. doi: 10.1207/S15327965PLI1302_01

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., and Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychol. Rev.* 108, 624–652. doi: 10.1037/0033-295X.108.3.624

Botvinick, M. M., and Cohen, J. (1998). Rubber hands 'feel' touch that eyes see. *Nature* 391, 756–756. doi: 10.1038/35784

Buetti, S., and Kerzel, D. (2008). Time course of the Simon effect in pointing movements for horizontal, vertical, and acoustic stimuli: evidence for a common mechanism. *Acta Psychol.* 129, 420–428. doi: 10.1016/j.actpsy.2008.09.007

Buetti, S., and Kerzel, D. (2009). Conflicts during response selection affect response programming: reactions toward the source of stimulation. *J. Exp. Psychol. Hum. Percept. Perform.* 35, 816–834. doi: 10.1037/a0011092

Dietrich, B. D. (1992). The ghost of the corpus callosum. *Semiotics* 279–287. doi: 10.5840/cpsem199211

Egner, T. (2014). Creatures of habit (and control): a multi-level learning perspective on the modulation of congruency effects. *Cognition* 5, 1247. doi: 10.3389/fpsyg.2014.01247

Egner, T., and Hirsch, J. (2005a). Cognitive control mechanisms resolve conflict through cortical amplification of task-relevant information. *Nat. Neurosci.* 8, 1784–1790. doi: 10.1038/nn1594

Egner, T., and Hirsch, J. (2005b). The neural correlates and functional integration of cognitive control in a Stroop task. *Neuroimage* 24, 539–547. doi: 10.1016/j.neuroimage.2004.09.007

Eriksen, B. A., and Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Percept. Psychophys.* 16, 143–149. doi: 10.3758/BF03203267

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. doi: 10.1037/0096-3445.121.4.480

Harper, R., and Mentis, H. (2013). "The mocking gaze: the social organization of kinect use," in *Proceedings of the 2013 Conference on Computer Supported Cooperative Work,* New York, NY, 167–180. Available at: http://dl.acm.org/citation.cfm?id = 2441797

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. doi: 10.1017/S0048577299001602

Hommel, B. (1996). The cognitive representation of action: automatic integration of perceived action effects. *Psychol. Res.* 59, 176–186. doi: 10.1007/BF00425832

Hommel, B. (2009). Action control according to TEC (theory of event coding). *Psychol. Res. PRPF* 73, 512–526. doi: 10.1007/s00426-009-0234-2

Hommel, B., Müsseler, J., Aschersleben, G., and Prinz, W. (2001). The Theory of Event Coding (TEC): a framework for perception and action planning. *Behav. Brain Sci.* 24, 849–878. doi: 10.1017/S0140525X01000103

Hommel, B., Proctor, R. W., and Vu, K.-P. L. (2004). A feature-integration account of sequential effects in the Simon task. *Psychol. Res.* 68, 1–17. doi: 10.1007/s00426-003-0132-y

Kahneman, D., Treisman, A., and Gibbs, B. J. (1992). The reviewing of object files: object-specific integration of information. *Cogn. Psychol.* 24, 175–219. doi: 10.1016/0010-0285(92)90007-O

Kuikkaniemi, K., Jacucci, G., Turpeinen, M., Hoggan, E., and Muller, J. (2011). From space to stage: how interactive screens will change urban life. *Computer* 44, 40–47. doi: 10.1109/MC.2011.135

Lenggenhager, B., Tadi, T., Metzinger, T., and Blanke, O. (2007). Video ergo sum: manipulating bodily self-consciousness. *Science* 317, 1096–1099. doi: 10.1126/science.1143439

Maack, L. H., and Mullen, P. E. (1983). The doppelgänger, disintegration and death: a case report. *Psychol. Med.* 13, 651–654. doi: 10.1017/S0033291700048066

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

Notebaert, W., and Verguts, T. (2008). Cognitive control acts locally. *Cognition* 106, 1071–1080. doi: 10.1016/j.cognition.2007.04.011

Pfister, R., Dignath, D., Hommel, B., and Kunde, W. (2013). It takes two to imitate anticipation and imitation in social interaction. *Psychol. Sci.* 24, 2117–2121. doi: 10.1177/0956797613489139

Scherbaum, S., Dshemuchadse, M., Fischer, R., and Goschke, T. (2010). How decisions evolve: the temporal dynamics of action selection. *Cognition* 115, 407–416. doi: 10.1016/j.cognition.2010.02.004

Schmidt, J. R., and Besner, D. (2008). The Stroop effect: why proportion congruent has nothing to do with congruency and everything to do with contingency. *J. Exp. Psychol. Learn. Mem. Cogn.* 34, 514–523. doi: 10.1037/0278-7393.34.3.514

Serrien, D. J., and Spapé, M. M. (2011). Motor awareness and dissociable levels of action representation. *Neurosci. Lett.* 494, 145–149. doi: 10.1016/j.neulet.2011.02.077

Simon, J. R., and Rudell, A. P. (1967). Auditory SR compatibility: the effect of an irrelevant cue on information processing. *J. Appl. Psychol.* 51, 300–304. doi: 10.1037/h0020586

Simons, D. J., and Levin, D. T. (1997). Change blindness. *Trends Cogn. Sci.* 1, 261–267. doi: 10.1016/S1364-6613(97)01080-2

Spapé, M. M., Band, G. P. H., and Hommel, B. (2011). Compatibility-sequence effects in the Simon task reflect episodic retrieval but not conflict adaptation: evidence from LRP and N2. *Biol. Psychol.* 88, 116–123. doi: 10.1016/j.biopsycho.2011.07.001

Spapé, M. M., and Hommel, B. (2008). He said, she said: episodic retrieval induces conflict adaptation in an auditory Stroop task. *Psychon. Bull. Rev.* 15, 1117–1121. doi: 10.3758/PBR.15.6.1117

Spapé, M. M., and Hommel, B. (2010). Actions travel with their objects: evidence for dynamic event files. *Psychol. Res.* 74, 50–58. doi: 10.1007/s00426-008-0219-6

Spapé, M. M., and Hommel, B. (2014). Sequential modulations of the Simon effect depend on episodic retrieval. *Front. Psychol.* 5:855. doi: 10.3389/fpsyg.2014.00855

Spapé, M. M., and Serrien, D. J. (2010). Interregional synchrony of visuomotor tracking: perturbation effects and individual differences. *Behav. Brain Res.* 213, 313–318. doi: 10.1016/j.bbr.2010.05.029

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *J. Exp. Psychol.* 18, 643–662. doi: 10.1037/h0054651

Treisman, A. (1996). The binding problem. *Curr. Opin. Neurobiol.* 6, 171–178. doi: 10.1016/S0959-4388(96)80070-5

Valle-Inclán, F. (1996). The locus of interference in the Simon effect: an ERP study. *Biol. Psychol.* 43, 147–162. doi: 10.1016/0301-0511(95)05181-3

van der Lubbe, R. H., Jaśkowski, P., Wauschkuhn, B., and Verleger, R. (2001). Influence of time pressure in a simple response task, a choice-by-location task, and the Simon task. *J. Psychophysiol.* 15, 241–255. doi: 10.1027//0269-8803.15.4.241

Van Steenbergen, H., Band, G. P., and Hommel, B. (2009). Reward counteracts conflict adaptation evidence for a role of affect in executive control. *Psychol. Sci.* 20, 1473–1477. doi: 10.1111/j.1467-9280.2009.02470.x

Wiegand, K., and Wascher, E. (2005). Dynamic aspects of stimulus-response correspondence: evidence for two mechanisms involved in the Simon effect. *J. Exp. Psychol. Hum. Percept. Perform.* 31, 453–464. doi: 10.1037/0096-1523. 31.3.453

Wilson, A., Benko, H., Izadi, S., and Hilliges, O. (2012). "Steerable augmented reality with the beamatron," in *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology,* New York, NY, 413–422. doi: 10.1145/2380116.2380169

Zajonc, R. B. (1965). Social facilitation. *Science* 149, 269–274. doi: 10.1126/science.149.3681.269

The virtual hand illusion is moderated by context-induced spatial reference frames

Jing Zhang[1,2†], Ke Ma[1†] and Bernhard Hommel[1]*

[1] Cognitive Psychology Unit, Institute for Psychological Research and Leiden Institute for Brain and Cognition, Leiden University, Leiden, Netherlands, [2] Center for the Study of Language and Cognition, Zhejiang University, Hangzhou, China

The tendency to perceive an artificial effector as part of one's own body is known to depend on temporal criteria, like the synchrony between stimulus events informing about the effector. The role of spatial factors is less well understood. Rather than physical distance, which has been manipulated in previous studies, we investigated the role of relative, context-induced distance between the participant's real hand and an artificial hand stimulated synchronously or asynchronously with the real hand. We replicated previously reported distance effects in a virtual reality setup: the perception of ownership increased with decreased distance, and the impact of synchrony was stronger for short distances. More importantly, we found that ownership perception and impact of synchrony were affected by previous distance: the same, medium distance between real and artificial hand induced more pronounced ownership after having experienced a far-distance condition than after a near-distance condition. This suggests that subjective, context-induced spatial reference frames contribute to ownership perception, which does not seem to fit with the idea of fixed spatial criteria and/or permanent body representations as the sole determinants of perceived body ownership.

Keywords: body image, self-recognition, sense of ownership, virtual hand illusion, spatial reference frame

INTRODUCTION

How do we perceive ourselves and what are the mechanisms underlying our ability to perceive our body as constituting our bodily self? A recent technique to investigate this issue is the rubber hand illusion (RHI) and its virtual-reality version, the virtual hand illusion (VHI). In the RHI/VHI, participants perceive an artificial physical or virtual hand as a part of their own body (Botvinick and Cohen, 1998; Ehrsson et al., 2004; Tsakiris and Haggard, 2005; Slater et al., 2008; Shimada et al., 2009). This illusion can be induced by synchronously stroking a rubber/virtual hand placed in front of a participant in such a way that it seems extend from the participant's body, while the corresponding real hand is hidden from view. After a short while of synchronous stroking or, as in the virtual case, of perceived synchrony between own and artificial hand, the participant starts to get the perceptual impression that the rubber/virtual hand becomes his or her own hand.

Temporal synchrony between multimodal input coming from the real and artificial hand is crucial for the illusion, as asynchronous conditions (in which one stimulus stream is delayed with respect to the other by several 100s of ms) commonly produce significantly lower ownership ratings. Interestingly, however, there is also evidence for spatial criteria for perceived ownership.

While the illusion is most pronounced with minimal gaps between real and artificial hand (e.g., Costantini and Haggard, 2007; Lloyd, 2007; Gentile et al., 2013), the illusion does survive some discrepancies. For example, Lloyd (2007) showed that the strength of the illusion declined significantly if the rubber hand is placed horizontally more than 27.5 cm away from the participant's real corresponding hand. However, Zopf et al. (2010) did not find a reduction in RHI strength with distances up to 45 cm between the real and fake hands, which might suggest that the illusion relies on reaching distance. Preston (2013) considered the possibility that it may not be the absolute distance between real and artificial hand that matters but, rather, the distance between real hand and trunk. She manipulated both the absolute distance between real and artificial hand and their relative distance from body midline. The finding is that the strength of the illusion is reduced only if the artificial hand is far from both the real hand and the trunk. Kalckert and Ehrsson (2014) varied the vertical instead of the horizontal distance. The illusion became weaker with increasing distance.

These and related findings were taken to suggest a role of spatial reference frames when considering whether an artificial hand is or is not part of one's body. Maravita et al. (2003) proposed that although visuotactile interactions are usually most pronounced for stimuli near the real body part, the space to be considered can be plastically modified with active tool-use. If so, the ownership-related spatial reference frame could be flexible. FMRI studies already showed some evidence for this possibility. Brozzoli et al. (2012) found that the hand-centered encoding of space was remapped when a rubber hand was perceived as one's own. In the present study, we were interested to see whether the situational context might also affect the spatial reference frame used to determine body ownership. The reasons for considering this possibility were some informal observations in other studies from our lab, where the order or presence/absence of conditions seemed to play a role (e.g., see Zhang and Hommel, 2015). Consider, for instance, a condition in which real hand (and body) and artificial hand are separated by a noticeable spatial gap. After just having experienced a condition with a closer connection between real and artificial hand, the artificial hand may now be perceived as rather distant, and the perception of ownership may be reduced. In contrast, after just having experienced a condition with an even greater gap between real and artificial hand, the artificial hand may now be perceived as rather closely connected to one's real hand or body and, thus, motivate rather high ownership ratings.

We tested this possibility by presenting participants with a condition with a noticeable but not extreme gap between real and artificial hand after having them presented with an even larger or with a smaller gap. That is, we used far-distance and near-distance conditions as *priming* conditions and a medium-distance condition as *test* condition. We used a VHI setup, in which participants wore a data glove and were presented with a 3D virtual hand. Tactile stimulation was applied through vibrators attached to the data glove, which avoids the rather artificial stroking procedure required for the traditional RHI setup. Given previous reports about divergent findings for different kinds of ownership-perception indicators (Rohde et al.,

2011), we used the standard ownership questionnaire (adapted for the virtual setup), in addition to proprioceptive drift and skin conductance response (SCR), two commonly used "objective" measures to assess the ownership illusion. Our prediction was that the same medium-distance test condition should produce lower ownership ratings after a near-distance priming condition than after a far-distance priming condition.

MATERIALS AND METHODS

Participants

There were 34 participants (three more were tested but did not complete the experiment), all of them were student volunteers (eight males; mean age = 23 years, SD = 2.38, range 18–28) from Leiden University, unfamiliar with the rubber/VHI, who participated in exchange for course credit or pay. Ethical approval for this study was obtained from the local Psychology Research Ethics Committee, and written informed consent was obtained from all participants.

Design

We used a 2-factorial within-participants design. The two factors were synchrony (synchronous vs. asynchronous) and distance-condition sequence (near-medium vs. far-medium). To avoid the influence of fatigue and response strategies, we divided the experiment into two sessions performed on different days (with 1.32 days on average in between). In the near-medium session, participants were exposed to a condition with a medium-sized gap between their real hand and a virtual hand on the screen in front of them after having been exposed to a condition with a small-sized gap. In the far-medium session, participants were exposed to the same medium-sized gap condition after having been exposed to a condition with a large-sized gap. All participants served in both sessions. Half of the participants participated in the near-medium session before the far-medium session while the other half participated in reversed order. In each of the two distance conditions per session, the participant would be exposed to a synchronous condition and an asynchronous condition. The order of these two synchrony conditions was the same for the two distance conditions for a given participant, but the order of synchronous and asynchronous conditions was balanced across participants.

Experimental Setup

The study was performed in a virtual reality environment. The setup consisted of a data glove (Cyberglove, measurement frequency = 100 Hz, latency = 10 ms), virtual reality software (Vizard), and a large projection screen of 212 cm × 133 cm, which was around 50 cm away from the participants. The Cyberglove had a vibrator on the palm, through which we were able to apply the tactile stimulation (vibration frequency = 0–125 Hz). Participants wore the glove on their right hand, which during the experiment was placed in a fixed position inside a black box (50 cm × 24 cm × 38 cm) with the palm facing up. A Biopac MP100 acquisition unit and AcqKnowledge software were used for the SCR data recording.

We used a virtual hand from Vizard character set and imported the tracker and data glove module into Vizard. The virtual hand was projected on the large screen in three different positions (always aligned with the participant's real hand): near (seemingly extending from the real hand), medium (22 cm horizontally away from the near position), and far (44 cm horizontally away from the near position), as shown in **Figure 1**. In the near conditions, the virtual hand was projected in alignment with the participant's real hand, which looked as if the virtual hand extended from the real hand; and in the far conditions, the virtual hand was 44 cm horizontally away from the near position.

Measurements

Subjective ownership perception was assessed by means of the standard ownership questionnaire developed by Botvinick and Cohen (1998), which we only adjusted to the virtual setup. Corresponding versions of this questionnaire have been used in various kinds of rubber/VHI experiments (Botvinick and Cohen, 1998; Makin et al., 2008; Zhang and Hommel, 2015). We also considered more objective measures for explorative purposes, namely, proprioceptive drift (Longo et al., 2008; Kammers et al., 2009; Riemer et al., 2013; Ma and Hommel, 2015b), and SCR (Armel and Ramachandran, 2003; Yuan and Steed, 2010; Ma and Hommel, 2013, 2015a,b). Subjective and objective measures have shown different outcomes in various cases (e.g., Ma and Hommel, 2013, 2015a), suggesting that they do not reflect the exact same mechanisms, and objective measures such as proprioceptive drift have been criticized for several reasons (Rohde et al., 2011; Folegatti et al., 2012). This makes

it difficult to make predictions for the more objective measures, but we nevertheless analyzed and report effects for all three measures.

Questionnaire

We used an adapted version (Slater et al., 2008; Padilla et al., 2010; Ma and Hommel, 2013) of the standard nine-item questionnaire (Botvinick and Cohen, 1998) to assess the strength of ownership illusion in our design. Q1–Q5 are related to the experience of perceiving the hand as one's own (Kalckert and Ehrsson, 2014; Ma and Hommel, 2015a,b), and Q6–Q9 assess possible side effects of the illusion. Each statement was scored on a 7-point Likert scale, ranging from 1 for "strongly disagree" to 7 for "strongly agree", and 4 for 'uncertain.' The questionnaire items are shown below:

Q1: I felt as if I was looking at my own hand.
Q2: I felt as if the virtual hand were my hand.
Q3: It seemed as if I were feeling the touch of the ball in the location where I saw the virtual hand touched.
Q4: It seemed as though the touch I felt was caused by the ball touching the virtual hand.
Q5: I felt as if the virtual were part of my body.
Q6: It felt as if my (real) hand were drifting toward the virtual hand on the screen.
Q7: It seemed as if I might have more than one right hand or arm.
Q8: I felt as if my real hand were turning virtual.
Q9: I felt as if my right hand had disappeared.

So far, no psychometrically analyzed version of the questionnaire has been developed and no absolute criteria
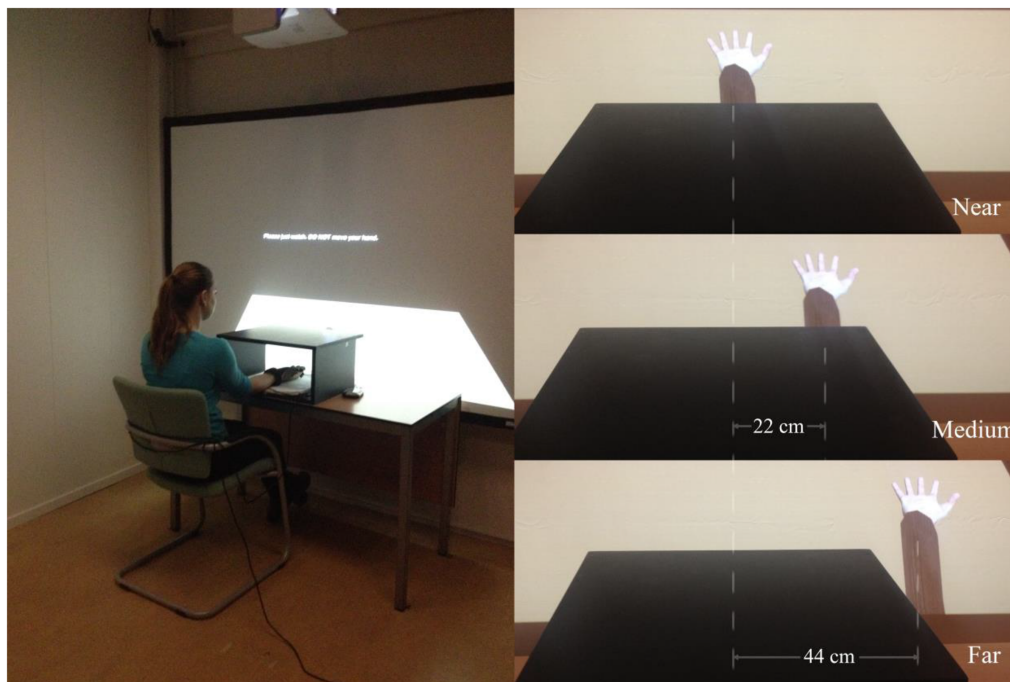

FIGURE 1 | The experimental setup (left) and the three different positions at which the virtual hand was shown on the screen (right).

for determining the absence or presence of an illusion have been suggested. We therefore used the comparison between synchronous and asynchronous conditions as a proxy. A significantly stronger ownership score in synchronous as compared to asynchronous conditions was thus taken to indicate a relative increase in perceived ownership, and the size of the increase was taken to reflect the strength of the impact of the corresponding factor.

Proprioceptive Drift

The method we used for the proprioceptive drift measurement was the same as in our earlier study (Ma and Hommel, 2015b). We presented an array of letters on the screen and asked participants to verbally report the felt location of their real right middle finger by choosing the particularly corresponding letter. To work against response strategies, the letters in the strings were presented in random order. The letter size differed depending on their alphabetic shape, with the biggest letter measuring approximately 2 cm. We recorded the corresponding letter before and after the illusion induction process (Botvinick and Cohen, 1998; Tsakiris and Haggard, 2005; Kalckert and Ehrsson, 2014). We calculated the distance between the letters and the screen side, and calculated the proprioceptive drift by subtracting the distance in the post-measure from the distance in pre-measure, so that positive values imply a drift toward the virtual hand.

SCR

The method we used for the SCR measurement was also the same as our earlier study (Ma and Hommel, 2015b). We measured SCR during a threat phase, in which a virtual knife appeared above the virtual hand on the screen and moved down to cut the virtual hand. It took 4 s to cut the virtual hand and another 4 s to move back to the original position. The cutting procedure was repeated five times. We defined a latency onset window between 1 and 6 s after stimulus/event onset, namely, when the virtual knife cut the virtual hand, with the skin conductivity level before event onset serving as baseline (see Boucsein, 1992; Figner and Murphy, 2010; Ma and Hommel, 2013, 2015a,b). We then calculated the magnitude of the event-induced SCR by subtracting baseline skin conductivity from the peak amplitude of the SCR during the analyzed time window.

Procedure

When participants arrived in the lab, they were asked to put the glove on their right hand and a SCR remote transmitter on their left wrist with a strap. Then they were seated in front of a desk and a projection screen (see **Figure 1**). They were instructed to put their right hand with palm upward into a box in between the participant and the screen, so they could not see their own right hand. Participants' right hands were placed at the middle position of the box, and they were asked not to move their right hand during the experiment.

As mentioned already, each session consisted of four blocks (e.g., far/synchronous, far/asynchronous, medium/synchronous, medium/asynchronous). The sequence of events was the same for each block. First, participants judged the location of the right middle finger of their real hand, as described above. Second, the

illusion was induced by means of visuotactile stimulation. The virtual hand was shown on the screen, seen as extending from the participant's right hand, and a small virtual ball appeared above the virtual hand. The ball took 4 s to move down to contact the virtual hand's palm, and then took another 4 s to return to its original position; this illusion induction procedure was repeated for 90 s. In the synchronous conditions, the contact between the virtual ball and hand was associated with the onset of the palm vibration stimulator of the glove, so as to apply synchronous visuotactile stimulation. In the asynchronous conditions, the vibration was delayed by 4 s, so that visual and tactile stimulation did not match. The vibration lasted for 1 s for every ball movement procedure in all conditions. Third, participants would again judge the location of their real right middle finger, and then fill in the ownership questionnaire on paper with his/her their unstimulated hand. Fourth, the same illusion induction procedure in the second step was implemented again, and then the virtual ball was replaced by a virtual knife, the threat phase started, SCR was measured while the virtual hand was threatened by the virtual knife, as described above. Finally, participants were asked to take a short break before they experienced the next block.

RESULTS

Priming Conditions (Near and Far)

All questionnaire items scores for priming conditions were submitted to 2×2 ANOVA with the factors synchrony (synchronous vs. asynchronous) and distance (near vs. far). Means and standard errors for each question item in each condition, F, P and effect size values for each question item, are shown in **Table 1**. The synchrony pattern of results is similar to previous studies (Botvinick and Cohen, 1998; Slater et al., 2008), ownership questions (Q1–Q5) showed significant synchrony effects, while control questions (Q6–Q9) did not (except for Q8).

Following Kalckert and Ehrsson (2014) and Ma and Hommel, 2015a,b), we aggregated the ownership questions (Q1–Q5) and computed their mean to represent sense of ownership. This score was analyzed by means of a 2×2 ANOVA with the factors synchrony (synchronous vs. asynchronous) and distance (near vs. far). There were significant main effects of synchrony, $F(1,33) = 71.470$, $p < 0.001$, $\eta_p^2 = 0.684$, indicating a stronger sense of body ownership for synchronous visuotactile stimulation ($M = 4.126$, $SE = 0.180$) than for asynchronous stimulation ($M = 2.535$, $SE = 0.159$); and of distance, $F(1,33) = 9.837$, $p = 0.004$, $\eta_p^2 = 0.230$, showing a stronger sense of body ownership for near ($M = 3.571$, $SE = 0.134$) than for far ($M = 3.091$, $SE = 0.184$) placement of the virtual hand. Importantly, the interaction between the two factors was also significant, $F(1,33) = 18.812$, $p < 0.001$, $\eta_p^2 = 0.363$, indicating that the synchrony effect was more pronounced for the near than the far condition, see **Figure 2**. Two tailed paired t-tests revealed that the synchrony effect was significant in both near and far positions, $t(33) = 8.980$, $p < 0.001$, $d = 1.995$, and $t(33) = 5.703$, $p < 0.001$, $d = 0.943$, respectively; and the distance effect was significant in synchronous conditions, $t(33) = 4.425$, $p < 0.001$,

TABLE 1 | Priming conditions (near and far): means (M) and standard errors (SE); F, P, and effect size values for all the questionnaire items scores, and also for the aggregate scores of Q1–Q5, with df = 33.

M/SE	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q1–Q5
Near-synchronous	4.35/0.27	4.38/0.26	5.32/0.19	4.88/0.24	4.00/0.23	3.32/0.26	2.23/0.19	3.00/0.26	1.19/0.15	4.59/0.19
Near-asynchronous	3.09/0.25	2.59/0.19	2.53/0.27	1.94/0.21	2.62/0.21	2.88/0.23	2.15/0.19	2.47/0.22	1.85/0.18	2.55/0.16
Far-synchronous	3.18/0.27	3.35/0.29	4.79/0.27	4.09/0.29	2.91/0.26	2.94/0.25	2.15/0.22	2.44/0.25	1.97/0.17	3.66/0.23
Far-asynchronous	2.73/0.28	2.79/0.24	2.59/0.28	2.06/0.22	2.41/0.23	2.71/0.26	1.82/0.17	2.29/0.21	1.88/0.21	2.52/0.19
$F/p/\eta_p^2$ (Distance)	11.10/0.002/0.25	3.35/0.076/0.09	1.30/0.262/0.03	2.52/0.122/0.07	11.67/0.002/0.26	1.75/0.195/0.05	1.64/0.210/0.05	4.51/0.041/0.12	0.17/0.681/0.01	9.84/0.004/0.23
$F/p/\eta_p^2$ (Synchrony)	16.24/<0.001/0.33	30.79/<0.001/0.48	76.22/<0.001/0.70	93.98/<0.001/0.74	23.45/<0.001/0.41	2.80/0.104/0.08	2.16/0.151/0.06	5.74/0.022/0.15	0.48/0.492/0.01	71.47/<0.001/0.68
$F/p/\eta_p^2$ (Interaction)	6.29/0.017/0.16	11.72/0.002/0.26	2.45/0.127/0.07	5.59/0.024/0.14	5.34/0.03/0.14	0.53/0.471/0.02	0.55/0.462/0.02	1.05/0.313/0.03	0.01/0.905/0.01	18.81/<0.001/0.36

$d = 0.764$, but not in asynchronous conditions, $t(33) = 0.227$, $p = 0.822$, $d = 0.034$.

Test Condition (Medium)

Questionnaire

All questionnaire items scores for the test condition were submitted to 2 × 2 ANOVA with the factors synchrony (synchronous vs. asynchronous) and context (near-medium vs. far-medium). Means and standard errors for each question item in each condition, F, P, and effect size values for each question item, are shown in **Table 2**.

The mean score for ownership (Q1–Q5) was analyzed by means of a 2 × 2 ANOVA with the factors synchrony (synchronous vs. asynchronous) and context (near-medium vs. far-medium). There were significant main effects of synchrony, $F(1,33) = 67.002$, $p < 0.001$, $\eta_p^2 = 0.670$, showing a stronger sense of ownership for synchronous ($M = 4.129$, $SE = 0.175$) than for asynchronous conditions ($M = 2.694$, $SE = 0.187$); and of context, $F(1,33) = 39.818$, $p < 0.001$, $\eta_p^2 = 0.547$, showing stronger ownership for the far-medium ($M = 3.768$, $SE = 0.156$) than the near-medium condition ($M = 3.056$, $SE = 0.179$). The interaction between the two factors was also significant, $F(1,33) = 7.192$, $p = 0.011$, $\eta_p^2 = 0.179$, suggesting that the synchrony effect was more pronounced in the far-medium than the near-medium condition, see **Figure 3**. Two tailed paired t-tests showed that the synchrony effect was significant in the near-medium [$t(33) = 6.271$, $p < 0.001$, $d = 0.974$] and the far-medium condition [$t(33) = 7.485$, $p < 0.001$, $d = 1.538$]; and the context effect was significant in both synchronous conditions,

FIGURE 2 | Mean Score of Ownership Questions (Q1–Q5) for priming condition. Error bars represent ±1 SE.
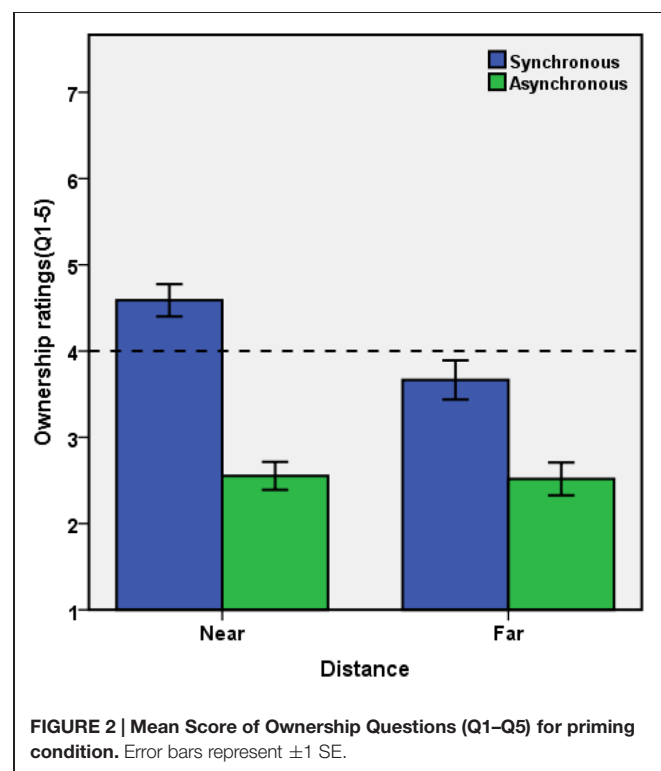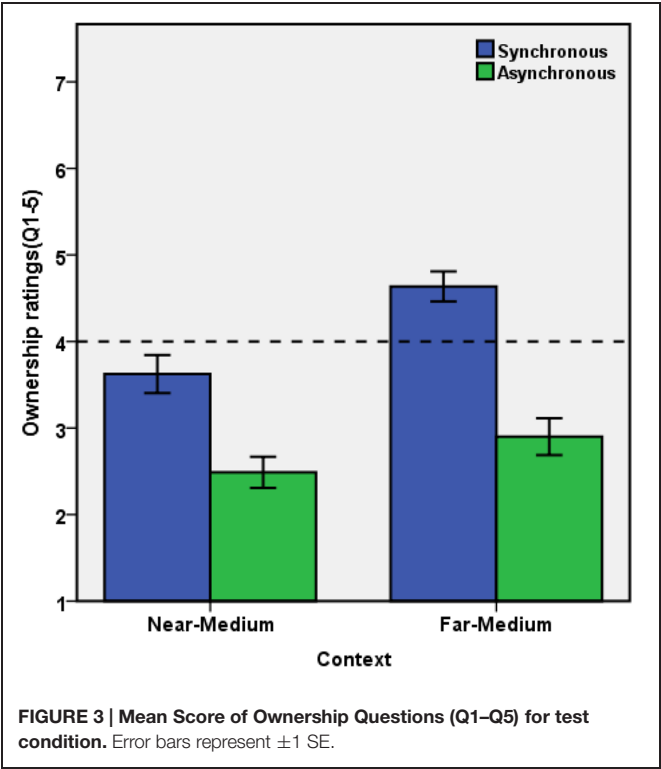
TABLE 2 | Test conditions (near-medium and far-medium): means (M) and standard errors (SE); F, P, and effect size values for all the questionnaire items scores, and also for the aggregate scores of Q1–Q5, with df = 33.

M/SE	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q1–Q5
Near-medium-synchronous	3.00/0.25	3.56/0.28	4.03/0.29	4.21/0.28	3.32/0.26	3.47/0.27	2.06/0.18	2.73/0.24	2.21/0.19	3.62/0.22
near-medium-asynchronous	2.71/0.25	2.73/0.23	2.35/0.24	2.09/0.19	2.56/0.20	2.88/0.26	2.23/0.21	2.53/0.21	1.97/0.19	2.49/0.18
far-medium-synchronous	4.41/0.23	4.47/0.22	5.23/0.17	4.94/0.21	4.12/0.23	3.38/0.28	2.29/0.20	3.09/0.25	2.21/0.24	4.63/0.17
far-medium-asynchronous	3.15/0.26	3.15/0.29	2.94/0.28	2.44/0.25	2.82/0.22	2.88/0.24	2.12/0.22	2.41/0.25	2.00/0.12	2.90/0.21
$F/p/\eta_p^2$ (Context)	23.45/<0.001/0.41	9.29/0.005/0.22	23.22/<0.001/0.41	12.45/0.001/0.27	9.84/0.004/0.23	0.05/0.817/0.01	0.11/0.743/0.01	0.33/0.571/0.01	0.01/0.929/<0.01	39.82/<0.001 0.55
$F/p/\eta_p^2$ (Synchrony)	10.23/0.003/0.24	21.71/<0.001/0.40	69.92/<0.001/0.68	105.84/<0.001/0.76	27.35/<0.001/0.45	8.07/0.008/0.20	01.000/0	8.96/0.005/0.21	2.36/0.134/0.07	67.00/<0.001 0.67
$F/p/\eta_p^2$ (Interaction)	7.29/0.011/0.18	1.77/0.193/0.05	2.67/0.111/0.07	1.30/0.262/0.04	3.13/0.086/0.09	0.08/0.779/0.01	2.42/0.129/0.07	2.81/0.103/0.08	0.01/0.908/< 0.01	7.19/0.011/0.18

FIGURE 3 | Mean Score of Ownership Questions (Q1–Q5) for test condition. Error bars represent ±1 SE.
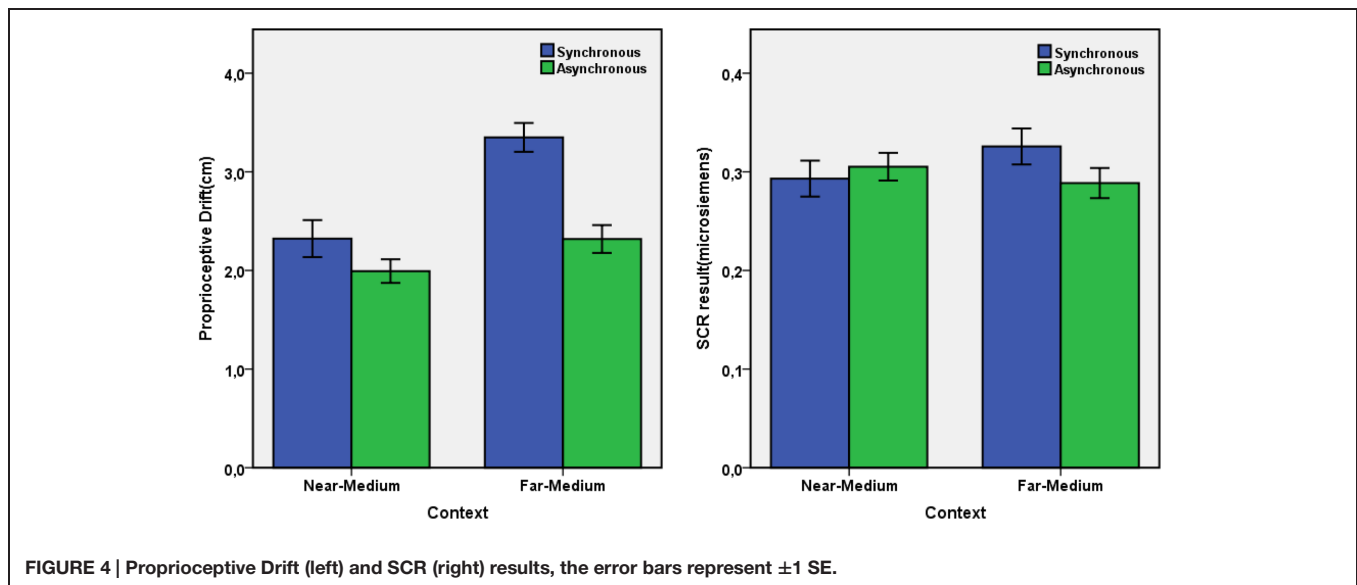
$t(33) = 5.458$, $p < 0.001$, $d = 0.882$, and asynchronous conditions, $t(33) = 3.244$, $p = 0.003$, $d = 0.359$.

Proprioceptive Drift

The proprioceptive drift results were log transformed and the normality of distribution was determined using the Shapiro–Wilk test, $p > 0.8$.

The transformed scores of proprioceptive drift for each condition were submitted to a 2×2 ANOVA with the factors synchrony (synchronous vs. asynchronous) and context frame (near-medium vs. far-medium). There were significant main effects of synchrony, $F(1,33) = 26.035$, $p < 0.001$, $\eta_p^2 = 0.441$, showing a stronger proprioceptive drift with synchronous ($M = 2.836$ cm, $SE = 0.107$) than asynchronous stimulation ($M = 2.156$ cm, $SE = 0.100$); and of context, $F(1,33) = 24.804$, $p < 0.001$, $\eta_p^2 = 0.429$, showing a stronger proprioceptive drift in the far-medium ($M = 2.834$ cm, $SE = 0.104$) than the near-medium condition ($M = 2.158$ cm, $SE = 0.105$). The interaction also reached significance, $F(1,33) = 4.170$, $p = 0.049$, $\eta_p^2 = 0.112$, indicating that the synchrony effect was more pronounced in the far-medium than the near-medium condition. As shown in **Figure 4**, the outcome pattern was comparable to that for the ownership questionnaire items. Two tailed paired t-tests showed that the synchrony effect was significant in the far-medium [$t(33) = 5.180$, $p < 0.001$, $d = 1.229$], but not the near-medium condition [$t(33) = 1.412$, $p = 0.167$, $d = 0.368$]; and the context effect was significant in synchronous conditions, $t(33) = 3.954$, $p < 0.001$, $d = 1.054$; but not in asynchronous conditions, $t(33) = 1.941$, $p = 0.061$, $d = 0.429$.

FIGURE 4 | Proprioceptive Drift (left) and SCR (right) results, the error bars represent ±1 SE.

SCR

The SCR results were log transformed and the normality of distribution was determined using the Shapiro–Wilk test, $p > 0.6$.

The transformed scores of SCR for each conditions were submitted to a 2×2 ANOVA with the factors synchrony (synchronous vs. asynchronous) and context (near-medium vs. far-medium). There was no main effect but the interaction was significant, $F(1,33) = 5.667$, $p = 0.023$, $\eta_p^2 = 0.147$, indicating that the synchrony effect was more pronounced in the far-medium than the near-medium conditions (see **Figure 4**). Two-tailed paired t-tests revealed that the synchrony effect was significant for the far-medium condition, $t(33) = 2.587$, $p = 0.014$, $d = 0.379$, but not for the near-medium condition, $t(33) = 0.723$, $p = 0.475$, $d = 0.128$; while the context effect was not significant in synchronous conditions, $t(33) = 1.821$, $p = 0.078$, $d = 0.306$, or asynchronous conditions, $t(33) = 1.135$, $p = 0.265$, $d = 0.194$.

DISCUSSION

Temporal relationships between different sources of intermodal stimulation are known to affect the degree of perceived body ownership. Spatial factors also play a role, but they are less well understood. In contrast to previous studies, which looked into the distance between real and artificial hands, we tested the possibility that the situational context has an impact on whether a given distance is perceived as short or long. We thus tested the same medium-distance condition after a near-distance and after a far-distance condition, to see whether ownership ratings are more pronounced in the latter than in the former condition. Our findings provide clear evidence that the situational context affects perceived ownership. In particular, our findings have three implications.

Firstly, the questionnaire results for the two priming conditions showed that we were able to replicate the distance effect reported by Lloyd (2007) in a virtual setup. When the virtual hand was placed in a near position, questionnaire scores were significantly higher than those in the far position. It is interesting to see that the absolute ownership scores were not very high in the present study, probably because our setup made the virtual hand look a little bit far away from the participants even in the near condition. This is also consistent with previous studies (Lloyd, 2007; Preston, 2013), which suggested important roles of both distance and reaching space. Hence, our findings can be taken to confirm that distance effects are rather robust.

Second, our results showed that the synchrony-induced increase in ownership perception was significantly stronger for a near than far placed virtual hand. This provides even more direct evidence for the idea that ownership perception takes the distance between real and artificial hand (and/or between real body and artificial hand) into account. This is consistent with previous observations and corresponding theoretical claims (Lloyd, 2007; Tsakiris, 2010; Preston, 2013; Kalckert and Ehrsson, 2014). As Tsakiris (2010) suggested, one criterion for the ownership perception may occur as a result of the comparison between current sensory input and body-related reference frames. Alternatively, a distance rule may apply. Such a rule may operate continuously, with the probability of ownership perception increasing with decreasing distance, discontinuously, with ownership perception being restricted to candidate effector is within reaching space, or reflect some interaction of both. Given that we observed interactions between distance and synchrony for both (near and far) priming conditions and (medium) test conditions, a merely discontinuous rule does not seem to be sufficient: given that all our conditions fell into reaching space, such a simple rule could not account for such interactions. This leaves a simple distance rule and an interaction between a distance rule and a discontinuous criterion as possibilities.

Third, in the test condition, perceived body ownership was affected by the perceptual context: While absolute distance was kept constant, the relative size of the ownership illusion varied as a function of the context-induced relative distance between real

hand (or body) and artificial hand. Given the impact of actual distance observed in the priming conditions, this should not be taken to rule out contributions from physical distance. However, relative distance that relates previous experiences to the current distance between real and artificial hand seems to contribute as well. This observation is not consistent with the assumption that ownership perception relies on objective situational variables and internal representations thereof alone. It also does not fit with assumptions that only objective spatial parameters, like reaching space, and/or stable pre-existing body models play a role. Rather, ownership perception seems to rely on various informational sources that include subjective impressions informed by previous experiences in the same situation (Ma and Hommel, 2015b).

One thing to note is that, in our experiment setup, the virtual hand seemed to extend from the participant's real hand into the screen, so that the virtual hand always looked longer than the real hand. Could that have affected our results? Even though we are unable to exclude main effects, there are two reasons why we do not consider it plausible that this aspect can account for our main observations. For one, the "virtual extension" was the same in all conditions, as we only manipulated the horizontal distance between the real and the artificial hand. This suggests that, even if there was some effect, it should have impacted all conditions equally. For another, previous RHI studies suggest that such kinds of "virtual extensions" do not seem to influence the synchrony effect significantly. For example, in Preston and Newport (2012), the experimenter pulled the participant's arm while participants viewed the pull in a real-time video of themselves. In the video, the arm looked like being stretched to twice of its normal length. Participants did have the impression of their arm being stretched and they overestimated reaching distance, but the actual reaches were unaffected. In one of Armel and Ramachandran's (2003) experiments, the arm looked like being stretched to 0.91 m, but the basic illusion was still obtained. Finally, Kilteni et al. (2012) found that participants experienced ownership illusions even for a virtual arm that was about three times as long as a real arm. As we mentioned before, the ownership scores in the present study are relatively low, an observation that we attribute to the arm extension design we used in our study. Similar observations have been made in previous studies (Armel and Ramachandran, 2003;

Kilteni et al., 2012), where ownership ratings were relatively low when the rubber/virtual hand seem to be much longer than the real arm.

In addition to the more theoretical implications, our observations are also of relevance methodologically. For one, they strongly suggest that sequence effect can play an important role in moderating the size of the ownership illusion. Our findings also provide convergent evidence for the conclusion that ownership questionnaires, proprioceptive drift, and SCR are not fully equivalent methods to assess perceive body ownership. In the present study, the questionnaire turned out to be much more sensitive to the impact of our manipulations on self-perception than the other two measures, which fits with previous observations (Rohde et al., 2011; Folegatti et al., 2012; Ma and Hommel, 2013, 2015a).

CONCLUSION

The present study extends our knowledge about the cognitive process underlying RHI/VHI by demonstrating the flexibility of spatial criteria for moderating perceived body ownership. This adds to previous evidence that ownership perception may not be a simple function of continuous or discontinuous distance rules or a cross-situationally stable body image. Rather, there is increasing evidence that multiple sources of information contribute to the illusion, so that the relative importance of a given source may very well depend on the situation and the existence of other informational sources. This again is consistent with previous claims that body representations are dynamic and continuously updated to reflect the present situation (e.g., Graziano and Botvinick, 2002; Ehrsson, 2012).

ACKNOWLEDGMENTS

REFERENCES

Armel, K. C., and Ramachandran, V. S. (2003). Projecting sensations to external objects: evidence from skin conductance response. *Proc. R. Soc. Lond. B Biol. Sci.* 270, 1499–1506. doi: 10.1098/rspb.2003.2364

Botvinick, M., and Cohen, J. (1998). Rubber hands 'feel' touch that eyes see. *Nature* 391, 756–756. doi: 10.1038/35784

Boucsein, W. (1992). *Electrodermal Activity*. Heidelberg: Springer.

Brozzoli, C., Gentile, G., and Ehrsson, H. H. (2012). That's near my hand! Parietal and premotor coding of hand-centered space contributes to localization and self-attribution of the hand. *J. Neurosci.* 32, 14573–14582. doi: 10.1523/JNEUROSCI.2660-12.2012

Costantini, M., and Haggard, P. (2007). The rubber hand illusion: sensitivity and reference frame for body ownership. *Conscious. Cogn.* 16, 229–240. doi: 10.1016/j.concog.2007.01.001

Ehrsson, H. H. (2012). "The concept of body ownership and its relation to multisensory integration," in *The New Handbook of Multisensory Processes*, ed. B. E. Stein (Cambridge, MA: MIT Press), 775–792.

Ehrsson, H. H., Spence, C., and Passingham, R. E. (2004). That's my hand! Activity in premotor cortex reflects feeling of ownership of a limb. *Science* 305, 875–877. doi: 10.1126/science.1097011

Figner, B., and Murphy, R. O. (2010). "Using skin conductance in judgment and decision making research," in *A Handbook of Process Tracing Methods for Decision Research: A Critical Review and User's Guide*, eds M. Schulte-Mecklenbeck, A. Küehberger, and R. Ranyard (New York, NY: Psychology Press), 163–184.

Folegatti, A., Farnè, A., Salemme, R., and De Vignemont, F. (2012). The rubber hand illusion: two's a company, but three's a crowd. *Conscious. Cogn.* 21, 799–812. doi: 10.1016/j.concog.2012.02.008

Gentile, G., Guterstam, A., Brozzoli, C., and Ehrsson, H. H. (2013). Disintegration of multisensory signals from the real hand reduces default limb self-attribution: an fMRI study. *J. Neurosci.* 33, 13350–13366. doi: 10.1523/JNEUROSCI.1363-13.2013

Graziano, M. S. A., and Botvinick, M. (2002). "How the brain represents the body: insights from neurophysiology and psychology," in *Common Mechanisms in Perception and Action: Attention and Performance*

XIX, eds W. Prinz and B. Hommel (Oxford: Oxford University Press), 136–157.

Kalckert, A., and Ehrsson, H. H. (2014). The spatial distance rule in the moving and classical rubber hand illusions. *Conscious. Cogn.* 30, 118–132. doi: 10.1016/j.concog.2014.08.022

Kammers, M. P., Longo, M. R., Tsakiris, M., Dijkerman, H. C., and Haggard, P. (2009). Specificity and coherence of body representations. *Perception* 38, 1804–1820. doi: 10.1068/p6389

Kilteni, K., Normand, J. M., Sanchez-Vives, M. V., and Slater, M. (2012). Extending body space in immersive virtual reality: a very long arm illusion. *PLoS ONE* 7:e40867. doi: 10.1371/journal.pone.0040867

Lloyd, D. M. (2007). Spatial limits on referred touch to an alien limb may reflect boundaries of visuo-tactile peripersonal space surrounding the hand. *Brain Cogn.* 64, 104–109. doi: 10.1016/j.bandc.2006.09.013

Longo, M. R., Schüür, F., Kammers, M. P., Tsakiris, M., and Haggard, P. (2008). What is embodiment? A psychometric approach. *Cognition* 107, 978–998. doi: 10.1016/j.cognition.2007.12.004

Ma, K., and Hommel, B. (2013). The virtual-hand illusion: effects of impact and threat on perceived ownership and affective resonance. *Front. Psychol.* 4:604. doi: 10.3389/Fpsyg.2013.00604

Ma, K., and Hommel, B. (2015a). Body-ownership for actively operated non-corporeal objects. *Conscious. Cogn.* 36, 75–86. doi: 10.1016/j.concog.2015.06.003

Ma, K., and Hommel, B. (2015b). The role of agency for perceived ownership in the virtual hand illusion. *Conscious. Cogn.* 36, 277–288. doi: 10.1016/j.concog.2015.07.008

Makin, T. R., Holmes, N. P., and Ehrsson, H. H. (2008). On the other hand: dummy hands and peripersonal space. *Behav. Brain Res.* 191, 1–10. doi: 10.1016/j.bbr.2008.02.041

Maravita, A., Spence, C., and Driver, J. (2003). Multisensory integration and the body schema: close to hand and within reach. *Curr. Biol.* 13, R531–R539. doi: 10.1016/S0960-9822(03)00449-4

Padilla, M. A., Pabon, S., Frisoli, A., Sotgiu, E., Loconsole, C., and Bergamasco, M. (2010). "Hand and arm ownership illusion through virtual reality physical interaction and vibrotactile stimulations," in *Haptics: Generating and Perceiving Tangible Sensations*, eds A. M. L. Kappers, J. B. F. van Erp, W. M. Bergmann Tiest, and F. C. T. van der Helm (Heidelberg: Springer), 194–199.

Preston, C. (2013). The role of distance from the body and distance from the real hand in ownership and disownership during the rubber

hand illusion. *Acta Psychol.* 142, 177–183. doi: 10.1016/j.actpsy.2012.12.005

Preston, C., and Newport, R. (2012). How long is your arm? Using multisensory illusions to modify body image from the third person perspective. *Perception* 41, 247–249.

Riemer, M., Kleinböhl, D., Hölzl, R., and Trojan, J. (2013). Action and perception in the rubber hand illusion. *Exp. Brain Res.* 229, 383–393. doi: 10.1007/s00221-012-3374-3

Rohde, M., Di Luca, M., and Ernst, M. O. (2011). The rubber hand illusion: feeling of ownership and proprioceptive drift do not go hand in hand. *PLoS ONE* 6:e21659. doi: 10.1371/journal.pone.0021659

Shimada, S., Fukuda, K., and Hiraki, K. (2009). Rubber hand illusion under delayed visual feedback. *PLoS ONE* 4:e6185. doi: 10.1371/journal.pone.0006185

Slater, M., Perez-Marcos, D., Ehrsson, H. H., and Sanchez-Vives, M. V. (2008). Towards a digital body: the virtual arm illusion. *Front. Hum. Neurosci.* 2:6. doi: 10.3389/neuro.09.006.2008

Tsakiris, M. (2010). My body in the brain: a neurocognitive model of body-ownership. *Neuropsychologia* 48, 703–712. doi: 10.1016/j.neuropsychologia.2009.09.034

Tsakiris, M., and Haggard, P. (2005). The rubber hand illusion revisited: visuotactile integration and self-attribution. *J. Exp. Psychol. Hum. Percept. Perform.* 31, 80–91.

Yuan, Y., and Steed, A. (2010). "Is the rubber hand illusion induced by immersive virtual reality?" in *IEEE Virtual Reality conference, VR 2010*, eds B. Lok, G. Klinker, and R. Nakatsu (Waltham, MA: IEEE Computer Society), 95–102.

Zhang, J., and Hommel, B. (2015). Body ownership and response to threat. *Psychol. Res.* doi: 10.1007/s00426-015-0698-1 [Epub ahead of print].

Zopf, R., Savage, G., and Williams, M. A. (2010). Crossmodal congruency measures of lateral distance effects on the rubber hand illusion. *Neuropsychologia* 48, 713–725. doi: 10.1016/j.neuropsychologia.2009.10.028

Subjective Significance Shapes Arousal Effects on Modified Stroop Task Performance: A Duality of Activation Mechanisms Account

Kamil K. Imbir*

Faculty of Psychology, University of Warsaw, Warsaw, Poland

Activation mechanisms such as arousal are known to be responsible for slowdown observed in the Emotional Stroop and modified Stroop tasks. Using the duality of mind perspective, we may conclude that both ways of processing information (automatic or controlled) should have their own mechanisms of activation, namely, arousal for an experiential mind, and subjective significance for a rational mind. To investigate the consequences of both, factorial manipulation was prepared. Other factors that influence Stroop task processing such as valence, concreteness, frequency, and word length were controlled. Subjective significance was expected to influence arousal effects. In the first study, the task was to name the color of font for activation charged words. In the second study, activation charged words were, at the same time, combined with an incongruent condition of the classical Stroop task around a fixation point. The task was to indicate the font color for color-meaning words. In both studies, subjective significance was found to shape the arousal impact on performance in terms of the slowdown reduction for words charged with subjective significance.

Keywords: activation mechanisms, duality of mind, resource competition, duality of activation

COGNITIVE CONTROL AND DUALITY OF MIND

The human mind's ability to control actions and plan for them in the context of goals and expectations for the future is a milestone in cultural development. Cognitive psychology provides us with a concept of cognitive control (Cooper, 2010; Juvina, 2011) as an easily measurable mind ability that has much in common with our goal realization over time. Recently, the duality of mind approach, which describes and compares two separate mental systems or ways that the mind processes information, namely, automatic and controlled (for review see: Schneider and Chein, 2003), has been gaining attention in the science community (Gawronski and Creighton, 2013). There are several duality of mind theories focusing on specific processes (e.g., persuasion: Petty and Cacioppo, 1986; attitude: Wilson et al., 2000), focusing on a specific domain (such as cognition Strack and Deutsch, 2004) or emotion (Jarymowicz and Imbir, 2015), and, finally, more generally describing mind systems (e.g., Epstein, 2003; Kahneman, 2003, 2011) resulting in different processes. The aim of this paper is to address questions arising from the duality of mind perspective as applied to cognitive control and its probable activation mechanisms.

Activation Mechanisms Underlying Two Mind Systems

One of the function of emotions is to enhance motivation and action (Frijda, 2007; Kagan, 2007; Damasio, 2010). That is why it is reasonable to search for activation mechanisms specific to each mental system rather than to focus on a single mechanism. Consequently, I distinguish between two activation mechanisms (Imbir, 2015), namely, arousal, specific to the experiential mind, on the one hand, and, subjective significance, specific to the rational mind on the other hand. The reason for distinguishing both types of activation mechanisms is the concept of activation itself. Our mind needs activation mechanisms to sustain motivation to deal with everyday problems. For example, without pleasurable excitation related to exploration of new objects, an organism would not want to get to know such objects, and, consequently, would fail to explore the environment. Psychological research has often demonstrated negative outcomes of activation (meaning arousal) for tasks involving cognitive control and interference control such as the Emotional Stroop Task (EST). But, in fact, such outcomes occur when activation mechanisms are not specific to the task. Duality of mind theory suggests that we should try to discover specific activation for more complex processing. For that reason, dual theories can bring an important contribution to understanding activation mechanisms in both experiential and rational minds.

Arousal is relatively well recognized in psychology. Epstein (2003) argued that arousal level is the factor responsible for shifting the balance between experiential and rational minds toward the former. Arousal can be understood as energy that appears when an organism has to deal with arousing stimuli. This energy activates simple processes, making it easier to run as fast as we can when being chased by a dangerous animal or by another person who wants to rob us. Arousal works on a highly automated level. We do not have to think in order to know that something is threatening our survival or is physically attractive. This recognition immediately comes to mind when we look at it (Kahneman, 2011). Arousal changes associations in our mind by enhancing them for things which are more rather than less arousing. Arousal also influences the quality and results of associations-based processing (Strack and Deutsch, 2014) modifying relations between objects and connections strengthened in the associative store.

Subjective significance is a relatively newly proposed mechanism. It is related to the concept of so-called will power (e.g., Baumeister et al., 1998). To activate and continue careful, energy consuming (Kahneman, 2011), rational and propositionally based (Strack and Deutsch, 2004) thinking, we should engage something which operates at the same level. Arousal is damaging to almost every rational or controlled process because of the shifting balance toward associative mind. Systematic thinking is a luxury (due to the amount of effort required) that, in everyday situations, is better not to use (Kahneman, 2011) lest we overly fatigue ourselves. But why do people engage in such difficult thinking? I argue because they simply want to or think that it is worth doing so. This is probably attributable to the fact that certain situations or ideas may be important from the point of view of one's goals and expectations for the future. Subjective significance is, thus, a type of attitude toward an object that renders it important and significant, thereby, meriting the investment of energy in accurate systematic processing. Subjective significance could also be referenced to the salience concept (e.g., Kahnt and Tobler, 2013) describing importance of outcomes. For example, in decision making, both gains and losses associated with different options are different in valence but similar in salience. This mean that people perceive them as important in comparison with neutral outcomes that are perceived as non-salient.

To measure arousal and subjective significance, Self Assessment Manikin (SAM) scales were developed, based on Lang's (1980) idea of pictorial representation of bodily sensations that do not require a verbal response. This is especially important when arousal is being measured. However, the nature of the rational mind is propositional (Strack and Deutsch, 2004, 2014); thus, to make both scales more comparable, to each was added a description providing context and an explanation of its meaning. Thus, the characteristics of both scales were combined in order to facilitate collection of comparable assessments. **Figure 1** presents both scales used in these normative studies concerning words [Affective Norms for Polish Words (ANPWs): Imbir, 2015].
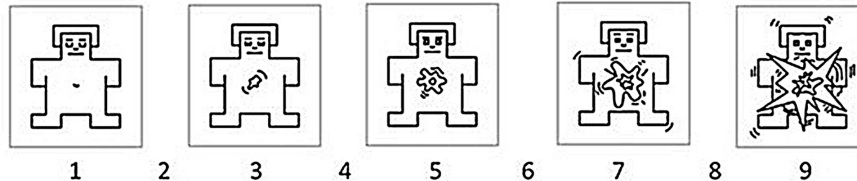
The ANPW study showed that both scales were reliable in terms of test–retest coherence and split half estimations (c.f. Imbir, 2015). Additionally, arousal and significant assessments were weakly correlated ($r = 0.24$) and, thus, measured different aspects of the activation properties of stimuli. Providing reliable measures of both variables enabled the testing of their contribution to EST processing.

Cognitive Control and Role of Emotion in Emotional Stroop Task

Stroop (1935) introduced a very simple paradigm allowing researchers to measure interference control (Nigg, 2000). The Stroop task is based on a task in which a participant is required to name the color of ink for different words. The task itself generates congruent trials (where the word RED is written in red ink) as well as incongruent trials (where the word GREEN is written in blue ink). The Stroop effects can be observed after subtracting reaction times in congruent trials from those in incongruent trials (Larsen et al., 2006). The difference derives from the interference of the two processes involved in task processing (c.f. **Figure 2**). The first is a controlled and effortful target task (ink color naming) which is rarely performed in everyday experience. The second relates to reading, and semantic access to, content, which, in the case of participants with extensive reading training in school, is highly automated, effortless and even uncontrolled. Both processes may work in the same direction where the probe is congruent or in opposite directions where the trial is incongruent, and where access to semantic meaning, which necessarily requires to be inhibited, gives us the wrong answer. This renders incongruent tests more difficult to perform and thus, reactions take longer.
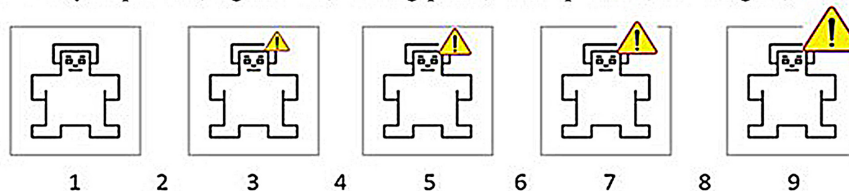
FIGURE 1 | Self Assessment Manikin (SAM) scales and descriptions used to measure arousal and subjective significance levels of verbal stimuli in ANPW (Imbir, 2015**) and ANPW_R (**Imbir, submitted**).**
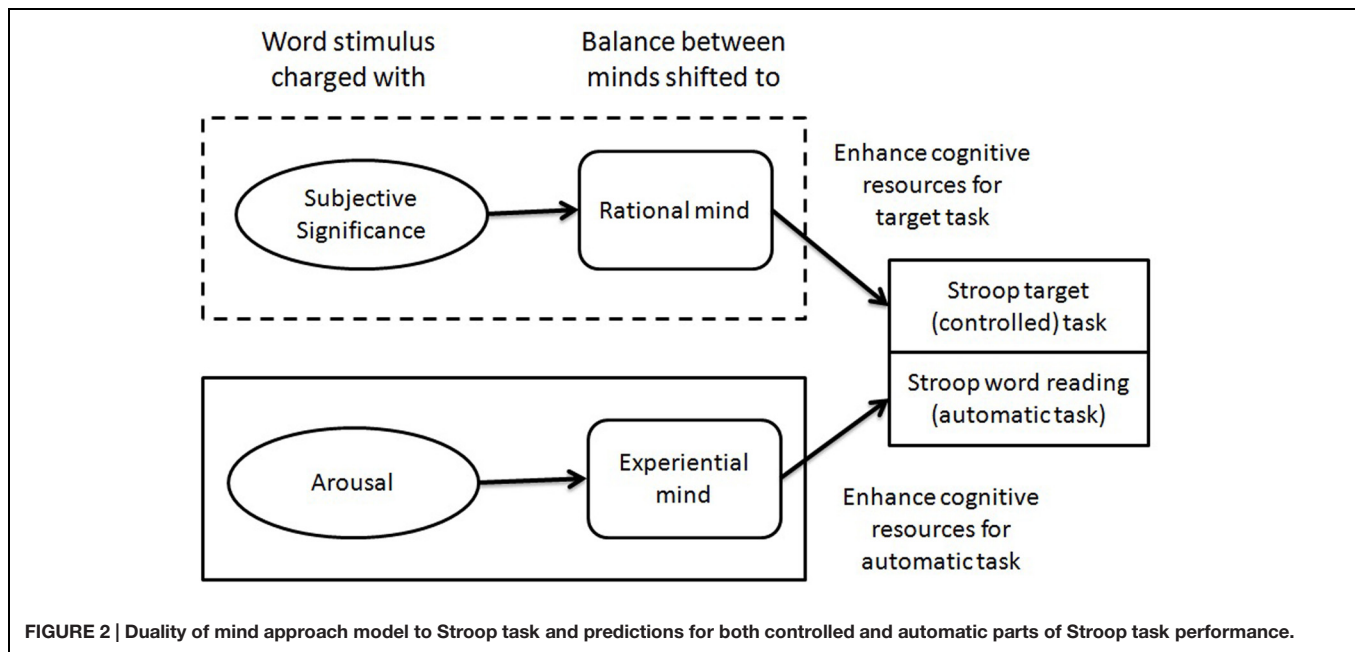
The EST is slightly different. The task itself is similar to the Stroop task but the trials are different experimental and controlled as opposed to congruent or incongruent. The semantic content of words is especially different, often emotional rather than neutral in meaning (e.g., Williams et al., 1996). EST is very sensitive to the properties of words used (Larsen et al., 2006); thus, when using this paradigm, the words must be carefully chosen for both experimental and control conditions in order to preclude the possibility of irrelevant factors potentially influencing the processing of the task.

There are many studies showing valence effects on EST performance, especially in the case of negatively valenced stimuli (Williams et al., 1996; c.f. McKenna and Sharma, 2004). Some clinical studies involving patients who had suffered traumatic experiences showed that EST slowdown was observed for words connected with traumatic experience (Watts et al., 1986; McKenna and Sharma, 1995, 2004). Apart from valence, EST performance is influenced by at least two important lexical variables. For example, Burt (2002) showed that word frequency influenced the color naming task in terms of response latencies. Less frequently occurring words resulted in longer reaction times as compared with more frequently occurring words. This was probably due to the greater resources required to process the less frequently occurring words and to the capacity of our cognition.

Larsen et al. (2006) demonstrated that, among 32 published EST studies, affective words used had lower frequency, longer length, and smaller orthographic neighborhood than the control (neutral) words. They concluded that this could have been the cause of the slowdown reported.

Another important variable is the arousal associated with each valence word used. For example, Dresler et al. (2009), by the careful use of factorial manipulation, showed that the arousal attributable to the word produced emotional interference, independently of valence. Other studies using neuroimaging techniques (fMRI) showed that, in healthy individuals, highly arousing stimuli elicited greater interference than stimuli with low arousal (Compton et al., 2003). Surprisingly this effect was greater for negative than for positive words. All of the above mentioned examples provide evidence that lexical word properties and activation mechanisms play a crucial role in the EST phenomenon; thus such factors must be carefully considered in experimental materials preparation.

Recent studies concerning the emotion duality model impact on cognitive control (Imbir and Jarymowicz, 2013) showed that the types of emotions it tested shaped the EST performance. The automatic emotions-related words (both negatively and positively valenced) generated slowdown in the case of EST as compared with neutral and reflective emotions-related words.

FIGURE 2 | Duality of mind approach model to Stroop task and predictions for both controlled and automatic parts of Stroop task performance.

This result convinced us to search for the mechanisms underlying cognitive control and duality of mind in order to explain the divergence.

Duality of Mind in Stroop Task

In the case of both the classical and modified Stroop tasks, two operations are competing for resources. The first one is the explicit task, addressing systematic processing (reflective like), which requires cognitive control to indicate the color of ink in which the word is displayed. The control is required because putting attention on the color of the words is not a spontaneous reaction. To achieve control, a participant must avoid the second process which is highly automated in nature. The reading of words in a visual field is a well-trained skill for any person able to read and practice the skill. Such reading is an excellent example of effortless processing (Kahneman, 2011), characteristic of an automated mind, and gives access to the semantic aspect of a word. In the classical Stroop task, the meaning of the word interferes with the required answer and generates slowdown in reaction times. In EST and modified Stroop task, some aspects of the word (e.g., arousal level or trauma-related content) attract attention and generate slowdown in answering. The dual nature of the Stroop task is expressed in **Figure 2**.

Such construction of Stroop like tasks allows for the measurement of interference control (Nigg, 2000) over automated function of reading and assimilation of semantic meaning. The control itself is an example of effortful processing (Kahneman, 2011) and should, thus, be sensitive to mechanisms characteristic of the rational mind. Previous studies have shown that valence or arousal included in stimuli can make the interference control difficult to maintain; thus, reaction times for controlled tasks are longer. The important question is whether there are some aspects of stimuli that can provide the activation for controlled processing? Taking into account

the duality of activation perspective presented in this study, subjective significance may play such a role.

I argue, that arousal of words is the pivotal factor influencing the automated part of Stroop task. More highly arousing words should capture more attention, necessarily meaning that cognitive control should be less effective. But subjective significance should provide the activation for the rational mind. This activation should influence the strength of control, making it easier to give an answer to the explicit task. The difference between arousal and subjective significance is that arousal is in fact a non-verbal, single way dimension (from sleep to excitement: c.f. Russell, 2003), whereas subjective significance is based on a conscious response to stimuli that can be neutral (of moderate subjective significance) but also strong (affirmative of subjective significance) or weak (non-affirmative of subjective significance). The semantic processing during Stroop task trials is associative rather than reflective in nature (c.f. Strack and Deutsch, 2004, 2014). Negation is a reflective operation (Deutsch et al., 2006) requiring time. At the first stage of semantic analysis, there is no time for negation processing; thus, stimuli triggering the significance concept should influence the cognitive control in the same way whether the stimuli be of low or high subjective significance. Only conscious analysis of meaning can trigger a negative association. In **Figure 2**, the duality of activation mechanisms in Stroop task performance is presented with the components influenced by two different aspects of activation.

Aim and Hypothesis

The aim of the present work was to investigate the activation mechanism underlying the EST effect. Although there is agreement that arousal is the most important factor modulating slowdown for emotional (especially negative) as compared with neutral words, this study sought to examine the duality of activation mechanism predictions. A higher level of slowdown

in color naming was expected to be observed in the case of high arousal as compared with low arousal stimuli. It was also expected that subjective significance presence (both low and high level) would modulate this relationship influencing the controlled part of modified Stroop task processing (c.f. **Figure 2**). A moderate level of subjective significance would mean the absence of this factor; thus, in these conditions only an arousal effect would be expected to occur.

MATERIALS AND METHODS

Apparatus

To present stimuli, a standard 15 inch laptop with Windows 7 operating system was used. The experiment script was prepared with E-Prime 2.0 software. Response keys were indicated by stickers with printed symbols: P for orange, C for red, Z for green, and N for blue (which are the first letters of the Polish words describing those colors: *Pomarańczowy, Czerwony, Zielony,* and *Niebieski*). Participants were instructed to use both hands when answering. They were also instructed to keep their fingers over the answer keys at all times during the experiment.

Materials

To create factorial manipulation, a list of 135 words (nouns) with checked affective qualities were chosen from among 4,905 words. This list was derived from ANPWs Reloaded (Imbir, submitted) which had been compiled using a methodology similar to a previous study concerning affective norms for lower number of words (Imbir, 2015). Two activation dimensions were examined: Arousal and Subjective Significance as well as control affective dimensions such as valence, concreteness and lexical word properties including frequency of appearance in the Polish language (based on Kazojć (2011)) and the number of letters (word length). Assessments for each word used in the current study are presented in Supplementary Materials 1. **Table 1** presents Mean values (*M*) and Standard Deviation (*SD*) for arousal and subjective significance manipulation groups.

To ensure that words chosen for factorial manipulation were correct, a 3 (arousal levels) × 3 (subjective significance levels) ANOVA was calculated. In the case of arousal ratings, a significant main effect of arousal, $F_{(2,126)} = 31.09$, $p < 0.001$, $\eta^2 = 0.83$, and a no significant main effect of subjective significance, $F_{(2,126)} = 0.88$, $p = 0.4$, $\eta^2 = 0.01$, were found. Taking into account subjective significance ratings, a no significance main effect of arousal, $F_{(2,126)} = 3.02$, $p = 0.053$, $\eta^2 = 0.04$, and a statistically significant main effect of subjective significance, $F_{(2,126)} = 35.62$, $p < 0.001$, $\eta^2 = 0.81$, were found. Although the *p*-value was slightly above 0.05, taking into account the huge differences in η^2, the factor manipulation was sufficient and independent. In both cases, no interaction effect was found.

To ensure that words chosen for factorial manipulation differed only in the case of manipulated variables, an additional 3 (arousal levels) × 3 (subjective significance levels) ANOVA was run controlling for affective (valence, concreteness) and lexical (natural logarithm of frequency, number of letters) dimensions. In the case of valence ratings, no significant main effect of arousal, $F_{(2,126)} = 1.46$, $p < 0.23$, $\eta^2 = 0.02$, and no significant main effect of subjective significance, $F_{(2,126)} = 1.89$, $p = 0.16$, $\eta^2 = 0.03$, were found. Taking into account concreteness ratings, neither a significant main effect of arousal, $F_{(2,126)} = 0.03$, $p = 0.97$, $\eta^2 < 0.001$ nor an effect of subjective significance, $F_{(2,126)} = 2.74$, $p = 0.07$, $\eta^2 = 0.04$, were found. In the case of frequency estimations, natural logarithm values from Kazojć's (2011) database concerning the right-skewed distribution (dataset consisted of a number of single word repetitions in a wide range of Polish texts) were analyzed. No significant main effect of arousal, $F_{(2,126)} = 1.24$, $p = 0.29$, $\eta^2 = 0.02$, and no effect of subjective significance, $F_{(2,126)} = 2.99$, $p = 0.054$, $\eta^2 = 0.05$, were found. Finally, word length (number of letters) was assessed and no significant main effect of arousal, $F_{(2,126)} = 0.57$, $p = 0.57$, $\eta^2 = 0.01$, and no effect of subjective significance, $F_{(2,126)} = 1.29$, $p = 0.28$, $\eta^2 = 0.02$, were found.

The above mentioned analyses revealed that the factorial manipulation used enabled the distinguishing between arousal and subjective significance levels of the words used. Furthermore, other factors that could have potentially influenced Stroop task performance such as frequency of appearance, word length, valence and concreteness were controlled. For this reason, observed differences may be attributed only to the designed manipulation.

Design

A within-subject 3 × 3 factorial design was applied by manipulating word arousal load (Low, Medium, and High) and subjective significance load (Low, Medium, and High). This generated nine groups of words, each containing 15 words. Other factors such as valence, concreteness, frequency of appearances, and numbers of letters in words were controlled and their level aligned between the groups.

Both studies were carried out in accordance with the recommendations of the bioethical committee of the Maria Grzegorzewska University without written informed consent from all subjects. Written consents were not collected as we had assured the participants of anonymity. The oral consent was made by participants in appearance of at least one lab staff member and documented in research diary. This procedure was suggested by the bioethical committee approving research. All subjects gave informed consent in accordance with the Declaration of Helsinki.

Experiment 1 – Modified Stroop Task
Participants

In the first experiment, 60 individuals (30 women) from different Warsaw universities (in equal proportion from the departments of social science, humanities, engineering, life science, and natural science) participated. The sample size of 60 participants was planned in advance. Only correct answers were analyzed; thus, six participants were excluded because they performed poorly and did not provide more than five correct answers in each of the nine conditions. The final analyses included 54 participants (26 women) aged 18–25 years ($M = 21.13$, $SD = 1.71$). All

TABLE 1 | Word properties (*M, SD*) for each manipulation group for experimental conditions (low or high) and control conditions (medium).

		Arousal					
		Low		Medium		High	
Subjective significance		*M*	*SD*	*M*	*SD*	*M*	*SD*
Low	Arousal	3.20	0.29	3.84	0.27	4.79	0.53
	Subjective significance	2.87	0.33	2.96	0.18	2.93	0.60
	Valence	5.25	0.48	5.11	0.44	5.01	0.63
	Concreteness	4.07	1.02	3.88	0.73	3.91	0.90
	LN frequency	6.04	1.86	6.32	1.41	5.95	1.36
	Length	5.73	2.09	6.20	1.32	6.07	2.02
Medium	Arousal	3.20	0.16	3.85	0.26	4.85	0.35
	Subjective significance	3.56	0.32	3.71	0.22	3.74	0.34
	Valence	5.42	0.47	5.38	0.62	5.10	0.66
	Concreteness	3.92	0.90	3.94	0.90	3.98	0.75
	LN frequency	6.50	1.52	6.29	1.58	5.68	1.91
	Length	6.27	2.09	6.13	1.68	6.87	2.45
High	Arousal	3.27	0.27	3.85	0.32	4.97	0.33
	Subjective significance	4.55	0.31	4.64	0.40	4.88	0.44
	Valence	5.38	0.36	5.41	0.35	5.31	1.11
	Concreteness	4.28	0.75	4.33	1.00	4.37	0.98
	LN frequency	6.99	2.02	7.08	1.23	6.55	1.90
	Length	6.47	2.13	6.67	1.95	6.87	2.00

participants were right handed and had normal or corrected to normal (by contact lens or glasses) vision. Before the experiment commenced, the participants were also assessed for normal color vision.

Procedure

Before the main experimental session, each participant filled out a socio-demographic questionnaire: age, sex, number of years of education, and academic field of interest. As a training session, each participant performed a standard Stroop task (Stroop, 1935) containing 20 trials (naming color bars squares displayed in one of the four target colors; and reading color meaning words and naming colors of font color-meaning words, both congruent and incongruent). Participants were encouraged to maximize speed of answering and accuracy at the same time. The training session ensured that participants understood the task and how to perform it correctly. Then the modified Stroop test was conducted as an experimental procedure. This test was presented to the participants as a set of 135 trials. Words appeared in a block design in fully random order across each of the nine conditions of the factorial manipulation. The block order was also fully random. A block design was chosen based on evidence showing that EST effects are especially visible in such types of presentations (c.f. Bar-Haim et al., 2007). The task was to indicate the font color in activation charged words. First, a fixation point was presented for a random time from 300 to 600 ms (with 10 ms intervals). This was applied to obviate preparation of expected time range event effects. Then randomly chosen words were displayed in the center of the screen. There was no time limit for response. After choosing the proper letter on the keyboard,

the word was replaced by a fixation point (+) for the trial to follow. The words in the entire experiment were presented to participants on a 15 inch monitor in 36 point size, Courier New font using E-Prime 2.0 software. For the entire time words were presented, four letters (P, C, Z, N), indicating possible answers, were displayed on the bottom of the screen.

Experiment 2 – Combined Stroop Task
Participants

In the second experiment, another 60 individuals (30 women) from different Warsaw universities (in equal proportion from the departments of social science, humanities, engineering, life science, and natural science) participated. The sample of 60 participants was planned in advance. Only correct answers were analyzed; thus two participants were excluded because they performed poorly and did not provide more than five correct answers in each of the nine conditions. The final analyses were conducted on 58 participants (29 women) aged 19–26 years ($M = 21.59$, $SD = 1.76$). All participants were right handed and had normal or corrected to normal (by contact lens or glasses) vision. Before the experiment, the participants were assessed for normal color vision.

Procedure

As in Experiment 1, participants filled out a socio-demographic questionnaire and then performed a training session with the standard Stroop task for 20 trials based on naming color bars, reading words displayed in black font and naming colors of font color-meaning words which were presented in the center of the screen. Participants were encouraged to work as quickly

and accurately as possible as both speed and accuracy were test variables.

Following the experimental session, a different version of the modified Stroop task was used. To create it, I used the paradigm modification introduced by Fackrell et al.,(2013). This combines classical Stroop task with EST. Target color-meaning words and emotional words were displayed together randomly on the screen 10% higher and lower than the center of the screen. Vertical display was used as Borkenau and Mauer (2006) showed that lateralized presentation elicited asymmetric processing. The task was to indicate the font color of color-meaning words. Each trial was prepared in only mismatch conditions (meaning and color were incongruent). No clue regarding what to do with the other words (displayed simultaneously in black font) was given to the participants. Non-color meaning words were displayed in a block design, with a fully random order of presentation inside each of the nine conditions of the factorial manipulation and sequence of blocks. The experimental session consisted of 135 trials, each timed in the same manner as in Experiment 1. First the fixation point was presented for a random time from 300 to 600 ms (with 10 ms intervals). Then randomly chosen words were displayed above or below the central location. There was no time limit for response. After choosing the proper letter on the keyboard, words were replaced by a fixation point (+) for the next trial. **Figure 3** presents the procedure used in Experiment 2.

The words in the entire experiment were presented to participants on a 15 inch monitor in 36 point size, Courier New font using E-Prime 2.0 software. For the entire time words were presented, four letters (P, C, Z, N) indicating possible answers were displayed on the bottom of the screen.

RESULTS

To investigate the impact of activation dimensions on performance in both modified Stroop tasks, a repeated measure 3×3 ANOVA was computed. Data were aggregated across the conditions for each subject. Each ANOVA was performed on a natural logarithm (LN) of reaction times. This procedure has been widely used in reaction times data and avoids problems with a right-skewed distribution (c.f. Heathcote et al., 1991). In **Table 2**, or in the text, raw reaction times are presented to facilitate better understanding of the observed differences.

Experiment 1 – Modified Stroop Task

The overall error rate in Experiment 1 was 4.92% and all error trials were excluded from further analysis. No significant main effect of arousal, $F(2,52) = 0.006$, $p = 0.99$, $\eta^2 < 0.001$, and no main effect of subjective significance, $F(2,52) = 1.44$, $p = 0.25$, $\eta^2 = 0.052$, were found. A statistically significant interaction effect of arousal and subjective significance was found, $F(4,50) = 3.21$, $p = 0.02$, $\eta^2 = 0.2$. **Table 2** presents mean reaction times and standard deviations for each of the experimental groups.

To explore the interaction effect obtained, four additional repeated measures ANOVAs were conducted for each of the subjective significance levels (three analyses of simple main effects of arousal) and for the high arousal level (one analysis of simple main effect of subjective significance). The reason for choosing these simple main effects was based on theoretical expectations connected with the fact that slowdown in reaction times should be observed only for high arousal stimuli. The Holm correction for multiple comparisons was applied. This is a
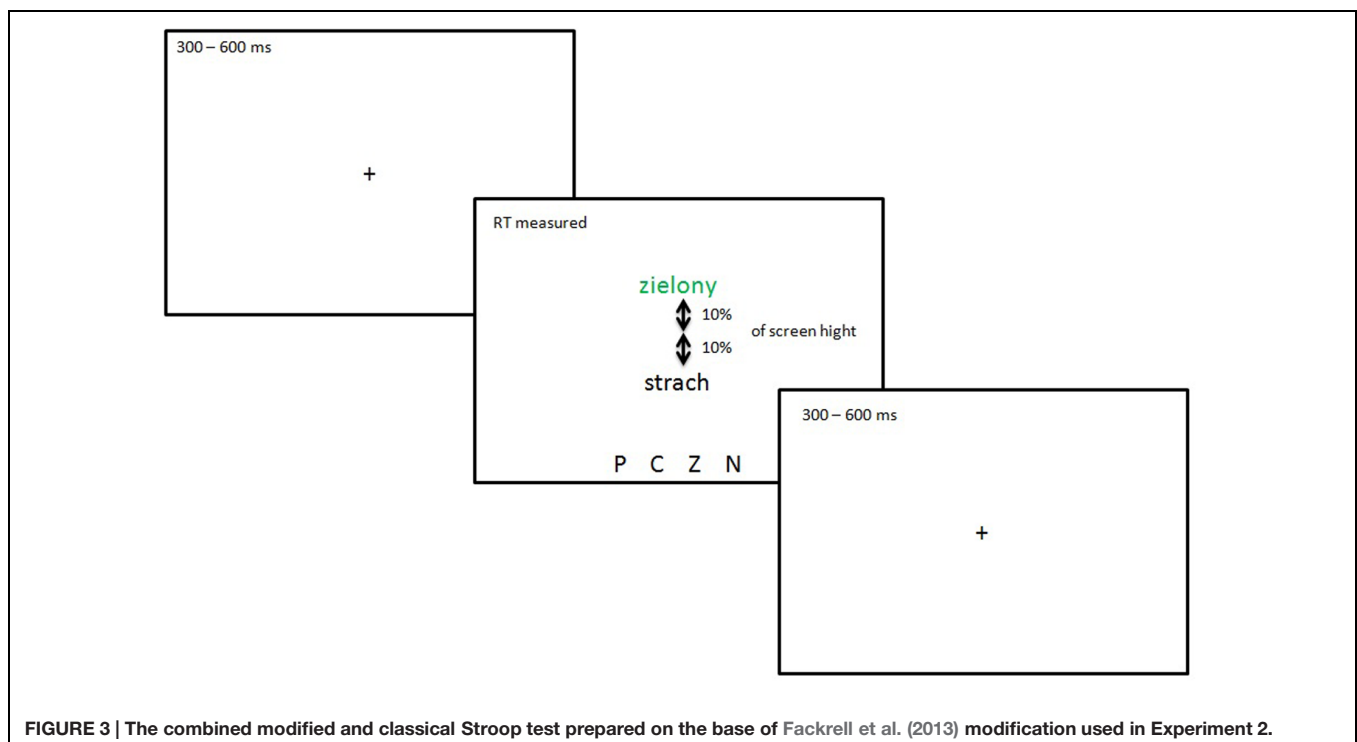

FIGURE 3 | The combined modified and classical Stroop test prepared on the base of Fackrell et al. (2013) **modification used in Experiment 2.**

TABLE 2 | Mean reaction times (in milliseconds) and standard deviations for each of the experimental conditions.

		Arousal		
		Low	Medium	High
		Experiment 1 – Modified Stroop Task		
Subjective significance	Low	848 (256)	826 (231)	763 (212)
	Medium	824 (258)	827 (288)	896 (235)
	High	799 (244)	812 (235)	822 (247)
		Experiment 2 – Combined Stroop Task		
Subjective significance	Low	1007 (344)	1019 (251)	1015 (283)
	Medium	994 (262)	981 (249)	1084 (251)
	High	1043 (263)	1058 (262)	997 (299)

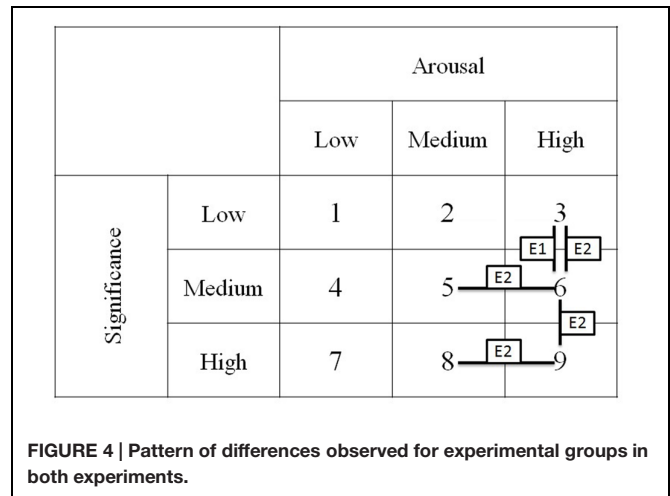**FIGURE 4 | Pattern of differences observed for experimental groups in both experiments.**

sequentially rejective version of the simple Bonferroni correction for multiple comparisons; thus, one has to divide the critical p-value by the number of tests performed at each stage of analysis. In this regard, we may assume that the critical p-values were as follows: for the first detected difference $p < 0.0125$ (where four tests were compared); for the second effect $p < 0.016$ (for three performed tests); and, for the third detected difference $p < 0.025$ (for two conducted tests). In each case, the difference contrast was applied to check for effects between low and medium or medium and high manipulating factor groups for each variable. To make the planned analyses more visible, in **Figure 4** one can find experimental design, numbers of manipulation conditions used in text description, and significant contrasts for both experiments.

Taking into account subjective significance simple main effect for high arousal stimuli (groups 3, 6, and 9), a statistically significant difference was found, $F(2,52) = 6.52$, $p = 0.003$, $\eta^2 = 0.2$. Difference contrast analysis showed statistically significant differences between groups of low (3) and medium (6) subjective significance stimuli among high arousal ones, $F(1,53) = 13.22$, $p = 0.001$, $\eta^2 = 0.2$. The remaining difference contrasts were not statistically significant and were not reported.

Taking into account arousal effects, a no statistically significant simple main effect for low subjective significant stimuli (groups 1, 2, and 3) was found, $F(2,52) = 3.70$, $p = 0.03$, $\eta^2 = 0.13$. Difference contrast analysis showed a statistically significant difference between groups of medium (2) and high (3) arousal level stimuli among low subjective significance ones, $F(1,53) = 6.75$, $p = 0.012$, $\eta^2 = 0.11$. No statistically significant simple main effect of arousal either in the case of medium subjective significance stimuli (see groups 4, 5, and 6), $F(2,52) = 2.53$, $p = 0.09$, $\eta^2 = 0.09$, or in the case of high subjective significance stimuli (groups 7, 8, and 9), $F(2,52) = 0.38$, $p = 0.7$, $\eta^2 = 0.014$, were found. No other difference contrast was found to be statistically significant and, thus, were not reported.

Experiment 2 – Combined Stroop Task

The overall error rate in Experiment 2 was 7.53% and all error trials were excluded from further analysis. Neither a significant main effect of arousal, $F(2,56) = 0.44$, $p = 0.64$, $\eta^2 = 0.016$

nor a main effect of subjective significance, $F(2,56) = 0.64$, $p = 0.53$, $\eta^2 = 0.022$, were found. A statistically significant interaction effect of arousal and subjective significance was found, $F(4,54) = 4.22$, $p = 0.005$, $\eta^2 = 0.24$. **Table 2** presents mean reaction times and standard deviations for each of the manipulation groups in Experiment 2.

To explore interaction effects, four additional repeated measures ANOVAs for each of the subjective significance levels and for high arousal level words were conducted. Holm correction for multiple comparisons was applied. In the case of subjective significance simple main effect in high arousal words (groups 3, 6, and 9) a statistically significant effect was found, $F(2,56) = 4.51$, $p = 0.015$, $\eta^2 = 0.14$. Difference contrast analysis showed statistically significant differences between groups of low (3) and medium (6) subjective significance stimuli, $F(1,57) = 4.61$, $p = 0.036$, $\eta^2 = 0.08$, as well as between groups of medium (6) and high (9) subjective significance stimuli, $F(1,57) = 4.05$, $p = 0.049$, $\eta^2 = 0.07$, among high arousal ones.

Taking into account arousal simple main effects, a no statistically significant effect for low subjective significance stimuli (groups 1, 2, and 3) was found, $F(2,56) = 0.76$, $p = 0.7$, $\eta^2 = 0.01$. A statistically significant effect in the case of medium subjective significance stimuli (see groups 4, 5, and 6) was found, $F(2,56) = 6.3$, $p = 0.003$, $\eta^2 = 0.18$. Difference contrast analysis showed a statistically significant difference between groups of medium (5) and high (6) arousal level stimuli among medium subjective significance ones, $F(1,57) = 12.43$, $p = 0.001$, $\eta^2 = 0.18$. Finally, a no statistically significant effect in the case of high subjective significance stimuli (groups 7, 8, and 9) was found, $F(2,56) = 2.32$, $p = 0.11$, $\eta^2 = 0.07$. Difference contrast analysis showed a statistically significant difference between groups of medium (8) and high (9) arousal level stimuli among high subjective significance ones, $F(1,57) = 4.58$, $p = 0.037$, $\eta^2 = 0.07$. No other difference contrasts were found to be statistically significant and were not reported.

Additional Analysis of Word Properties

The results obtained suggested checking whether observed results concerning medium and high arousal stimuli could be

derived from subtle differences in word properties among each manipulation conditions. To do so, a *t*-test for independent samples taking into account manipulated (arousal and subjective significance) and controlled (concreteness, valence, LN of frequency and length) aspects was conducted. Analyses comparing each of depicted in **Figure 4** as statistically significant contrast effects, namely, as between manipulation conditions 5 and 6, 8 and 9, 3 and 6 as well as 6 and 9 were conducted. For all of these described comparisons, no significant contrasts among controlled variables were found whereas expected significant ($p < 0.05$) contrasts among manipulated variables were found; namely conditions 5 and 6, 8 and 9 differed in the case of arousal ratings whereas conditions 3 and 6, 6 and 9 differed in the case of subjective significance ratings.

DISCUSSION

This is the first study to combine a duality of mind perspective on activation mechanisms in order to investigate the manner in which arousal and subjective significance shape cognitive control in the case of interference control in the modified Stroop task. Both experiments presented were based on carefully chosen verbal material, contrasting in arousal and subjective significance ratings, but matched in the case of many potentially important variables such as valence, concreteness (c.f. Siakaluk et al., 2014), frequency and length (c.f. Burt, 2002).

In general, no main effects for either variable were found, but an interaction between them was found. In both experiments, differences mostly concerned groups of high arousal words. This pattern of results, perhaps, is due to carefully chosen materials. In some way, this also confirms the validity of proposed factors impacting on Stroop task performance. Simply, it is probable that the main effect may disappear when a new dimension of subjective significance is controlled. An alternative explanation for the lack of arousal effect may be that words included in the lists were in fact moderate arousal ones taking into account a nine-point Likert scale (c.f. **Table 1**). Simply, it is possible that the slowdown could be observed better for higher arousal levels. In the current study, words were chosen in a way that allowed the comparison of three different, increasing levels of arousal, but the highest one was at least moderate. Unfortunately, at this stage, due to both dimensions correlations and correlations between them and controlled variables (Imbir, 2015), it was the only way to prepare a list allowing for the manipulation of both arousal and subjective significance, at the same time as controlling other potentially important factors (valence, concreteness, frequency, and length).

In Experiment 1, slowdown for (relatively) high arousal words was reduced in the low subjective significance group as compared with the moderate group. In Experiment 2, slowdown for (relatively) high arousal words was reduced in both low and high subjective significance groups in comparison with moderate groups and, in both cases, results were statistically significant (c.f. **Figure 4** and **Table 2**). This indicated that the presence of subjective significance factors (low or high) neutralized arousal impact. Subjective significance presence could have influenced

cognitive control, motivating and enhancing resources needed for the controlled target task of naming the color of ink (c.f. **Figure 2**). It is interesting that the effect was observed both when the explicit task concerned activation charged words (Experiment 1) and when words were not the subject of the task (Experiment 2). The effects observed in low and high subjective significance groups suggest that the construct of subjective significance is not in fact analogical to unimodal arousal construct, but represents rather bimodal structure (negation and affirmation of subjective significance). The modified Stroop task used in current studies does not allow for the processing of information in a reflective way (Strack and Deutsch, 2014) mostly because quick answers are required and task specificity, namely, the explicit task, is not to read the words, but to ignore their content. For that reason, the results showed no effects of negation. In fact, low and high significant words produced similar outcomes for task performance and lack of significance factor (in moderate groups) resulting in the slowdown observed for high arousal words. Further research is needed to understand the mechanisms of slowdown reduction in a modified Stroop task by subjective significance, but the effects demonstrated in this paper cannot be attributed to mismatching word groups in crucial dimensions such as valence, concreteness, frequency, and length.

Differences in response times in Experiments 1 and 2 were observed. Participants in the first experiment pressed response keys more quickly than in the second experiment. This was due to the nature of the task. In the first experiment, the task was simple because the target stimuli appeared in the center of the screen. In the case of the modified EST, on each occasion, the participants had to find the target stimuli either above or below the fixation point and, thus, had to take more time. The decision to display words in Experiment 2 above or below the fixation point (based on Fackrell et al., 2013) obviated potential lateralization effects demonstrated earlier (c.f. Borkenau and Mauer, 2006).

CONCLUSION

The current research showed a new phenomenon concerning Stroop task performance. This was based on a duality of mind approach, which distinguishes between two mechanisms of activation specific to non-verbal experiential system processing and verbalized, rational, and propositional system processing. The Stroop task is a good example of interference between both processes contributing to behavior. Recent findings have shown that valenced or arousal words cause slowdown in response times for ink-color naming task. This study showed that the activation mechanism specific to the controlled part of the Stroop task can reduce arousal level effects.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and approved it for publication.

FUNDING

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpsyg.2016.00075

REFERENCES

Bar-Haim, Y., Lamy, D., Pergamin, L., Bakermans-Kranenburg, M. J., and Van Ijzendoorn, M. H. (2007). Threat-related attentional bias in anxious and nonanxious individuals: a meta-analytic study. *Psychol. Bull.* 133, 1–24. doi: 10.1037/0033-2909.133.1.1

Baumeister, R. F., Bratslavsky, E., Muraven, M., and Tice, D. M. (1998). Ego depletion: is the active self a limited resource? *J. Pers. Soc. Psychol.* 74, 1252–1265. doi: 10.1037/0022-3514.74.5.1252

Borkenau, P., and Mauer, N. (2006). Processing of pleasant, unpleasant, and neutral words in a lateralized Emotional Stroop Task. *Cogn. Emot.* 20, 866–877. doi: 10.1080/02699930500348131

Burt, J. S. (2002). Why do non-color words interfere with color naming? *J. Exp. Psychol. Hum. Percept. Perform.* 28, 1019–1038. doi: 10.1037/0096-1523.28.5.1019

Compton, R. J., Banich, M. T., Mohanty, A., Milham, M. P., Herrington, J., Miller, G. A., et al. (2003). Paying attention to emotion: an fMRI investigation of cognitive and emotional stroop tasks. *Cogn. Affect. Behav. Neurosci.* 3, 81–96. doi: 10.3758/CABN.3.2.81

Cooper, R. P. (2010). Cognitive control: componental or emergent? *Top. Cogn. Sci.* 2, 598–613. doi: 10.1111/j.1756-8765.2010.01110.x

Damasio, A. (2010). *Self Comes to Mind: Constructing the Conscious Mind*. New York, NY: Pantheon.

Deutsch, R., Gawronski, B., and Strack, F. (2006). At the boundaries of automaticity: negation as reflective operation. *J. Pers. Soc. Psychol.* 91, 385–405. doi: 10.1037/0022-3514.91.3.385

Dresler, T., Mériau, K., Heekeren, H. R., and Van der Meer, E. (2009). Emotional Stroop task: effect of word arousal and subject anxiety on emotional interference. *Psychol. Res.* 73, 364–371. doi: 10.1007/s00426-008-0154-6

Epstein, S. (2003). "Cognitive-experiential self-theory of personality," in *Handbook of Psychology VOL5: Personality and Social Psychology*, eds T. Millon and M. J. Lerner (Hoboken, NJ: John Wiley & Sons, Inc).

Fackrell, K., Edmondson-Jones, M., and Hall, D. A. (2013). A controlled approach to the emotional dilution of the stroop effect. *PLoS ONE* 8:e80141. doi: 10.1371/journal.pone.0080141

Frijda, N. H. (2007). *The Laws of Emotion*. Mahwah, NJ: Laurence Erlbaum Association.

Gawronski, B., and Creighton, L. A. (2013). "Dual-process theories," in *The Oxford Handbook of Social Cognition*, ed. D. E. Carlston (New York, NY: Oxford University Press), 282–312.

Heathcote, A., Popiel, S. J., and Mewhort, D. J. (1991). Analysis of response time distributions: an example using the Stroop task. *Psychol. Bull.* 109, 340–347. doi: 10.1037/0033-2909.109.2.340

Imbir, K. (2015). Affective Norms for 1,586 Polish Words (ANPW): duality of mind approach. *Behav. Res. Methods* 47, 860–870. doi: 10.3758/s13428-014-0509-4

Imbir, K., and Jarymowicz, M. (2013). The Effect of Automatic vs. Reflective Emotions on cognitive control in antisaccade tasks and the emotional Stroop test. *Pol. Psychol. Bull.* 44, 137–146. doi: 10.2478/ppb-2013-0016

Jarymowicz, M., and Imbir, K. (2015). Toward a human emotions taxonomy (based on their automatic vs. reflective origin). *Emot. Rev.* 7, 183–188. doi: 10.1177/1754073914555923

Juvina, I. (2011). Cognitive control: componental and yet emergent. *Top. Cogn. Sci.* 3, 242–246. doi: 10.1111/j.1756-8765.2011.01144.x

Kagan, J. (2007). *What is Emotion? History, Measures, and Meanings*. New Haven, CT: Yale University Press.

Kahneman, D. (2003). A perspective on judgment and choice: mapping bounded rationality. *Am. Psychol.* 58, 697–720. doi: 10.1037/0003-066X.58.9.697

Kahneman, D. (2011). *Thinking. Fast and Slow*. New York, NY: Farrar, Straus and Giroux.

Kahnt, T., and Tobler, P. N. (2013). Salience signals in the right temporoparietal junction facilitate value-based decisions. *J. Neurosci.* 33, 863–869. doi: 10.1523/JNEUROSCI.3531-12.2013

Kazojć, J. (2011). *Słownik Frekwencyjny Języka Polskiego [Polish Language Dictionary of Attendance]*. Available at: http://www.elektronikjk.nazwa.pl/slowniki/lista-frekwencyjna.pdf

Lang, P. J. (1980). "Behavioral treatment and bio-behavioral assessment: computer applications," in *Technology in Mental Health Care Delivery Systems*, eds J. B. Sidowski, J. H. Johnson, and T. A. Williams (Norwood, NJ: Ablex), 119–l37.

Larsen, R. J., Mercer, K. A., and Balota, D. A. (2006). Lexical characteristics of words used in emotional Stroop experiments. *Emotion* 6, 62–72. doi: 10.1037/1528-3542.6.1.62

McKenna, F. P., and Sharma, D. (1995). Intrusive cognitions: an investigation of the emotional Stroop task. *J. Exp. Psychol. Learn. Mem. Cogn.* 21, 595–607. doi: 10.1037/0278-7393.21.6.1595

McKenna, F. P., and Sharma, D. (2004). Reversing the emotional Stroop effect reveals that it is not what it seems: the role of fast and slow components. *J. Exp. Psychol. Learn. Mem. Cogn.* 30, 382–392. doi: 10.1037/0278-7393.30.2.382

Nigg, J. T. (2000). On inhibition/disinhibition in developmental psychopathology: views from cognitive and personality psychology and a working inhibition taxonomy. *Psychol. Bull.* 126, 220–246. doi: 10.1037/0033-2909.126.2.220

Petty, R. E., and Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. *Adv. Exp. Soc. Psychol.* 19, 123–205. doi: 10.1016/S0065-2601(08)60214-2

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychol. Rev.* 110, 145–172. doi: 10.1037/0033-295X.110.1.145

Schneider, W., and Chein, J. M. (2003). Controlled & automatic processing: behavior, theory, and biological mechanisms. *Cogn. Sci.* 27, 525–559. doi: 10.1207/s15516709cog2703_8

Siakaluk, P. D., Knol, N., and Pexman, P. M. (2014). Effects of emotional experience for abstract words in the Stroop task. *Cogn. Sci.* 38, 1698–1717. doi: 10.1111/cogs.12137

Strack, F., and Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Pers. Soc. Psychol. Rev.* 8, 220–247. doi: 10.1207/s15327957pspr0803_1

Strack, F., and Deutsch, R. (2014). "The reflective–impulsive model," in *Dual-Process Theories of the Social Mind*, eds J. W. Sherman, B. Gawronski, and Y. Trope (New York, NY: Guilford Press), 92–104.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *J. Exp. Psychol.* 18, 643–662. doi: 10.1037/h0054651

Watts, F. N., McKenna, F. P., Sharrock, R., and Trezise, L. (1986). Colour naming of phobia-related words. *Br. J. Psychol.* 77, 97–108. doi: 10.1111/j.2044-8295.1986.tb01985

Williams, J. M. G., Mathews, A., and MacLeod, C. (1996). The emotional Stroop task and psychopathology. *Psychol. Bull.* 120, 3–24. doi: 10.1037/0033-2909.120.1.3

Wilson, T. D., Lindsey, S., and Schooler, T. Y. (2000). A model of dual attitudes. *Psychol. Rev.* 107, 101–126. doi: 10.1037/0033-295X.107.1.101

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read,
for greatest visibility

COLLABORATIVE PEER-REVIEW
Designed to be rigorous
– yet also collaborative,
fair and constructive

FAST PUBLICATION
Average 85 days from
submission to publication
(across all journals)

COPYRIGHT TO AUTHORS
No limit to article
distribution and re-use

TRANSPARENT
Editors and reviewers
acknowledged by name
on published articles

SUPPORT
By our Swiss-based
editorial team

IMPACT METRICS
Advanced metrics
track your article's impact

GLOBAL SPREAD
5'100'000+ monthly
article views
and downloads

LOOP RESEARCH NETWORK
Our network
increases readership
for your article

Find us on