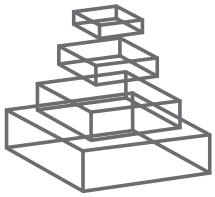
frontiers

RESEARCH TOPICS

CONSCIOUSNESS AND NEURAL PLASTICITY

Hosted by
Morten Overgaard and Mads Jensen

frontiers in PSYCHOLOGY



frontiers

Cover image provided by Ibbl sarl, Lausanne CH

ISSN 1664-8714
ISBN 978-2-88919-007-2
DOI 10.3389/978-2-88919-007-2

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

CONSCIOUSNESS AND NEURAL PLASTICITY

Hosted By
Morten Overgaard, Aalborg University, Denmark
Mads Jensen, Aarhus University Hospital, Denmark

In contemporary consciousness studies the phenomenon of neural plasticity has received little attention despite the fact that neural plasticity is of still increased interest in neuroscience. We will, however, argue that neural plasticity could be of great importance to consciousness studies. If consciousness is related to neural processes it seems, at least prima facie, that the ability of the neural structures to change should be reflected in a theory of this relationship "Neural plasticity" refers to the fact that the brain can change due to its own activity. The brain is not static but rather a dynamic entity, which physical structure changes according to its use and environment. This change may take the form of growth of new neurons, the creation of new networks and structures, and change within network structures, that is, changes in synaptic strengths.

Plasticity raises questions about the relation between consciousness and brain functions. If consciousness is connected to specific brain structures (as a function or in identity) what happens to consciousness when those specific underlying structures change? It is therefore possible that the understanding and theories of neural plasticity can have direct consequences for a theory about consciousness. For instance, theories of strict identity between consciousness and structure may face the serious dilemma to either accept that, say, the experience of the colour red is fundamentally different in one individual over time due to cortical changes or to abandon the strong identity thesis altogether.

Were one to pursue a theory according to which consciousness is not an epiphenomenon to brain processes, consciousness may in fact affect its own neural basis. The neural correlate of consciousness is often seen as a stable structure, that is, something that is stable over time.

Considering neural plasticity, this is not necessarily so. The NCC might change and hence literally change the way a person is conscious. What it is about the NCC that can and might change is, even though it can be relevant for the relation between the brain and consciousness is, still an unanswered question.

There are, hence, a lot of questions that might shed light upon the relevant but unknown relations between consciousness and the brain. Therefore, We hereby propose to do a special issue on consciousness and neural plasticity to shed light on these underestimated issues.

Table of Contents

frontiers in
PSYCHOLOGY

EDITORIAL
published: 01 September 2011
doi: 10.3389/fpsyg.2011.00191

Neural plasticity and consciousness

Mads Jensen* and Morten Overgaard

Cognitive Neuroscience Research Unit, Hammel Neurorehabilitation and Research Center, Aarhus University Hospital, Aarhus, Denmark
**Correspondence: mads@cnru.dk*

Neural plasticity has, despite a growing interest in neuroscience over the last decades, received little interest in consciousness research. This might be because neural plasticity has no clear and strict definition; it is used to describe the ability of neuronal tissue to adjust activities and physical characteristics in order to adapt to changes in the environment or changes in their use – in relation to, e.g., behavioral patterns, cognition, or metabolic demands. The term traces back in history and was discussed by William James (1910–1942) and Ramón y Cajal (1852–1934) who provided evidence to argue for local plasticity in structure and function in otherwise fixed neural circuits.

Neural plasticity covers a number of different types of changes in brain tissue or organization, which arguably can be described at a high or a low level and at different time scales. At "lower" levels of description, a very simple example of plasticity is facilitation, i.e., the increase in amplitude of a postsynaptic potential after rapid activations from other, presynaptic neurons. Long-term potentiation (LTP), being the persistency of the postsynaptic potential after the facilitating stimulus has subsided, may lead to increased postsynaptic activities for days. Other examples of "low level processes" that could be mentioned are, changes in individual neurons or small circuits involve synaptogenesis (the creation of individual or groups of synapses to build connections between neurons) and neuronal migration (where neurons "travel" to become part of circuits distant from their site of "birth").

At a "higher" level of description, reorganizations related to changes in cognition may be investigated. In cognitive neuroscience, an impressive amount of evidence suggests a functional localization of mental phenomena in the brain, so that different cognitive and emotional functions are associated with specific networks or regions of the brain (e.g., Frackowiak et al., 2003). Functional reorganization involves a change in how, for example, a specific cognitive function is realized in the brain. Such a change will in some instances be the result of neural repair mechanisms following an injury to the brain. In such cases, functional reorganizations will typically be considered cases of functional recovery where a wide range of functions (e.g., language) seemingly are realized in different brain regions after injury to regions previously associated with the function (Overgaard and Mogensen, 2011).

The concept of consciousness is even more complicated and has even more different interpretations than the concept of neural plasticity. We, like several others in this field of research, take the concept to mean subjective or phenomenal experience. One primary question very frequently asked among those interested in consciousness is which specific neural regions or networks that are necessary and/or sufficient for a person to be conscious. This is referred to as the neural correlate of consciousness (NCC).

The NCC is traditionally defined as "the minimal set of neuronal events and mechanisms jointly sufficient for a specific conscious percept" (Koch, 2004, p. 16) or "A neural system N is an NCC if the state of N correlates directly with states of consciousness" (Chalmers, 2000, p. 18). At a first look, such a statement claims nothing more than a neural system N at a given point in time correlates directly with states of consciousness. However, we assume the claim is intended to be stronger than that, namely that an NCC in some fundamental aspects is the same over time and between individuals. In other words, the NCC for a conscious experience X is not just the NCC at *a* time – it is the NCC at *any* time.

Research on neural plasticity can be taken to imply that a simple one-to-one mapping between brain structure and mental phenomenon might not be an optimal strategy in the attempt to reveal the neural underpinnings of consciousness. If it proves to be the case that "one and the same" mental state may be related to several different physical processes, an explanation should rather look for common factors among different sufficient "realisers" of consciousness than it would merely identify brain regions.

A classical interpretation of the relation between conscious experience (qualia) and brain processes has been that qualia are determined by the brain activity, such that if the brain activity had been different then the conscious experience would have been different as well. Such an interpretation, in the light of research on neural plasticity, seems very difficult. Rather, research on neural plasticity seems more in support of a weaker view on localization, arguing that mechanisms or cognitive structure is "realized in" brain activity or brain structure, which is not restricted to a certain localization.

The concept of neural plasticity captures a diverse and complex range of phenomena, as argued above, with different levels of explanation. The integration of theories about neural plasticity at different levels and theories of consciousness is therefore complex, carrying important implications for the plausibility of any conception of the relationship of consciousness and brain processes. With this introductory paper, we hope to start an otherwise neglected debate, unfolded in the Special Topic "Consciousness and Neural Plasticity."

ACKNOWLEDGMENTS

Mads Jensen was supported by a grant from The Danish Council for Independent Research | Humanities. Morten Overgaard was supported by a starting grant from The European Research Council.

REFERENCES

Chalmers, D. J. (2000). "What is a neural correlate of consciousness?" in *Neural Correlates of Consciousness*, ed. T. Metzinger (Cambridge: MIT Press), 17–39.

Frackowiak, R., Friston, K. J., Frith, C., Donal, R., Price, C. J., Zeki, S., Ashburner, J., and Penny, W. D. (2003). *Human Brain Function*, 2nd Edn. San Diego: Academic Press.

Koch, C. (2004). *The Quest for Consciousness: A Neurobiological Approach.* Englewood, CO: Roberts and Co.

Overgaard, M., and Mogensen, J. (2011). A framework for the study of multiple realizations: the importance of levels of analysis. *Front. Psychol.* 2:79. doi: 10.3389/fpsyg.2011.00079

Received: 20 June 2011; accepted: 27 July 2011; published online: 01 September 2011.

Citation: Jensen M and Overgaard M (2011) Neural plasticity and consciousness. Front. Psychology ***2****:191. doi: 10.3389/fpsyg.2011.00191*

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

frontiers in
PSYCHOLOGY

REVIEW ARTICLE
published: 20 January 2011
doi: 10.3389/fpsyg.2011.00006

What does neural plasticity tell us about role of primary visual cortex (V1) in visual awareness?

Juha Silvanto[1]* and Geraint Rees[2,3]*

[1] *Brain Research Unit, Low Temperature Laboratory and Advanced Magnetic Imaging Centre, School of Science and Technology, Aalto University, Espoo, Finland*
[2] *Institute of Cognitive Neuroscience, University College London, London, UK*
[3] *Wellcome Trust Centre for Neuroimaging, University College London, London, UK*

Edited by:
Morten Overgaard, Aarhus University, Denmark

Reviewed by:
Niko Busch, Humboldt Universität Berlin, Germany
Søren Kyllingsbaek, University of Copenhagen, Denmark

****Correspondence:***
Juha Silvanto, Brain Research Unit, Low Temperature Laboratory and Advanced Magnetic Imaging Centre, School of Science and Technology, Aalto University, Espoo, Finland.
e-mail: silvanto@neuro.hut.fi;
Geraint Rees, Institute of Cognitive Neuroscience, University College London, 17 Queen Square, London WC1N 3AR, UK.
e-mail: g.rees@fil.ion.ucl.ac.uk

The complete loss of visual awareness resulting from a lesion to the primary visual cortex (V1) suggests that this region is indispensable for conscious visual perception. There are however a number cases of conscious perception in the absence of V1 which appear to challenge this conclusion. These include reports of patients with bilateral V1 lesions sustained at an early age whose conscious vision has spontaneously recovered, as well as stroke patients who have recovered some conscious vision with the help of rehabilitation programs. In addition, the phenomenon of hemianopic completion and percepts induced by brain stimulation suggest that V1 may not be necessary for conscious perception in all circumstances. Furthermore, that the visual abilities in the cat are associated with the recovery of normal extrastriate tuning properties rather than emulation of V1 functions suggests that there is nothing unique about the functional properties of this region in visual awareness. Rather, the dramatic effect of a V1 lesion on visual awareness may be due to its role in providing the majority of extrastriate visual input, the loss of which abolishes normal neural responsiveness throughout the visual cortex.

Keywords: visual awareness, plasticity, V1, blindsight, transcranial magnetic stimulation, consciousness, extrastriate

A lesion to the primary visual cortex (V1) abolishes all conscious visual perception in the corresponding part of the visual field. However, the ability of some patients to unconsciously detect and discriminate stimuli within the field defect (see Cowey, 2010 for reviews of "blindsight"), implies that the loss of perceptual experience is not simply due to information failing to reach the visual system, and strongly implicate V1 in visual awareness. The role of V1 in determining the content of subjective awareness is also supported by studies carried out in neurologically normal observers (see Tong, 2003; Ro, 2010 for reviews).

There are however a number cases of conscious visual perception in the absence of V1 which appear to challenge the view that this region plays an indispensable role in visual awareness. These include reports of children born without V1, stroke patients who have recovered some conscious vision with the help of rehabilitation programs, as well as the phenomenon of hemianopic completion and percepts induced by brain stimulation. Such cases are important to the study of neural correlates of consciousness (NCC): if conscious perception of a given attribute is initially abolished by a lesion but subsequently recovers, this indicates that the NCC of that attribute is not fixed to the affected brain region. In this review we will consider the implications of plasticity to understanding the role of V1 in visual awareness.

CONSCIOUS PERCEPTION AFTER A BILATERAL V1 LESION

For the present discussion, the critical question is the extent to which visual awareness can recover in the complete absence of V1, as this would indicate that V1 is not indispensable for awareness. Cases of bilateral V1 lesion are thus of paramount interest. However, reports of complete bilateral lesions of occipital cortex are rare, as in many cases some occipital cortex may remain functional, which complicates interpretation. For example, Bova et al. (2008), reported a patient who suffered bilateral occipital lobe infarction at the age of 2 years 6 months. In the acute phase, his visual behavior was consistent with complete bilateral visual loss; fixation was absent, there was no blink response to threat, and he was only able to recognize objects and family members using compensatory (touch and auditory) abilities. However, 18 months after the lesion, he could see and reach small objects on the ground, name colors, and recognize himself and his relatives even in photographs. At 6 years 8 months, visual recognition acuity was 10/10 in both eyes and neuro-ophthalmological examination was normal, except for persistence of the visual field defect in the upper hemifield and a selective impairment for complex visuospatial skills. This case appears to suggest that conscious perception can almost fully recover in the absence of V1. However, MRI scan suggested a partial sparing of the striate cortex, and this was confirmed by the presence of some residual occipital response of the VEPs, suggesting the presence of some functional striate cortex. Therefore, while the presence of some residual V1 may not fully explain conscious visual perception in this patient, it may have facilitated recovery (see, e.g., Payne and Rushmore (2003)) and complicates the implications of this study for understanding the role of V1 in visual awareness. The vast majority of V1 lesions are incomplete, and thus the number of cases from which strong theoretical conclusions can be drawn is limited.

There are however few cases in which the absence of early visual cortex has been verified. Amicuzi et al. (2006) reported the case of a 5-year-old girl in whom ultrasonography 16 days after birth showed an enlargement of the occipital horns of the lateral ventricles, mostly on the left side. At the age of 5, MRI scans disclosed an

absence of occipital pole bilaterally, as well as of occipitotemporal regions in the right hemisphere and occipitoparietal regions in the left hemisphere. Functional examinations revealed that basic visual abilities such as visual acuity, contrast sensitivity and visual field were present although somewhat compromised. On visual perceptual assessment, a selective impairment of figure-ground segregation was found, whereas color, form, and orientation discrimination were almost normal. Visual recognition impairment was present for objects, faces, actions, and scenes. While the neural correlates of conscious vision in this subject are unclear due to the lack of neuroimaging evidence, it is clear that it must be the remaining extrastriate cortex which enables conscious perception to arise. This was supported by diffusion tensor imaging (DTI) evidence (Amicuzi et al., 2006) which highlighted connections from lateral geniculate nucleus to the intact extrastriate cortex which bypass the damaged areas. Expansion of pathways that can bypass V1 and directly connect subcortical nuclei with extrastriate visual structures is believed to be critical in neural plasticity following the loss of V1 and may also mediate blindsight (see, e.g., Cowey, 2004; Payne and Lomber, 2002).

A second intriguing case involves a 21-year-old male with extensive bilateral damage present since birth (Giaschi et al., 2003). This case is highly informative because the absence of the striate cortex was confirmed with anatomical MRI scans as well as using and functional MRI and high-resolution EEG. Furthermore, the neural basis of conscious perception was investigated using fMRI.

The visual impairment, likely to be due to hypoxic brain insult at birth, was severe. At the age of 2 years, he was diagnosis of congenital cortical visual impairment. Full-field flash stimuli evoked no potentials from the occipital cortex and no alpha rhythm and or responses on eye opening and closure was seen in EEG. When examined at the age of 21, he could identify a few simple shapes, by scanning them repeatedly close to his eyes. The patient was able to name the colors of large stimuli, but he made many errors in color discrimination on the Quantitative Color Vision Test. He could appreciate size differences and count coins, without touch, only when they were widely separated. The patients was however able to detect stimuli moving at high speeds; at low speeds (below 3.5°/s), his accuracy fell to chance performance. Importantly, he reported that he could "see" the motion, indicating that his motion detection ability was not blindsight; his detection ability was accompanied by phenomenal awareness. The neural basis of this motion perception was investigated with fMRI. Although the motion-selective area V5/MT appeared to be anatomically intact, no functional activation was observed in this region. Instead, fast radial motion activated regions in right premotor cortex, right medial posterior cingulate, right medial precuneus, left anterior superior temporal gyrus, and left and right posterior superior temporal sulcus. No activation was induced by the slow stimuli.

This case is interesting for several reasons. Firstly, although conscious residual vision for motion in the absence of awareness of other stimulus attributes has been previously reported (Riddoch, 1917), this is the only reported case in which the absence of striate cortex has been verified. Secondly, this case provides both anatomical and neuroimaging evidence indicating that some visual perception is possible in the *complete* absence of V1; no striate cortex tissue was visible on anatomic MRI scan, and no evoked responses were found with EEG or fMRI. Thirdly, it appears that parietal and temporal regions gave rise to conscious perception in this subject.

PLASTICITY FOLLOWING UNILATERAL V1 LESIONS

While the evidence from bilateral occipital lesions demonstrates that conscious vision can recover in the complete absence of V1, the neural basis of that recovery is unclear. More is known about the development of new visual field maps accompanying visual function recovery after unilateral V1 lesions. Muckli et al. (2009) report a subject whose development of the right cerebral hemisphere terminated before the 7th week of embryonic gestation. Despite the complete loss of right hemisphere at birth, the patient had close to normal vision in both hemifields and successfully mastered activities requiring bilateral coordination such as roller skating and bike riding. fMRI mapping revealed that the patient's remaining hemisphere contained visual field representations of not only he contralateral visual hemifield (as is the case in neurologically normal development) but, surprisingly, also maps of the ipsilateral (left) visual hemifield: islands of ipsilateral visual field representations were located along the representations of the vertical meridian. In V1, smooth and continuous maps from contra- and ipsilateral hemifield overlap each other, whereas in ventral V2 and V3 ipsilateral quarter field representations invaded small distinct cortical patches. This is significant because normal subjects do not show ipsilateral representations in early visual areas.

Brain plasticity tends to be more evident after congenital or early acquired brain injury (see, e.g., Werth, 2008; Guzzetta et al., 2010 for reviews). While potential for plasticity is greatly reduced in adulthood, the fact that visual field rehabilitation protocols have at least limited success in shrinking visual field defects (e.g., Zihl and von Cramon, 1986; Kerkhoff et al., 1994; Kasten et al., 1997; Schmielau and Wong, 2007) indicates that some functionally significant neural plasticity does occur. Furthermore there is also electrophysiological evidence of plastic changes in the adult monkey visual cortex. For example, after a few months of wearing prisms that laterally reversed the visual field, neurons in the macaque V1 respond to both ipsi- as well as to contralateral visual fields (Sugita, 1996).

The link between neural plasticity and functional recovery of conscious vision was recently investigated by Henriksson et al. (2007), who trained a 61-year-old patient with homonymous hemianopia with flicker stimulation. The training was performed 3 years after the stroke, at a stage in which the subject showed a stable homonymous hemianopia, with no evoked neuromagnetic responses in response to visual stimulation of his blind hemifield. During training, the patient became conscious of stimuli presented in the blind hemifield; form vision emerged, and the far periphery of the blind hemifield brightened. Changes in neural responsiveness during training were documented with magnetoencephalography, and the cortical organization after training was examined with fMRI. The key finding was that, after training, visual information from both hemifields was processed mainly in the intact hemisphere. Specifically, fMRI mapping revealed a representation of both the blind and the normal hemifields in the same set of cortical areas in the intact hemisphere, more specifically in the visual motion-sensitive area V5/MT, in a region around the superior temporal sulcus and in retinotopic visual areas V1, V2, V3, and V3a.

This case differs from that reported by Muckli et al. (2009) in that restored function and the coinciding involvement of the contralesional hemisphere were due to therapeutic intervention in an elderly patient rather than spontaneous recovery occurring in early life. Nevertheless, there is a remarkable similarity between them: in both cases, conscious vision appears to be associated with the development of ipsilateral visual field representations in the early visual cortex (although the extent to which visual recovery depends particularly on V1 is not known). It thus appears that, as long as some functional V1 remains, it can participate in the mediation of visual awareness from spatial positions that it did not encode prior to the lesion. These unilateral cases are thus not inconsistent with the view that retinotopic representation of the visual field in V1 plays a necessary role in visual awareness.

PHENOMENAL AWARENESS IN BLINDSIGHT

In the above examples, at least some visual awareness was possible after a V1 lesion. However, in most patients visual information fails to reach awareness although it can in some circumstances unconsciously detected. Perhaps the most frequently studied blindsight subject is GY, who has a large unilateral lesion in the left medial occipital lobe, caused by a traffic accident at the age of 8. Striate cortex is absent in the left hemisphere, except at the occipital pole corresponding to about three to four degrees of macular sparing. Although GY (and other blindsight patients) perform well in numerous visual tasks (see Stoerig and Cowey, 1997 for reviews), there is no phenomenal awareness associated with these functions in the blind field.

Interestingly, unlike in the subjects reported by Muckli et al. (2009) and Henriksson et al. (2007), V1 in GY's intact hemisphere is not activated by stimuli presented in the blind field that induce blindsight. In contrast, extrastriate areas in the damaged as well as intact hemisphere (such as V5/MT) do show BOLD responses to stimuli presented in the blind field. However, these extrastriate activations seem to be unable to give rise to conscious perception in the absence of V1. Taken together with the findings of Henriksson et al. (2007), it seems that after a unilateral V1 lesion, conscious vision tends to primarily recover through the development of retinotopic representations of the affected field in the contralesional V1. Representations of the blind field in the extrastriate (either in the intact or damaged hemisphere) do not seem to enable awareness to arise, at least when the lesion is sustained in adulthood.

There are however circumstances in which hemianopic patients with a V1 lesion sustained in adulthood can experience qualia in the blind field. An example of this is hemianopic completion, which refers to the perceptual completion of figures located across the vertical meridian in the context of hemianopia, such that one half of the figure falls within the blind hemifield (e.g., Bender and Kahn, 1949). Marcel (1998) used after-images to demonstrate veridical conscious perception of shape in GY's blind field when it was accompanied by a shape in the sighted field. When only the normal hemifield was exposed to visual stimulation, an afterimage appeared in the normal hemifield; in contrast, when the blind field was stimulated, no afterimage was perceived. Critically, when both the normal and the blind hemifields were stimulated in combination, such that the formed a good Gestalt, the after-image appeared in both hemifields. This result shows that visual stimulation of GY's blind field can give rise to visual experiences, if the blind field stimulus is part of a larger bilateral stimulus.

The neural basis of such completion has been recently studied by Weil et al. (2009) in a patient with a homonymous hemianopia following occipital stroke that was associated with hemianopic perceptual completion across the blind field. Completion was found to be associated with activation anterior to retinotopic cortex in the lingual gyrus in the right occipital cortex, contralateral to the lesion, ipsilateral to the illusory edge of the stimulus. This region was located in visually responsive ventral visual cortex near to reported coordinates for the human lateral occipital complex (Avidan et al., 2002). Importantly, activity in the early visual cortical areas V1–V3 was not associated with completion, demonstrating that visual awareness does not always depend on activity in V1.

A further example of conscious perception without the development of an ipsilateral visual field representation in the intact V1 has come from studies using transcranial magnetic stimulation (TMS). When TMS is applied over the extrastriate visual areas, neurologically normal observers perceive flashes of light (phosphenes). However, when TMS is applied unilaterally over presumably intact extrastriate areas in GY' damaged hemisphere, he does not perceive phosphenes; this is consistent with the view that extrastriate activation cannot reach awareness in the absence of V1. However, when TMS was applied over the extrastriate area V5/MT in both the damaged and intact hemisphere in close temporal proximity, GY perceived bilateral phosphenes that extend into the blind hemifield. Using this technique, GY could even experience colored percepts in his blind field (Silvanto et al., 2008).

At the perceptual level, TMS-induced bilateral phosphenes seem similar to hemianopic completion, and it may be that they are mediated by the intact extrastriate cortex via callosal connections. Consistent with this view, a DTI study by Bridge et al. (2008) found a substantial cortico-cortical connection in GY between V5/MT bilaterally, consistent with the possibility that information can be transferred from the damaged to the intact hemisphere. While the neural basis of these blind field phosphene is unclear, the critical point is that GY's intact V1 cannot be directly involved as it does not contain a retinotopic representation of the blind visual field. The main implication of these finding therefore is that a retinotopic representation of the visual field in V1 cannot be necessary for phenomenal awareness.

WHAT DOES PLASTICITY TELL ABOUT THE ROLE OF V1 IN AWARENESS?

The evidence can be summarized as follows: after a unilateral V1 lesion, spontaneous or training-induced plasticity can lead to the development of ipsilateral visual field maps in the unaffected V1, in both children and adults. The functional recovery resulting from spontaneous plasticity is robust in childhood, with visual functions being almost normal (Muckli et al., 2009). When the lesion occurs in adulthood after the visual system has fully developed, the extent of visual recovery is much smaller, and the development of new visual field maps in the intact V1 may require behavioral rehabilitation programs (Henriksson et al., 2007). These cases support the view that V1 is necessary for visual awareness, as the development of ipsilateral visual field maps in the intact early visual cortex appears to be necessary for the recovery of conscious vision.

Importantly however, the cases in which V1 is missing bilaterally (Giaschi et al., 2003; Amicuzi et al., 2006) demonstrate that plastic changes which enable some conscious perception to recover do not necessarily require V1. While visual functions may remain compromised, the fact that conscious vision can develop at all indicates that the brain does not need V1 to give rise to conscious visual percepts. Finally, the phenomena of hemianopic completion (e.g., Bender and Kahn, 1949; Marcel, 1998; Weil et al., 2009) and blind field phosphene perception (Silvanto et al., 2007, 2008) also demonstrate that that *some* conscious vision is possible without the development of an ipsilateral visual field representation in the contralesional V1. While the extent to which plasticity plays role in these latter phenomena is not clear, they are nevertheless inconsistent with the view that a retinotopic representation of the visual field in V1 is a prerequisite of *all* visual awareness.

ARE FUNCTIONAL PROPERTIES OF V1 NECESSARY FOR AWARENESS?

A brain recovering from lesions sustained early in life have undergone a very different developmental path than that of a neurologically normal subject, and one might question the usefulness of referring in this context to visual areas which are found in normally developed visual cortex. While this may be true, the other side of the coin is that there appears to be nothing unique about specific anatomically defined areas such as V1 in enabling conscious perception. Of course, what might make V1 necessary for awareness may not be its anatomical position at the back head, but rather, the functional properties and anatomical connectivity of its neurons. If we define V1 not in terms of its anatomical location as the portion of the brain at the occipital pole around the calcarine sulcus, but rather in terms of the functional properties of its neurons, the question becomes: is the recovery of conscious perception a consequence of unaffected cortex taking over the functional properties of V1? If so, this would suggest that V1 (in its functional definition) is indeed necessary for visual awareness.

It is impossible to offer a conclusive answer to this question, but electrophysiological evidence obtained in the cat may offer some clues. Neurons in posteromedial lateral suprasylvian area, an extrastriate visual area in the cat, lose neural tuning after a damaged to the visual cortex. However, neural responsiveness can recover, especially if the lesion has been sustained in childhood (e.g., Spear and Baumann, 1979; Tong et al., 1984). The important question is whether this recovery leads to neurons developing the response properties of the damaged cortex. This question has been addressed in a number of studies (e.g., Maffrei and Fiorentini, 1973; Movshon et al., 1978; Guido et al., 1992). The key finding is that the spatial or temporal response properties in this region are not different from those found in normal cats, and they are unlike those of normal striate cortex cells. In other words, the outcome of neural plasticity is the development of normal tuning properties. That these neurons did not take over the neural properties of the damaged cortex suggests visual recovery in the cat does not rely on V1 functions being emulated by other brain areas.

That the functional properties of V1 are not emulated by other areas does not imply that the functional role of V1 in a given visual skill cannot be taken over by the unaffected cortex. Indeed, there is evidence that visual abilities that are normally based on V1 activity may rely on other areas after a V1 lesion (see Payne and Lomber, 2002). However, this does imply a change in neural tuning – a visual area with somewhat similar neural tuning as a normal V1 may take over the role. Consequently, the particular visual ability may be impaired because the neural properties in that region may be less suited for this purpose than those in the normal V1.

Interestingly, neural tuning in the extrastriate cortex of the monkey shows rather less recovery after a V1 lesion. Rodman et al. (1989) found weak responses to visual stimuli in V5/MT (with only 5% of neurons responding strongly to visual stimulation), and receptive fields were difficult to localize precisely. A different conclusion was reached by Girard et al. (1992), who used cooling to deactivate V1 and found that 80% of neurons remained responsive to visual stimulation. However, the cooling plate did not cover the whole of V1 (affecting only approximately 4° of the lower quadrant); thus the majority of V5/MT neurons did not lose their V1 input, which could explain their high level of responsiveness. Subsequent studies (e.g., Azzopardi et al., 2003; Collins et al., 2003) on monkeys found no or severely reduced neural responsiveness in V5/MT. For example, in the study by Azzopardi et al. (2003) the neural responses to moving stimuli in V5/MT were as small as the V5/MT responses to static stimuli in the normal brain. More recently, functional MRI in monkeys demonstrates weak responses of extrastriate cortex to visual stimulation following a chronic V1 lesion, with overall activity evoked by visual stimulation approximately one-fifth of normal levels (Schmid et al., 2010). Interestingly, this residual activity was critically dependent on the lateral geniculate nucleus as it disappeared when that structure was inactivated. It remains possible however that in monkeys with V1 damage in infancy can develop more normal neural selectivity in extrastriate cortex, as there is significantly more residual vision compared to monkeys with lesions sustained in adulthood (Moore et al., 1996). Nevertheless, the electrophysiological evidence from lesioned monkeys suggests that in the absence of V1, the functioning of the extrastriate cortex is severely compromised.

In summary, while animal studies cannot conclusively answer the question of whether the recovery of visual awareness requires the development of functional properties of a normal V1 in another brain area, they suggest that this is not required. Rather, it appears that as long as extrastriate areas have normal functionality, visual perception can arise.

CONCLUSION

The objective of this review was to consider the implications of neural plasticity for understanding the role of V1 in visual awareness. Although a lesion to this region abolishes all phenomenal awareness, conscious vision can recover; the extent of this recovery appears to be linked to the age at which the lesion was sustained. The ability of the visual system to create conscious visual percepts in the absence of V1 suggests that there is fundamentally nothing unique about specific anatomically defined areas such as V1 in enabling conscious perception. Furthermore, that the extrastriate neural plasticity enabling visual recovery in the cat is associated with the recovery of normal tuning properties rather than emulation of functional properties of V1 neurons suggests that there is nothing unique about the neural selectivity of this region in visual awareness. Rather, the dramatic effect of a V1 lesion on visual awareness may be due to its role as the provider of input to most of extrastriate cortex that enables normal extrastriate neural responsiveness.

REFERENCES

Amicuzi, I., Stortini, M., Petrarca, M., Di Giulio, P., Di Rosa, G., Fariello, G., Longo, D., Cannatà, V., Genovese, E., and Castelli, E. (2006). Visual recognition and visually guided action after early bilateral lesion of occipital cortex: a behavioral study of a 4.6-year-old girl. *Neurocase* 12, 263–279.

Avidan, G., Harel, M., Hendler, T., Ben-Bashat, D., Zohary, E., and Malach, R. (2002). Contrast sensitivity in human visual areas and its relationship to object recognition. *J. Neurophysiol.* 87, 3102–3116.

Azzopardi, P., Fallah, M., Gross, C. G., and Rodman, H. R. (2003). Response latencies of neurons in visual areas MT and MST of monkeys with striate cortex lesions. *Neuropsychologia* 41, 1738–1756.

Bender, M. B., and Kahn, R. L. (1949). After-imagery in defective fields of vision. *J. Neurol. Neurosurg. Psychiatr.* 12, 196–204.

Bova, S. M., Giovenzana, A., Signorini, S., La Piana, R., Uggetti, C., Bianchi, P. E., and Fazzi, E. (2008). Recovery of visual functions after early acquired occipital damage. *Dev. Med. Child Neurol.* 50, 311–315.

Bridge, H., Thomas, O., Jbabdi, S., and Cowey, A. (2008). Changes in connectivity after visual cortical brain damage underlie altered visual function. *Brain* 131(Pt 6), 1433–1444.

Collins, C. E., Lyon, D. C., and Kaas, J. H. (2003). Response of neurons in the middle temporal visual area after long-standing lesions of the primary visual cortex in the adult new world monkeys. *J. Neurosci.* 23, 2251–2264.

Cowey, A. (2004). The 30th Sir Frederick Bartlett lecture. Fact, artefact, and myth about blindsight. *Q. J. Exp. Psychol. A.* 57, 577–609.

Cowey, A. (2010). The blindsight saga. *Exp. Brain Res.* 200, 3–24.

Giaschi, D., Jan, J. E., Bjornson, B., Young, S. A., Tata, M., Lyons, C. J., Good, W. V., and Wong, P. K. (2003). Conscious visual abilities in a patient with early bilateral occipital damage. *Dev. Med. Child Neurol.* 45, 772–781.

Girard, P., Salin, P. A., and Bullier, J. (1992). Response selectivity in area MT of the macaque monkey during reversible inactivation of area V1. *J. Neurophysiol.* 67, 1437–1446.

Guido, W., Spear, P. D., and Tong, L. (1992). How complete is physiological compensation in extrastriate cortex after visual cortex damage in kittens. *Exp. Brain Res.* 91, 455–466.

Guzzetta, A., D'Acunto, G., Rose, S., Tinelli, F., Boyd, R., and Cioni, G. (2010). Plasticity of the visual system after early brain damage. *Dev. Med. Child Neurol.* doi: 10.1111/j.1469-8749.2010.03710.x. [Epub ahead of print].

Henriksson, L., Raninen, A., Näsänen, R., Hyvärinen, L., and Vanni, S. (2007). Training-induced cortical representation of a hemianopic hemifield. *J. Neurol. Neurosurg. Psychiatry* 78, 74–81.

Kasten, E., Strasburger, H., and Sabel, B. A. (1997). Programs for diagnosis and therapy of visual field deficits in vision rehabilitation. *Spat. Vis.* 10, 499–503.

Kerkhoff, G., Münssinger, U., and Meier, E. K. (1994). Neurovisual rehabilitation in cerebral blindness. *Arch. Neurol.* 51, 474–481.

Maffrei, L., and Fiorentini, A. (1973). The visual cortex as a spatial frequency analyser. *Vision Res.* 13, 1255–1267.

Marcel, A. J. (1998). Blindsight and shape perception: deficit of visual consciousness or of visual function? *Brain* 121, 1565–1588.

Moore, T., Rodman, H. R., Repp, A. B., Gross, C. G., and Mezrich, R. S. (1996). Greater residual vision in monkeys after striate cortex damage in infancy. *J. Neurophysiol.* 76, 3928–3333.

Movshon, J. A., Thompson, I. D., and Tolhurst, D. J. (1978). Spatial and temporal contrast sensitivity of neurons in areas 17 and 18 of the cat's visual cortex. *J. Physiol.* 283, 101–120.

Muckli, L., Naumer, M. J., and Singer, W. (2009). Bilateral visual field maps in a patient with only one hemisphere. *Proc. Natl. Acad. Sci. U.S.A.* 106, 13034–13039.

Payne, B. R., and Lomber, S. G. (2002). Plasticity of the visual cortex after injury: what's different about the young brain? *Neuroscientist* 8, 174–185.

Payne, B. R., and Rushmore, R. J. (2003). Animal models of cerebral neglect and its cancellation. *Neuroscientist.* 9, 446–454.

Riddoch, G. (1917). Dissociation of visual perceptions due to occipital injuries, with especial reference to appreciation of movement. *Brain* 40, 15–57.

Ro, T. (2010). What can TMS tell us about visual awareness? *Cortex* 46, 110–113.

Rodman, H. R., Gross, C. G., and Albright, T. D. (1989). Afferent basis of visual response properties in area MT of the macaque. I. Effects of striate cortex removal. *J. Neurosci.* 9, 2033–2050.

Schmid, M. C., Mrowka, S. W., Turchi, J., Saunders, R. C., Wilke, M., Peters, A. J., Ye, F. Q., and Leopold, D. A. (2010). Blindsight depends on the lateral geniculate nucleus. *Nature* 466, 373–377.

Schmielau, F., and Wong, E. K. Jr. (2007). Recovery of visual fields in brain-lesioned patients by reaction perimetry treatment. *J. Neuroeng. Rehabil.* 4, 31.

Silvanto, J., Cowey, A., Lavie, N., and Walsh, V. (2007). Making the blindsighted see. *Neuropsychologia* 45, 3346–3350.

Silvanto, J., Cowey, A., and Walsh, V. (2008). Inducing conscious perception of colour in blindsight. *Curr. Biol.* 18, R950–R951.

Spear, P. D., and Baumann, T. P. (1979). Effects of visual cortex removal on receptive field properties of neurons in the lateral suprasylvian visual area of the cat. *J. Neurophysiol.* 42(Pt. 1), 31–56.

Stoerig, P., and Cowey, A. (1997). Blindsight in man and monkey. *Brain* 120(Pt 3), 535–559 (review).

Sugita, Y. (1996). Global plasticity in adult visual cortex following reversal of visual input. *Nature* 380, 523–526.

Tong, F. (2003). Primary visual cortex and visual awareness. *Nat. Rev. Neurosci.* 4, 219–229.

Tong, L., Kalil, R. E., and Spear, P. D. (1984). Critical periods for functional anatomical compensation in the lateral suprasylvian visual area following removal of visual cortex in cats. *J. Neurophysiol.* 52, 941–960.

Weil, R. S., Plant, G. T., James-Galton, M., and Rees, G. (2009). Neural correlates of hemianopic completion across the vertical meridian. *Neuropsychologia* 47, 457–464.

Werth, R. (2008). Cerebral blindness and plasticity of the visual system in children. A review of visual capacities in patients with occipital lesions, hemispherectomy or hydranencephaly. *Restor. Neurol. Neurosci.* 26, 377–389.

Zihl, J., and von Cramon, D. (1986). Visual field rehabilitation in the cortically blind? *J. Neurol. Neurosurg. Psychiatry* 49, 965–966.

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

Received: 22 November 2010; paper pending published: 16 December 2010; accepted: 04 January 2011; published online: 20 January 2011.

Citation: Silvanto J and Rees G (2011) What does neural plasticity tell us about role of primary visual cortex (V1) in visual awareness? Front. Psychology ***2****:6. doi: 10.3389/fpsyg.2011.00006*

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

frontiers in
PSYCHOLOGY

REVIEW ARTICLE
published: 14 February 2011
doi: 10.3389/fpsyg.2011.00019

The nature of consciousness in the visually deprived brain

Ron Kupers[1]*, Pietro Pietrini[2], Emiliano Ricciardi[2] and Maurice Ptito[3,4]

[1] *Institute of Neuroscience and Pharmacology, Panum Institute, University of Copenhagen, Copenhagen, Denmark*
[2] *Laboratory of Clinical Biochemistry and Molecular Biology, University of Pisa, Pisa, Italy*
[3] *Chaire Harland Sanders en Sciences de la Vision, École d'Optométrie, Université de Montréal, Montréal, QC, Canada*
[4] *Danish Research Center for Magnetic Resonance, Hvidovre Hospital, Hvidovre, Denmark*

Edited by:
Morten Overgaard,
Aarhus University, Denmark

Reviewed by:
Ryota Kanai,
University College London, UK
Steven Laureys,
University of Liège, Belgium

****Correspondence:***
Ron Kupers, Institute of Neuroscience and Pharmacology, Panum Institute, Copenhagen University, Blegdamsvej 3B, DK-2200 Copenhagen, Denmark.
e-mail: ron@pet.rh.dk

Vision plays a central role in how we represent and interact with the world around us. The primacy of vision is structurally imbedded in cortical organization as about one-third of the cortical surface in primates is involved in visual processes. Consequently, the loss of vision, either at birth or later in life, affects brain organization and the way the world is perceived and acted upon. In this paper, we address a number of issues on the nature of consciousness in people deprived of vision. Do brains from sighted and blind individuals differ, and how? How does the brain of someone who has never had any visual perception form an image of the external world? What is the subjective correlate of activity in the visual cortex of a subject who has never seen in life? More in general, what can we learn about the functional development of the human brain in physiological conditions by studying blindness? We discuss findings from animal research as well from recent psychophysical and functional brain imaging studies in sighted and blind individuals that shed some new light on the answers to these questions.

Keywords: vision, blindness, consciousness, qualia, cross-modal plasticity, supramodality, rewiring

> *If we could splice the nerves so that the excitation of the ear fed the brain centre concerned with seeing, and vice versa, we would "hear the lightning and see the thunder"*
>
> *William James, Principles of Psychology, Dover, New York, 1890*

INTRODUCTION

The study of brain function in individuals with congenital blindness provides a unique approach to understand how consciousness develops in the absence of sight. Classically, sight has always been regarded as the most important sense for humans to interact with the environment. In the ancient Greek language the verb "to know" (oi\da) was the past tense of the verb "to see" (oJravw), that is, "*I saw and thus I know*." The relevance of sight is also clearly reflected in the mental attitude senses of the lexicon of vision. Consider everyday linguistic uses such as "*I see what you mean*," "*Do you see my point?*" etc. At the same time, the amount of brain surface devoted to visual function in primates is remarkably high, accounting for almost one-third of the whole cortex. This predominance of vision in primates raises some critical questions. How do individuals who never had any visual experience since birth form a conscious representation of a world that they have never seen? How do their brains behave? What happens to vision-devoted brain structures in individuals who are born deprived of sight or who lose vision at different ages? To what extent is visual experience truly necessary for the brain to develop its functional architecture? What does the study of blind individuals teach us about the functional organization of the sighted brain in physiological conditions? We attempt to shed some new light on these old questions by reviewing evidence from studies conducted in animals and in humans.

ANIMAL STUDIES

Early experience plays a key role in the development of the nervous system as the brain has to adapt continuously to environmental changes that threaten its harmonious development. This adaptation phenomenon is called brain plasticity and refers to the lifelong changes in the structure of the brain that accompany experience (experience-dependent plasticity). This concept suggests that the brain is pliable, like plastic, and can be molded into different forms in response to changes in the environment such as sensory deprivation, brain injuries or abnormal development. Indeed, several regimens of visual deprivation such as dark-rearing, enucleation, or eye-lid suturing lead to alterations of the visual system. Conversely, environmental enrichment also produces regional changes in brain anatomy such as increased dendritic space for synapses, increased cortical thickness, and elevated gene expression (Ptito and Desgent, 2006).

CROSS-MODAL PLASTICITY

The loss of a particular sense leads to the invasion of the deprived cortical area by inputs originating from other modalities, illustrating the remarkable capacity of the cerebral cortex for plasticity and reorganization (Ptito and Desgent, 2006; Pietrini et al., 2009; Merabet and Pascual-Leone, 2010). These inter-modal connections result from a phenomenon called cross-modal plasticity. As early as 1977, Rebillard et al. (1977) reported that the primary auditory cortex in congenitally deaf cats is driven by visual stimuli. Conversely, studies on the microphthalmic mole rat (*Spalax ehrenbergi*) have shown that auditory stimulation can drive neurons in the primary visual cortex (Bronchti et al., 2002). Cells in the primary visual cortex of visually deprived mice, rats, cats, or monkeys respond to somatosensory or auditory inputs, suggesting cross-modal

reorganization (Toldi et al., 1994). For example, neurons in visual cortical areas 19 in monkeys with early visual deprivation respond to somatic inputs such as manipulating the experimenter's hand in search for food (Hyvarinen et al., 1991). Peripheral inputs play a pivotal role in the organization of the neocortex, as cortical territories usually involved in visual processing are invaded by the auditory and somatosensory system. It seems therefore that the visual cortex is capable of rewiring in order to accommodate these non-visual inputs. In the case of early brain damage, however, abnormal neuronal connectivity patterns can be produced and an alternative approach to study cross-modal plasticity resides in the tampering with "blue prints" during prenatal development. Relevant to this approach are the numerous studies on "rewiring" in hamsters (Ptito and Desgent, 2006) and in ferrets (reviewed in Lyckman and Sur, 2002).

"REWIRING" THE BRAIN OF HAMSTERS

If brain damage occurs during development, abnormal neuronal connectivity patterns can develop. It is thus possible to induce, by lesioning central retinal targets, the formation of new and permanent retinofugal projections into non-visual sites such as the thalamic auditory nucleus (Frost and Metin, 1985; Ptito et al., 2001a; **Figure 1A**). These surgically induced retinal projections are retinotopically organized and make functional synapses (Metin and Frost, 1989). Neurons in the somatosensory cortex of animals with ectopic retinal projections have visual response properties similar to those of neurons in the primary visual cortex of normal sighted animals (Metin and Frost, 1989). Ferrets without visual cortex but with retinofugal projections to the auditory thalamus appear to perceive light stimuli as visual (von Melchner et al., 2000). The question concerning the parallelism between a different brain organization (produced by lesions) and a behavioral recovery is still debated although recent experiments in rewired ferrets and hamsters suggest a large degree of recovery in visual functions (reviewed in Ptito et al., 2001a). For example, responses to visual stimuli have been observed in the auditory cortex of hamsters with robust and permanent projections to the auditory thalamic nucleus (medial geniculate body) but lacking a visual cortex. Single neurons in the auditory cortex of these animals respond to visual stimuli and some respond equally well to visual as to auditory stimuli. Moreover,

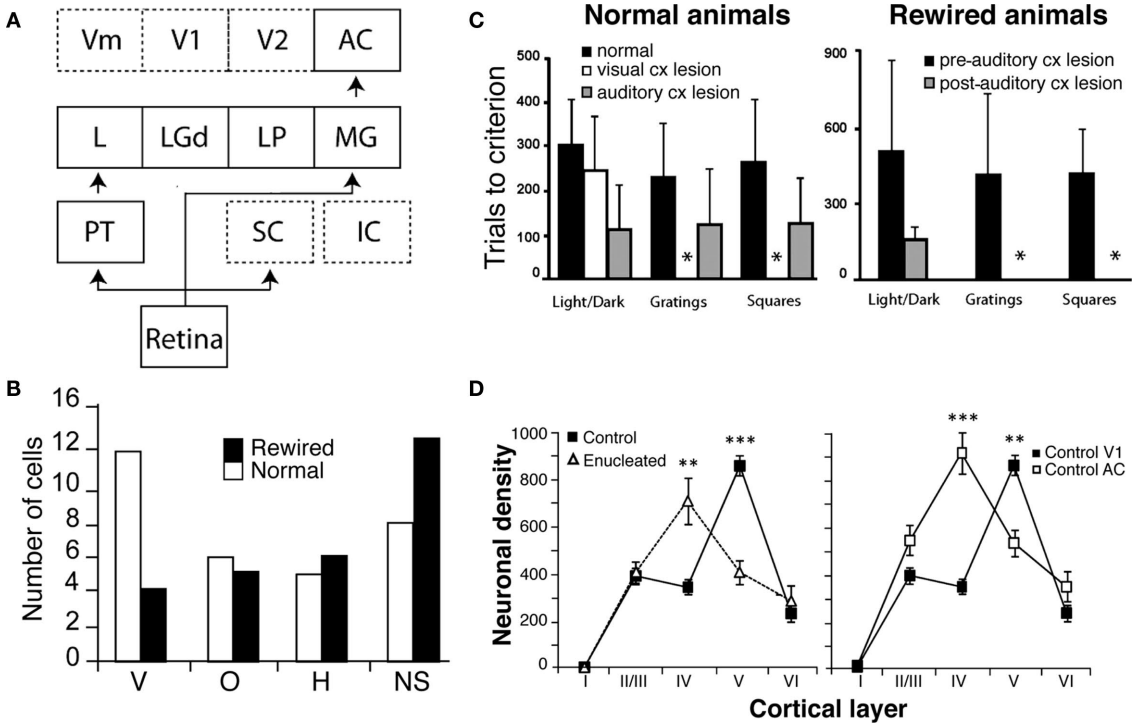


FIGURE 1 | Behavioral and physiological effects of surgically created ectopic retinal projections to the auditory cortex in hamsters. (A) Neonatal lesions of primary retinal targets (SC) and subsequent destruction of cortical visual areas (V1, V2, and Vm) produce a strong projection from the retina to the auditory cortex (AC). **(B)** Comparison of the distribution of preferred orientations for cells in area V1 (normal animals) and area A1 (rewired animals). The auditory cortex in rewired hamsters adopts receptive field properties of the primary visual cortex such as orientation selectivity. Neurons within A1 of rewired hamsters respond to horizontal (H), vertical (V), and oblique (O) stimulus orientations. All orientations are represented but with no preference bias for vertical orientations as seen in normal animals and with an increased number of non-specific cells (NS). **(C)** Rewired hamsters learn visual discrimination tasks as well as normal animals. In normal animals, a bilateral V1 lesion abolishes pattern discrimination. The same result is obtained with a lesion of the auditory cortex in rewired animals (shown by the asterisks). **(D)** Neuronal density of parvalbumin-immunoreactive cells (PV-Ir) within layers of the auditory and visual cortices in normal and enucleated hamsters. In V1 of control animals, PV-Ir are mainly found in layer V and in lesser density in layer IV. Following enucleation, the density of PV-Ir neurons is now greater in layer IV than the controls. This contrasts with the lower density of PV-Ir neurons in layer V. The resulting pattern of laminar distribution of PV-Ir cells in the primary visual cortex resembles that of the primary auditory cortex (adapted from Ptito et al., 2001a; Ptito and Desgent, 2006; Desgent et al., 2010).

cells responding to visual stimuli show orientation selectivity and motion and direction sensitivity (**Figure 1B**). These receptive field properties compare favorably well with those obtained from cells in the visual cortex of normal hamsters (Ptito et al., 2001a).

At the behavioral level, rewired hamsters can learn visual discrimination tasks as well as normal animals and a lesion of the auditory cortex abolishes this function (**Figure 1C**; Frost et al., 2000). In fact, rewired hamsters with auditory cortex lesions exhibit cortical blindness similar to non-rewired hamsters with visual cortex lesions.

POSSIBLE ROLE OF GABA AND CALCIUM BINDING PROTEINS

The mechanisms responsible for this cross-modal processing of sensory information in the cortex are not yet fully understood. Recent work carried out in our laboratory has lead to suggest that the observed changes may be due to modifications in GABAergic interneurons expressing the calcium binding proteins parvalbumin and calbindin (Desgent et al., 2010). In deaf and cross-modal rewired ferrets, qualitative changes occur in the morphology and proportion of parvalbumin and calbindin containing interneurons (Pallas, 2001, 2002). Since the laminar distribution of these proteins is significantly different in the primary visual and auditory cortices of normal hamsters (Desgent et al., 2005), specific sensory processing would require adapted cortical microcircuits and would therefore support the hypothesis of a non-stereotypical organization of cortical interneurons. The induction of aberrant connectivity to these cortices is also evident at the neurochemical level. Indeed, hamsters enucleated at birth show not only a reduction in visual cortex volume but also changes in the distribution of calcium binding proteins within that visual cortex (Desgent et al., 2010). Compared to intact hamsters, the density of parvalbumin-immunoreactive neurons in V1 of enucleated animals is higher in layer IV and lower in layer V, whereas the density of calbindin-immunoreactive cells is significantly lower in layer V (**Figure 1D**). These results suggest that the affected primary visual cortex may adopt chemical features of the auditory cortex through cross-modal rewiring and therefore support the non-stereotypical organization of cortical interneurons.

HOW ABSENCE OF VISION RESHAPES THE HUMAN BRAIN

How does absence of vision since birth affect the macrostructural organization of the human brain and through which pathways can non-visual information be funneled to the occipital cortex in the visually deprived brain? In recent years, magnetic resonance imaging (MRI)-based brain imaging techniques such as voxel-based-morphometry (VBM), diffusion tensor imaging (DTI), and diffusion tensor tractography (DTT) have been successfully applied for the *in vivo* investigation of alterations in gray and white matter in the blind human brain. The results of these studies concur that there are changes in both gray and white matter. In addition, metabolic changes have also been documented.

CHANGES IN GRAY MATTER

The results of VBM studies have revealed a significant gray matter atrophy of all brain structures of the visual pathways, including the lateral geniculate nucleus (LGN), the posterior pulvinar, the striate and extrastriate visual areas and the inferior temporal gyrus and lateral orbital cortex, regions that are part of the ventral stream which is involved in object recognition (Noppeney et al., 2005; Shimony et al., 2006; Pan et al., 2007; Ptito et al., 2008b; **Figure 2A**). These changes are massive with volume reductions ranging from 20% in extrastriate visual areas up to 25% in the primary visual cortex (Ptito et al., 2008b). Gray matter reductions also occur in non-visual areas such as the hippocampus (Chebat et al., 2007; Fortin et al., 2008), the extrapyramidal motor system (caudate, lenticular nuclei, and fornix), the prefrontal cortex and the posterior insula. Besides these volumetric reductions in gray matter, congenitally blind subjects show an increase in cortical thickness in the cuneus (**Figure 2B**) which is likely due to a reduction in pruning during the early maturation phase of the cortex, resulting from the absence of visual input.

CHANGES IN WHITE MATTER

White matter changes include atrophy of the optic tracts and the optic chiasm, the optic radiations, the splenium of the corpus callosum (Noppeney et al., 2005; Shimony et al., 2006; Pan et al., 2007; Ptito et al., 2008b) and the inferior longitudinal fasciculus (Ptito et al., 2008b), a pathway connecting the occipital cortex with the temporal lobe. Lesions of this pathway may induce visual agnosia, prosopagnosia, and disturbances in visual recent memory (Tusa and Ungerleider, 1985). No studies found direct evidence for the establishment of new fiber pathways, although volume increases in the occipito-frontal fasciculus, the superior longitudinal fasciculus and the genu of the corpus callosum have been reported (Ptito et al., 2008b). There is also indirect evidence for an increased functional connectivity between somatosensory and visual areas in the congenitally blind brain (Wittenberg et al., 2004; Ptito et al., 2005; Kupers et al., 2006). A recent study using dynamic causal modeling of functional MRI (fMRI) data acquired during an auditory discrimination task points in the direction of increased effective connectivity between the primary auditory and visual cortices in congenitally blind subjects (Klinge et al., 2010). Results with respect to increased thalamocortical connectivity were inconsistent, arguing against strengthening of thalamocortical connections in congenital blindness. It remains a matter of debate whether the increased connectivity between the primary visual cortex and the primary auditory and somatosensory cortical regions involves a direct monosynaptic or an indirect polysynaptic pathway via multisensory association areas (Fujii et al., 2009; Klinge et al., 2010). Taken together, since no *de novo* tracts have been demonstrated in congenitally blind subjects, the data suggest that cross-modal functionality of the visual cortex in early blindness is primarily mediated by preserved or strengthened cortico-cortical connections.

METABOLIC CHANGES

Besides structural changes, the occipital cortex of congenitally blind individuals also undergoes important metabolic changes. We measured brain glucose metabolism by using positron emission tomography (PET) in combination with 18-fluoro-D-deoxyglucose (18-FDG) in a group of 12 congenitally blind and 13 sighted individuals. We found that glucose metabolism at rest is increased by around 15% in both the striate and the extrastriate cortex of congenitally blind compared to sighted individuals. **Figure 2C** shows resting-state glucose metabolism in representative examples of a congenitally

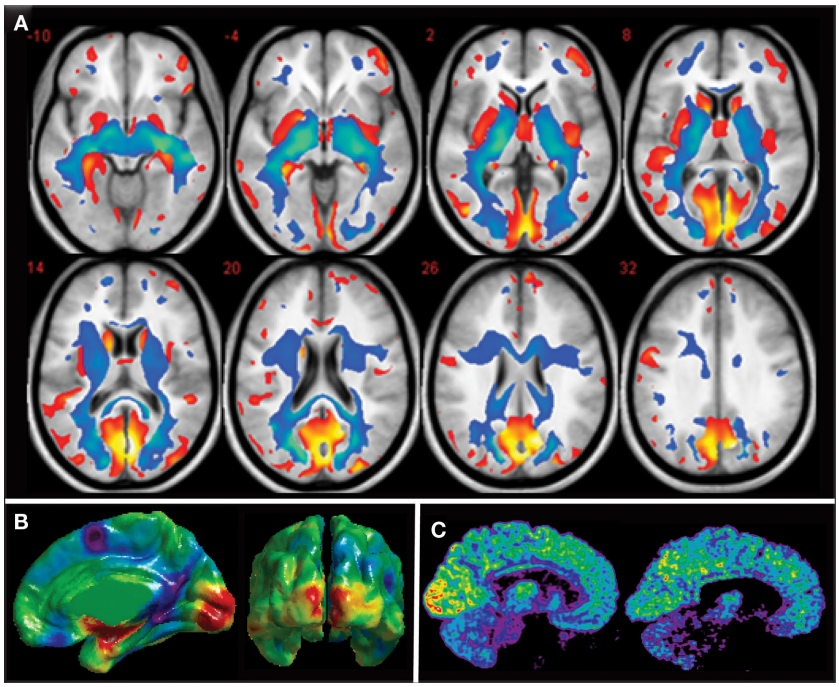


FIGURE 2 | Anatomical and metabolic changes in the congenitally blind brain. (A) Axial brain slices showing reductions in gray (red) and white matter (blue) in congenitally blind compared to matched sighted control subjects. All components of the visual system in the blind are reduced in volume (after Ptito et al., 2008b). **(B)** Differences in cortical thickness between congenitally blind and sighted control subjects. Despite a reduction in volume of the occipital cortex, cortical thickness of the cuneus is increased in congenitally blind subjects (unpublished data from our lab). **(C)** Mid-sagittal brain slices showing increased resting-state glucose metabolism in the congenitally blind brain. Illustrative examples of cerebral glucose metabolism in a congenitally blind (left) and a normal sighted control (right) subject (Kupers et al., 2009).

blind and a blindfolded control subject. Several mechanisms can be put forward to explain this increased glucose metabolism. One of the hypotheses that we favor is that it reflects internally guided cognitive activity during the FDG-uptake period while subjects are in the "resting" state. Functional brain imaging protocols and studies using transcranial magnetic stimulation (TMS) indeed have shown that the occipital cortex in the congenitally blind is recruited by and effectively involved in a variety of cognitive functions such as lexical, semantic and phonological processing, attention, verbal memory, working memory, etc. (Amedi et al., 2003; Pietrini et al., 2004; Kupers et al., 2007, 2010; Raz et al., 2007; Stevens et al., 2007; Bonino et al., 2008; Cattaneo et al., 2008; Renier et al., 2010).

HOW DO WE ACQUIRE KNOWLEDGE OF THE EXTERNAL WORLD?

THE FUSIFORM FACE AREA AND THE PARAHIPPOCAMPAL PLACE AREA…: IS THERE A BRAIN AREA FOR EVERYTHING?

How we make sense of the infinite number of distinct objects that unfold in front of our eyes has been a matter of fascinating debates for philosophers and scientists since the early days. In the past two decades, based on the results from single-cell recordings in non-human primates and functional brain imaging studies in humans, some authors have proposed the existence of a "fusiform face area" and a "parahippocampal place area," specifically devoted to recognition of faces and places, respectively (Kanwisher et al., 1997; McCarthy et al., 1997).

An alternative interpretation proposes that different areas in the extrastriate ventrotemporal cortex are specialized in different types of perceptual processes. According to this theory, the fusiform face area would be responsible for expert recognition of items from any category, not merely faces (Gauthier et al., 1999, 2000). Thus, the fusiform face area would respond to a face not just because it is a face but because of the fact that all of us are "face-experts" as we begin to look at faces since the very first days of life.

While for face and place recognition, and perhaps a few additional categories, it appears plausible that evolution might have led to the selection of dedicated neural systems given their biological meaning for survival, this certainly cannot be true for the vast majority of object categories. In the first place, the total amount of cortical surface in the ventral object vision pathway is rather limited; second, an evolution-based selection of specific neuronal groups that respond in an all-or-none manner to each distinct category would be simply impossible in terms of time required for evolutionary selection mechanisms to operate. A further strong objection comes from recent studies using fMRI that have demonstrated that looking at a face not only elicits a peak response in the lateral fusiform gyrus, but also additional activations outside of the fusiform gyrus (Ishai et al., 1999). On the other hand, other non-face object categories do activate the lateral fusiform gyrus, although to a lesser degree than faces (Ishai et al., 1999).

OBJECT FORM TOPOLOGY

To resolve this issue, Haxby et al. (2001) used fMRI to measure brain responses to different object categories, including human faces, cats, houses, and man-made objects. They demonstrated that different object categories elicit specific neural response patterns that are not restricted to selected subregions within the extrastriate ventrotemporal cortex but are rather widely distributed and overlapping. In other words, the extrastriate ventrotemporal cortex is able to produce an infinite number of neural response patterns specific for each category of objects being viewed (Haxby et al., 2001). Indeed, response patterns were so specific that it was possible to predict what the subject was actually looking at. Moreover, the specificity of the patterns changed only minimally even when the voxels with the maximal response to a given category had been removed from the analysis, indicating that the specificity of the neural response is truly distributed within the extrastriate ventrotemporal cortex and is not due to activity in a restricted area that drives the correlation. The functional architecture proposed by this model, named "object form topology," embodies the capacity of the inferior surface of the temporal lobe to generate unique representations for a virtually unlimited number of object categories.

DOES VISUAL CORTEX REQUIRE VISION TO DEVELOP AND FUNCTION?

The demonstration that the representation of a face or object is sustained by a widely distributed neural activity within the ventral temporal cortex raises further questions. Is *object form topology* in these cortical areas strictly visual or does it represent a more abstract, supramodal, functional organization? Next, is visual experience a mandatory prerequisite for this functional organization to develop?

To address these questions, we used fMRI to measure brain responses in a group of blindfolded sighted subjects while they performed non-visual object recognition tasks. Tactile recognition of face-masks and man-made objects of daily use elicited distinct category-specific patterns of neural response in the extrastriate ventral temporal cortex, that were similar to those elicited by visual recognition (Haxby et al., 2001) of the same object categories (**Figure 3A**; Pietrini et al., 2004). Furthermore, the neural response patterns elicited by tactile perception of bottles or shoes, the two man-made object categories in the study, correlated significantly with those evoked by visual perception of the same object category, indicating that these neural response patterns are supramodal in nature; that is, that they are not merely restricted to visual perception (Pietrini et al.,

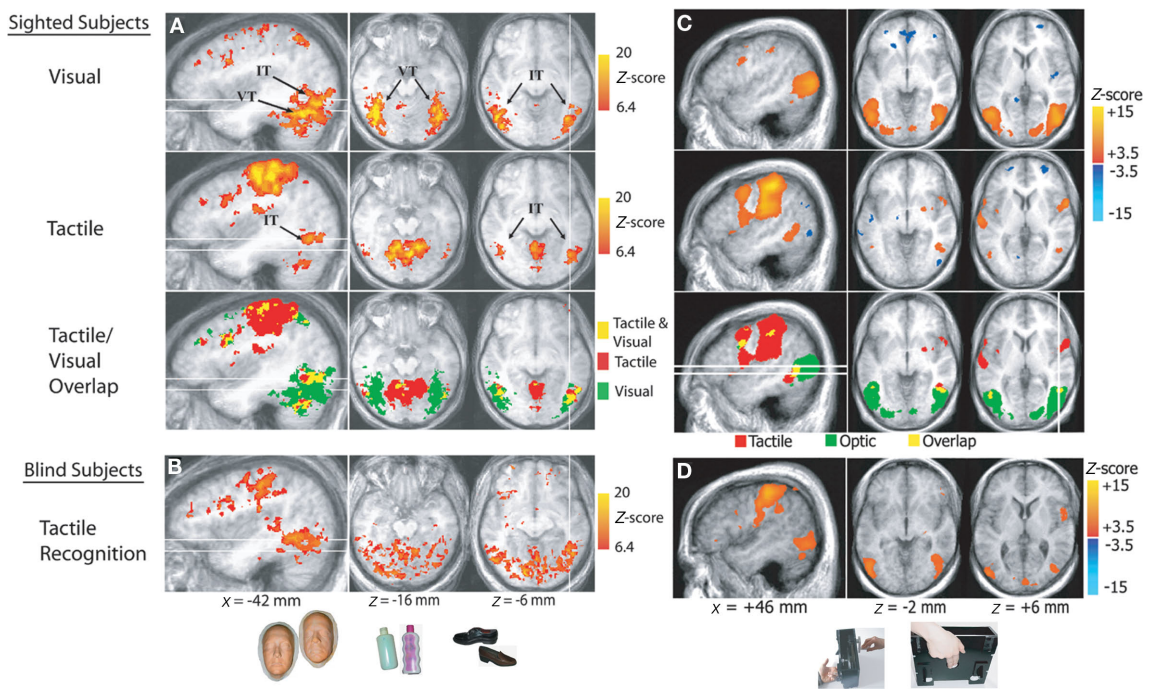


FIGURE 3 | Supramodal neural response patterns in the human brain. **(A,B)** Supramodal neural response in extrastriate ventral temporal cortex. Below, examples of stimuli (life masks of faces, plastic bottles, and shoes) used during tactile and visual recognition of different object categories in sighted and congenitally blind subjects. Brain areas that responded during tactile and/or visual object perception in sighted subjects and during tactile perception in blind individuals. The inferior temporal (IT) and ventral temporal (VT) regions activated by tactile and visual object perception are indicated. The tactile/visual overlap map shows the areas activated by both tactile and visual perception (shown in yellow), as well as the areas activated only by tactile (red) and visual (green) perception. The white lines correspond to the locations of the sagittal and axial slices. **(C,D)** Supramodal neural response in hMT+ cortex. Braille-like dot patterns moved on a plastic surface to provide translational and rotational tactile flow stimulation. Subjects' hands lay on the table with the index and middle fingers touching the plastic surface with dot patterns, as shown at the bottom of the picture. Brain areas are shown that responded during tactile or optic flow perception in sighted subjects and during tactile flow perception in blind subjects. The tactile/visual overlap map shows the areas activated by both tactile and optic flow perception (shown in yellow), as well as the areas activated only by tactile (red) and optic (green) perception (modified from Pietrini et al., 2004; Ricciardi et al., 2007).

2004, 2009). Of note, congenitally blind individuals showed category-specific neural response patterns in the ventral temporal extrastriate cortex during the same tactile object recognition task that were similar to those measured in sighted controls (**Figure 3B**).

Our findings expand results from other laboratories that in sighted individuals visual and tactile object perception activate the dorsal part of the lateral occipital cortex (LOC proper; Amedi et al., 2001, 2002; James et al., 2002) by showing a cross-modal correlation of response patterns between the two sensory modalities. In addition, our results in congenitally blind individuals clearly indicate that visual imagery cannot account for the visual cortex activation during the tactile task (James et al., 2002; Sathian and Zangaladze, 2002; Matteau et al., 2008). Indeed, while it has been shown that seeing an object or recalling the image of that object through visual imagery leads to similar neural responses in the brain (Ishai et al., 2000; O'Craven and Kanwisher, 2000; Ishai, 2010), visually based imagery is by definition absent in congenitally blind or early blind subjects with no recollection of visual experience (though congenitally blind subjects do have imagery! See review by Cattaneo et al., 2008).

The findings in the congenitally blind subjects are important also because they indicate that the development of topographically organized, category-related representations in the extrastriate visual cortex does not require visual experience. Experience with objects acquired through other sensory modalities appears to be sufficient to support the development of these patterns. Thus, at least to some extent, the visual cortex does not require vision to develop its functional architecture that makes it possible to acquire knowledge of the external world.

SUPRAMODAL FUNCTIONAL ORGANIZATION OF THE DORSAL STREAM

In light of the evidence discussed above indicating a supramodal functional organization independent from visual experience in the ventral "what" pathway, we asked whether the dorsal "where" pathway, that is devoted to spatial processing (Ungerleider and Mishkin, 1982; Haxby et al., 1994) also shows a similar property. To test this hypothesis, we compared fMRI responses while sighted subjects and individuals with congenital or early blindness and no recollection of visual experience performed distinct spatial perception tasks through the visual and/or tactile modalities. In a study using a one-back spatial discrimination task, sighted and congenitally/early blind subjects had to indicate whether a visually and/or tactilely presented matrix was the same or different as compared with the previous one. In the tactile spatial discrimination task, both groups activated a fronto-parietal network that extended bilaterally from dorsolateral prefrontal and anterior cingulate cortex toward fronto-parietal sensorimotor and posterior parietal cortical regions, including precuneus and intraparietal sulci (Ricciardi et al., 2006a; Bonino et al., 2008). A similar fronto-parietal network was activated during the visual version of the spatial task in the sighted subjects, indicating that the dorsal cortical pathway is capable of processing both visual and non-visual spatial information.

PROCESSING OF MOTION BY VISION OR TOUCH

Concurrent processing of spatial information and motion discrimination is crucial to perceive and react to dynamic environmental changes. Is motion perception processed also in a *supramodal* fashion? Furthermore, which is the effect of visual experience on the development of the functional architecture in motion responsive cortical regions? To investigate these questions, we compared brain responses in sighted and congenitally or early blind individuals during passive perception of visual and/or tactile motion (Ricciardi et al., 2007). Perception of dynamic stimuli in the visual and tactile sensory modalities shares fundamental psychophysical principles that can be explained by similar computational models. Both optic and tactile motion provide information about object form, position, orientation, consistency and movement, and also about the position and movement of the self in the environment (Bicchi et al., 2008). In sighted subjects, visual motion perception induced activation in the human middle temporal (hMT+) complex in the posterior inferior temporal cortex bilaterally, as expected on the basis of previous studies (Watson et al., 1993; Ptito et al., 2001a). In contrast, tactile motion perception activated the anterior part but deactivated a more posterior part of the hMT+ complex (**Figure 3C**). In blind subjects, tactile motion activated not only the anterior portion but also the more posterior part of the hMT+ complex, which was activated by visual motion and deactivated by tactile motion in sighted subjects (**Figure 3D**). In a separate study we demonstrated that repetitive TMS (rTMS) over the more anterior portion of hMT+ impairs the ability to detect changes in rotational velocity during a tactile motion discrimination task in blindfolded sighted subjects, indicating that recruitment of hMT+ is not a merely an epiphenomenon but it is truly necessary for tactile motion discrimination (Ricciardi et al., 2011).

ROLE OF VISUAL EXPERIENCE ON THE FUNCTIONAL ARCHITECTURE OF THE hMT+ COMPLEX

The observation that the hMT+ complex can be activated by tactile motion perception even in congenitally blind subjects demonstrates that recruitment of the hMT+ cortex is not mediated by visual-based mental imagery and that visual experience is not necessary for its development. Visual experience, however, does affect the functional architecture of hMT+ as it leads to the segregation of this region into a more anterior part, that is involved in the representation of both optic and tactile motion, and a more posterior part, that is uniquely involved in the representation of optic flow. When hMT+ develops in the absence of visual experience, the entire structure becomes involved in the representation of tactile motion. Moreover, hMT+ develops different networks of connectivity with the rest of the brain in sighted and in congenitally blind individuals (Sani et al., 2010). These results suggest that competitive interactions between visual and tactile inputs lead to functional specialization within the hMT+ complex only under conditions of physiological development.

THE hMT+ COMPLEX PROCESSES MOTION *PER SE*

Activation of area hMT+ was also shown in both sighted and congenitally blind individuals while listening to auditory stimuli that elicit the apparent perception of sounds moving right-to-left, front-to-back, and self-rotating (Poirier et al., 2006; Ricciardi et al., 2006b). Moreover, area hMT+ in blindfolded sighted and blind subjects responded to motion stimuli applied to the tongue by the means of an electrotactile device (Matteau et al., 2010). These findings expand the data discussed above indicating that hMT+

can process information about motion that is not acquired through the visual sensory modality, including tactile and auditory flow. It follows that hMT+ is capable of processing motion-related information *per se*, even when motion stimuli are delivered to body structures, such as the tongue, that, at least in humans, are not primarily devoted to the perception of movement or spatial location (Matteau et al., 2010).

Altogether, the results of these studies strongly indicate that "visual" association cortical regions are capable of processing and interpreting information carried by non-visual sensory modalities. This is not merely the consequence of a phenomenon of plastic functional reorganization in the brain of subjects deprived of sight since birth or soon afterward, as this ability is also present in sighted subjects, suggesting that these cortical structures are supramodal in nature. Not surprisingly, however, sighted and congenitally blind individuals do show differences in the extension and magnitude of the activation of the recruited areas that are likely due to rearrangements that follow the lack of sight. The supramodal nature of this functional cortical organization may explain how individuals who have never had any visual experience are able to acquire normal knowledge about objects and their position in space, form mental representations of and interact effectively with the external world (Pietrini et al., 2004, 2009, Ricciardi et al., 2007; Fiehler et al., 2009).

UNDERSTANDING WITHOUT SEEING: THE MIRROR NEURON SYSTEM IN BLIND SUBJECTS

ACTION UNDERSTANDING

Understanding actions carried out by other individuals is crucial for survival and for social organization in human and non-human primates. A particular class of neurons, discovered in the monkey premotor and parietal cortex, discharges both when performing a goal-directed action and when observing another individual performing the same action. These neurons have been named "mirror neurons" (Gallese et al., 1996; Rizzolatti et al., 1996; Rizzolatti and Fadiga, 1998; Rizzolatti and Craighero, 2004). The fact that this mirror neuron system is able to transform visual information into motor knowledge raises the hypothesis that this system may also play a significant role in action understanding. This has been confirmed by several animal studies showing that the mirror neuron system is also recruited when monkeys receive non-visual clues which allow them to understand the meaning, and create a mental representation of the occurring actions, such as when listening to action sounds (Kohler et al., 2002). In fact, a subclass of auditory–visual mirror neurons respond both while monkeys perform hand or mouth actions and while they listen to sounds of similar actions.

A similar "mirror" system has been identified in humans by using fMRI and is thought to play a major role not only in action and intention understanding, but also in learning by imitation, empathy, and language development (Buccino et al., 2004a,b; Rizzolatti and Craighero, 2004; Rizzolatti, 2005; Fabbri-Destro and Rizzolatti, 2008; Rizzolatti and Sinigaglia, 2008). The human mirror neuron system is activated during the observation of actions done by others and recruits a complex network formed by occipital, temporal, and parietal areas, and the inferior frontal cortex. As in non-human primates, the human mirror system is also recruited when individuals receive clues to understand the meaning of the occurring action with no access to visual features, such as when they only listen to the sound of actions (Gazzola et al., 2006; Lahav et al., 2007) or action-related sentences (Baumgaertner et al., 2007; Galati et al., 2008). These findings however do not address the question of whether activation within the mirror system in response to aurally presented stimuli may be mediated by sound-elicited visual imagery of the represented action or is rather independent from vision.

HOW BLIND SUBJECTS "SEE" THE ACTION OF OTHERS

We recently asked whether an efficient mirror neuron system exists in individuals who have never had any visual experience, and whether this action recognition-oriented network shares common neural patterns in sighted and blind individuals (Ricciardi et al., 2009). We measured neural response patterns in congenitally or early blind and sighted volunteers during the auditory presentation of hand-executed action or environmental sounds, and the motor pantomime of manipulation tasks. Sighted volunteers also performed an additional visual action recognition task. Congenitally blind individuals activated a premotor-temporo-parietal cortical network in response to aurally presented actions. This network overlapped both with mirror system areas found in sighted subjects in response to visually and aurally presented stimuli, and with the brain response patterns elicited by motor pantomime of the same actions (**Figure 4**). Furthermore, the mirror system showed a significantly larger response to familiar as compared to unfamiliar action sounds in both sighted and blind individuals (Ricciardi et al., 2009).

These findings indicate that visual experience is not a necessary precondition for the functional development of an efficient mirror neuron system, and suggest that a more abstract representation of actions by others also takes place through non-visual sensory modalities. This may contribute to explain the ability of congenitally blind individuals to learn by imitation of others.

SEEING WITHOUT UNDERSTANDING: BLINDSIGHT

In the preceding sections, we have shown that modifications of the input structures and alterations of the retino-recipient subcortical structures lead to a rearrangement of brain architecture. These altered brain structures can process sensory information and lead, in some instances, to quasi-normal behavioral functions. Our understanding of the mechanisms in recovery from large cortical lesions associated with brain plasticity is still unclear and remains an upmost challenge. With respect to vision, it remains difficult to explain in a satisfactory manner the remaining visual functions in patients with lesions restricted to the primary visual cortex (area V1) or with massive lesions encompassing the entire visual cortex of one hemisphere as in hemispherectomy. Destruction of area V1 has devastating effects and induces cortical blindness, i.e., a contralesional loss of the visual field (homonymous hemianopia). There is evidence that hemianopic human subjects and monkeys possess residual visual capacities in the blind part of their visual field (Cowey and Stoerig, 1995; Cowey, 2010). For example, in forced-choice paradigms where participants have to respond to stimuli presented in their blind field, performance is always above chance level

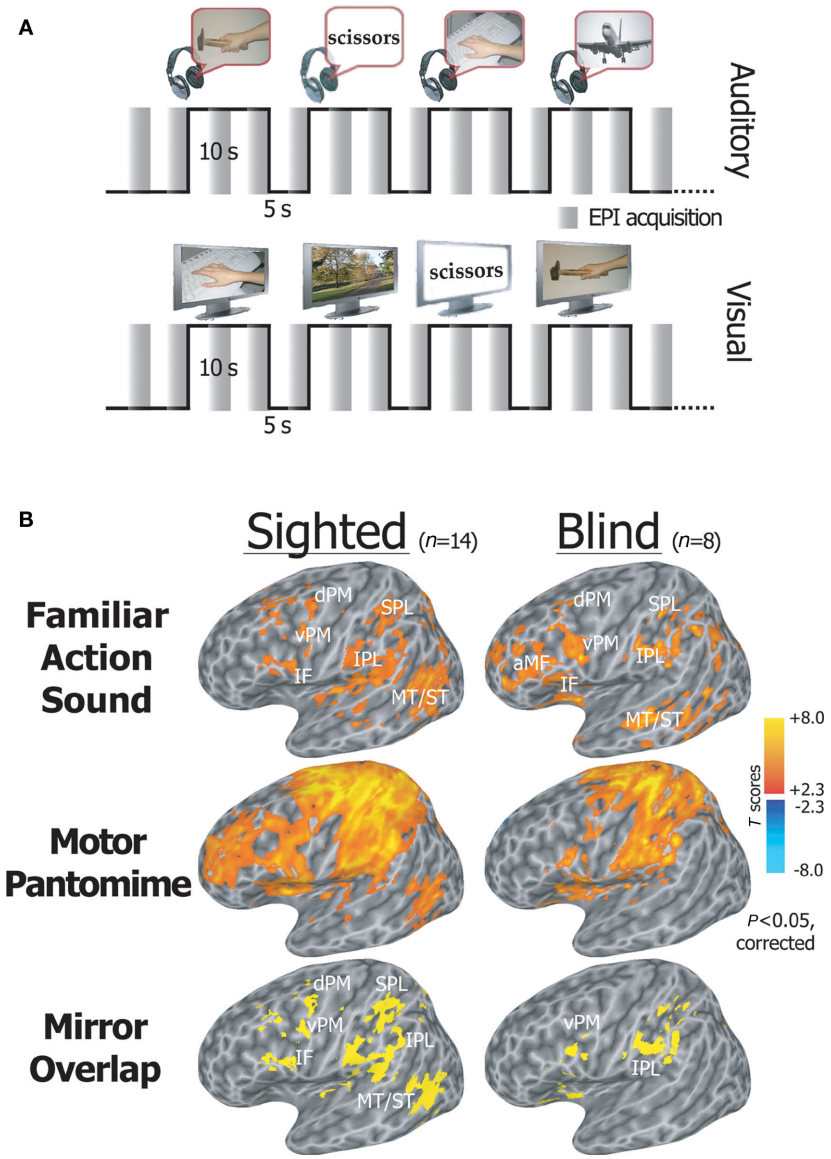


FIGURE 4 | Activation of the mirror system by action sounds. (A) fMRI experimental paradigm. An fMRI sparse sampling block design was used to examine neural activity in congenitally blind and sighted volunteers, while they alternated between the random presentation of hand-executed action or environmental sounds/movies, and the motor pantomime of a "virtual" tool or object manipulation task. **(B)** Statistical maps showing brain regions activated during listening to familiar action as compared to environmental sounds, and during the motor pantomime of action as compared to rest. Auditory mirror voxels are shown in yellow as overlap between the two task conditions (bottom row). Spatially normalized activations are projected onto a single-subject left hemisphere template in Talairach space. aMF, anterior middle frontal gyrus; IF, inferior frontal gyrus; vPM, ventral premotor cortex; dPM, dorsal premotor cortex; MT/ST, middle temporal and superior temporal cortex; IPL, inferior parietal lobule; SPL, superior parietal lobule (modified from Ricciardi et al., 2009).

although subjects consistently deny having seen the stimulus. This lack of acknowledged awareness has been termed *blindsight* (Weiskrantz et al., 1974) and has received considerable attention within the neuroscience community. The preserved visual abilities that have been reported include target detection and localization by eye-movement or manual pointing, movement and direction detection, two-color discrimination, as well as relative velocity discrimination. These residual functions have been ascribed to the spared extrastriate cortices of the lesioned hemisphere that maintain "normal" anatomical connections with their subcortical targets (Cowey, 2010), though some claims that these residual abilities are due to the sparing of minimal portions of V1 cortex (Radoeva et al., 2008). There is recent evidence from fMRI studies in monkeys with V1 lesions that ascribe blindsight to extrastriate activation via a residual pathway from the LGN to the extrastriate visual cortex (Schmid et al., 2010). In agreement with this observation, we recently showed a direct functional connection between the thalamus and the hMT+ complex in humans, that would enable motion information to reach directly hMT+, thereby bypassing V1 (Gaglianese et al., 2010).

VISUAL AWARENESS FOLLOWING HEMISPHERECTOMY

Hemispherectomy patients offer an alternative and unique model to study blindsight. In this condition, all the visual cortical areas of one hemisphere have been surgically removed, preventing the possibility that spared remnants of the visual cortex or extrastriate visual areas contribute to residual vision (Ptito and Leh, 2007). In addition, hemispherectomy allows for the investigation of the contribution of the remaining hemisphere through rewiring of the subcortical visual pathways. When hemispherectomized patients are asked to respond to a stimulus presented in their intact hemifield, they respond faster when an additional stimulus is presented at the same time in their blind hemifield, indicating a spatial summation effect, in spite of the fact that they are not aware that a stimulus was presented in the blind hemifield (Tomaiuolo et al., 1997). An fMRI study showed that these patients activate ipsilateral striate and extrastriate areas V3/V3A and V5 following stimulation of their blind hemifield (Bittar et al., 1999). In contrast, following stimulation of the intact hemifield, hemispherectomized patients activate the same brain areas in the contralateral hemisphere as normal sighted controls. Recent DTI studies revealed that the information originating in the blind hemifield is routed through the intercollicular commissure to the contralateral superior colliculus, then to the pulvinar to reach the visual cortex of the remaining hemisphere (Leh et al., 2006, 2008). These results are supported by neuroanatomical data showing a remarkable preservation of both superior colliculi in hemispherectomized monkeys (Théoret et al., 2001). The remaining hemisphere therefore seems to play a role in the implicit processing of visual information presented in the blind hemifield. If the primary visual cortex is necessary for visual consciousness, how can we explain that blindsight is still present in hemispherectomized patients?

CONSCIOUS VERSUS UNCONSCIOUS PERCEPTS

Models of conscious visual perception have evoked recurrent processing of visual information between visual cortical areas (Lamme, 2006) or a preconscious treatment of the activation within the visual areas (Dehaene et al., 2006). In both cases, the recruitment of fronto-parietal areas leads to a reportable conscious percept. Following selective damage of area V1, visual information from the blind hemifield is funneled directly to extrastriate cortical regions via subcortical structures (Cowey 2010; Schmid et al., 2010). We believe that in this case, the recurrent activation within the visual cortical areas is diminished to subliminal levels, leading to an unconscious percept (blindsight type I). The additional recruitment of the fronto-parietal network increases the activation level to threshold level, by recurrent activation, leading to visual awareness (blindsight Type II). In the case of hemispherectomy, the information from the blind field is transferred to the remaining hemisphere (Ptito and Leh, 2007) and the same recurrent activations might operate in the remaining hemisphere, explaining the two types of blindsight reported in these patients.

The same recurrent activation model of Lamme (2006) can also be applied to the condition of persistent vegetative state (PVS). For example, in PVS patients, sensory information stays confined to primary sensory cortex and is functionally uncoupled from activation in higher order association areas, including the posterior parietal cortex (Laureys et al., 2002). It is therefore plausible that recurrent activation processing does not occur, leaving the system in a subliminal state. This may explain why awake PVS patients may show basic visual functions such as eye movements and/or visual fixation without showing signs of visual consciousness (Owen et al., 2009; Bruno et al., 2010).

SUBJECTIVE EXPERIENCE (QUALIA) ASSOCIATED WITH ACTIVATION OF THE VISUAL CORTEX

As it has become clear from the studies described above, the occipital cortex in congenitally blind individuals is recruited by a wide variety of different types of sensory stimuli and cognitive tasks. It is generally accepted that cortical activity in a certain area produces a subjective sensation within the same domain. Thus, electrical stimulation experiments showed that stimulation of the somatosensory cortex (SI) induces tactile sensations referred to a particular body area. These stimulation studies further showed that the body is somatotopically mapped as inputs from neighboring body parts are encoded in adjacent parts within SI (Penfield and Boldrey, 1937). TMS is a technique that allows to stimulate the cortex in a non-invasive manner, either with the aim to interfere with brain activity within a specific brain area or to assess cortical excitability by evoking subjective sensations such as phosphenes or overt motor responses (Cowey and Walsh, 2000). For instance, the phosphene threshold is reduced after a brief period of blindfolding, indicating an increase in the excitability of the visual cortex (Boroojerdi et al., 2000). In view of the earlier described cross-modal responses in the occipital cortex of the blind, the question arises which type of sensations will be induced when applying TMS over the occipital cortex in these subjects. Studies in rewired ferrets seem to suggest that activation in the auditory cortex may be perceived as visual and not as auditory (von Melchner et al., 2000). We addressed this question in a series of studies (Kupers et al., 2006; Ptito et al., 2008a). In a first study, we exploited a tactile-to-vision sensory substitution model to examine the subjective character of experience associated with the activation of occipital cortex before and after the establishment of cross-modal plasticity (Kupers et al., 2006). More specifically, we wanted to test whether stimulation of the occipital cortex induces subjective sensations associated with the new (tactile) input. Thereto, we applied TMS over the occipital cortex before and after training with the tongue display unit (TDU) in a group of congenitally blind and blindfolded sighted control subjects. The TDU is a sensory substitution device that captures a visual image, taken by a camera or generated on a computer, and translates it into electrotactile pulses which are applied to the tongue (Bach-y-Rita and Kercel, 2003). With training, subjects learn to use the TDU to perform basic visual tasks such as orientation discrimination, motion and form detection, and spatial navigation. Our results showed that the effects of TMS depend on group and training. Before training, TMS of the occipital cortex elicited phosphenes in control subjects but not in blind subjects, which is in line with earlier findings (Cowey and Walsh, 2000). In sharp contrast, following 1 week of training, occipital TMS-induced "tactile sensations" on the tongue in blind subjects, which were described as short-lasting experiences of distinct tingling, varying in intensity, extent, and topography depending on the locus of the occipital cortex which was stimulated (**Figure 5A**). The number of occipital sites from which TMS-induced tactile sensations could be induced correlated significantly with

the size of occipital cortex that blind participants activated during a sensory substitution task with the TDU, as measured with PET (Kupers et al., 2006). None of the trained blindfolded subjects reported TMS-induced tactile sensations on the tongue.

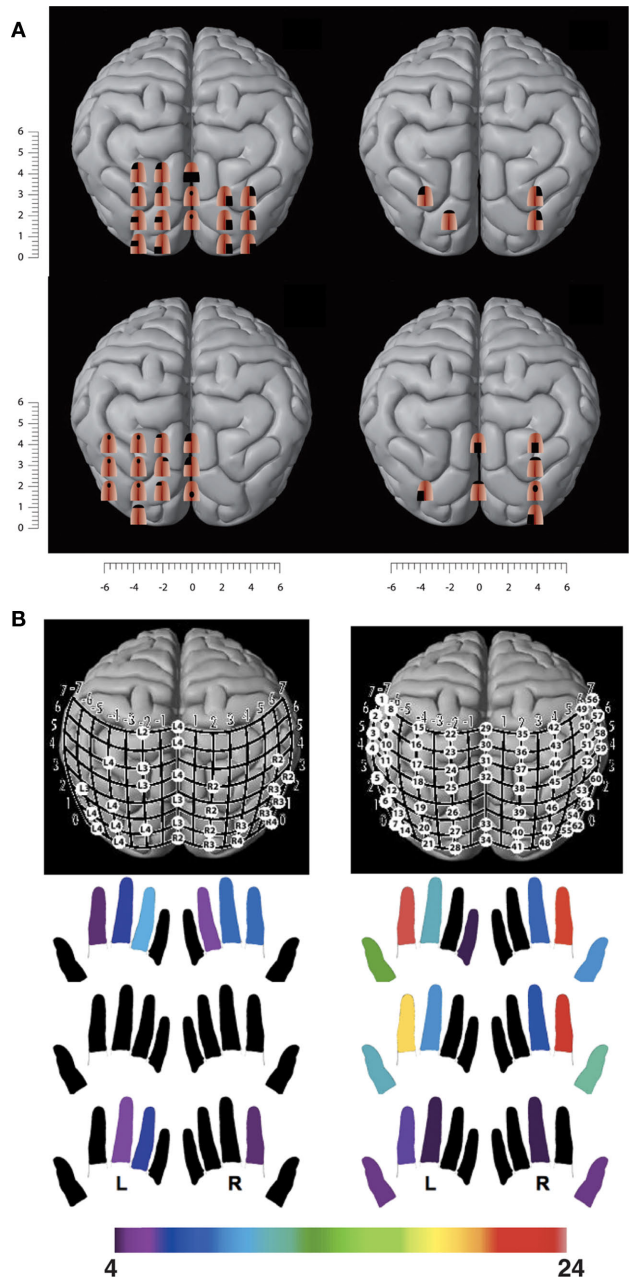


FIGURE 5 | TMS of the visual cortex in congenitally blind subjects induces tactile sensations. (A) TMS-induced tactile sensations referred to the tongue in congenitally blind subjects following a 1-week training with the TDU. **(B)** TMS-induced tactile sensations referred to the fingertips in two congenitally blind proficient Braille readers. The color map below indicates the fingers in which the subject experienced TMS-induced tactile sensations. The number of visual cortex sites from which paresthesiae could be induced in a particular finger is color-coded, with red indicating the highest number of cortical sites that induced paresthesiae in a particular finger and purple the lowest number (adapted after Kupers et al., 2006; Ptito et al., 2008a).

If tactile sensations referred to the tongue can be induced by TMS of the occipital cortex in blind subjects trained for only 1 week with the TDU, TMS should also induce tactile sensations referred to the fingertips in proficient blind Braille readers. We addressed the question of remapping of the fingers onto the visual cortex in a group of blind Braille readers and Braille-naive normal sighted controls (Ptito et al., 2008a). Again, TMS of the occipital cortex in control subjects only evoked phosphenes. In contrast, blind subjects reported tactile sensations in the fingers, varying in intensity, extent, and topography depending on the stimulated occipital area (**Figure 5B**). We found again important inter-individual differences with respect to the number of sites from which tactile sensations could be induced and in the topography of the referred sensations. The subject with the highest amount of finger paresthesiae was the one with the best Braille reading performance.

CORTICAL DOMINANCE AND CORTICAL DEFERENCE

The results of these TMS studies constitute the first direct demonstration that the subjective experience of activity in the visual cortex after sensory remapping is tactile, not visual. These findings provide new insights into the long-established scientific debate on cortical dominance or deference (James, 1890; Hurley and Noë, 2003). What is the experience of a subject in whom areas of cortex receive input from sensory sources not normally projecting to those areas? Our studies suggest that the qualitative character of the subject's experience is not determined by the area of cortex that is active (cortical dominance), but by the source of input to it (cortical deference). Our results are in line with evidence that sensory cortical areas receive input from multiple sensory modalities early in development (Falchier et al., 2002; Rockland and Ojima, 2003; Wallace et al., 2004).

CORTICAL REORGANIZATION OR UNMASKING?

Two competing hypotheses have been put forward to explain the neural mechanism driving this cross-modal plasticity. According to the *cortical reorganization* hypothesis, cross-modal brain responses are mediated by the formation of new pathways in the sensory-deprived brain. Studies in animals have shown that when the brain is deprived of visual input at an early age, tactile and other non-visual information is re-routed to the visual cortex (e.g., Hyvarinen et al., 1991; Chabot et al., 2008). This claim is largely supported by results of functional brain imaging studies showing activation of visual cortex in early blind subjects during Braille reading (Sadato et al., 1996; Burton et al., 2002) and other forms of tactile stimulation (Burton et al., 2004; Ptito et al., 2005). The importance of visual deprivation in early life is underscored by the observation that brain activity patterns in occipital cortex evoked by tactile stimulation are significantly stronger in early blind compared to late blind subjects (Cohen et al., 1999; Burton et al., 2002). According to the *unmasking* hypothesis, loss of a sensory input induces unmasking and strengthening of existing neuronal connections. The rapid onset of cross-modal responses with the TDU (within 1 week) excludes indeed the possibility that these are mediated by the establishment of new anatomical connections and therefore favors the unmasking hypothesis. One possibility is that training unmasks and strengthens pre-existing connections between the parietal and the occipital cortices. This claim is supported by

electrophysiological (Fishman and Michael, 1973) and anatomical (Falchier et al., 2002; Rockland and Ojima, 2003) data showing that primary visual cortex in normal mammals receives input not only from the visual thalamus, but also from somatosensory and auditory modalities. Single unit recordings in the cat visual cortex have shown that neurons in areas 17 and 18 receive both visual and auditory input (Fishman and Michael, 1973). In addition, there is evidence for direct projections from the auditory cortex to area 17 in the macaque monkey (Falchier et al., 2002). There are also direct projections from parietal association areas to areas V1 and V2 in the calcarine fissure (Rockland and Ojima, 2003). Under normal circumstances, these non-visual inputs conveying tactile and auditory inputs to occipital cortex can modulate the processing of visual information (Macaluso et al., 2000), while not giving rise to subjective non-visual sensations due to masking by the dominant visual input. This is illustrated by a study showing that TMS over the visual cortex impairs tactile discrimination of grating orientation in normal seeing subjects, without producing subjective tactile sensations (Zangaladze et al., 1999). Thus, in our trained control subjects, TMS over occipital cortex produced phosphenes and not tactile sensations. However, under certain circumstances such as following early visual deprivation, non-visual processing in the occipital cortex is strengthened or unmasked. In line with the dynamic sensorimotor hypothesis, training with the TDU results in new highly specific learned dynamic interaction patterns between sensory stimulation and active movement (O'Regan and Noë, 2001), thereby further strengthening and unmasking existing connections between the parietal and occipital cortices.

A DARWINIAN STRUGGLE FOR SURVIVAL?

There is now a wealth of data showing that the occipital cortex in the blind is activated in a variety of tasks, including lexical and phonological processing (Röder et al., 2002; Burton et al., 2003; Amedi et al., 2004), verbal memory (Burton et al., 2003; Raz et al., 2005), repetition priming (Kupers et al., 2007), auditory discrimination (Röder et al., 1999; Weeks et al., 2000; Gougoux et al., 2005), selective attention (Stevens et al., 2007), working memory (Pietrini et al., 2004; Bonino et al., 2008), motion detection (Ricciardi et al., 2007), and spatial navigation (Kupers et al., 2010). How to interpret this multiplicity of sensory and cognitive functions of the occipital cortex in the blind? Does it reflect some kind of Darwinian principle of struggle for survival? As humans, we are living in a world in which vision has a central role. This is already reflected by the fact that the visual cortex in primates covers about 30% of the total cortical surface. Therefore, the loss of vision is one of the most incapacitating events that can occur to a person. In order to survive, blind subjects need to rely on other senses and develop these in a supranormal manner to compensate for their loss of vision. Functional brain imaging studies have shown that enhanced practice leads to an enlargement of cortical representations (Gaser and Schlaug, 2003; Draganski et al., 2004; Scholz et al., 2009). In the sighted brain, this results in an enlargement of the cortex that is normally involved in the execution of the task (e.g., an expansion of the motor cortex in musicians; Lotze et al., 2003; Bengtsson et al., 2005) and not by the recruitment of novel cortex. In contrast, in the case of loss of a sensory input, the brain recruits *de novo* cortex that is normally not involved in execution of that particular task. This may represent a Darwinian mechanism for survival. Recruitment of the visually deprived occipital cortex is a much more cost-effective computational solution since it does not put extra demands on the cortex which is normally used for executing that task, thereby leaving sufficient resources available for situations of increased demand. The pathways through which this occurs are available since birth but, in the sighted brain, the activity is masked by the dominant specific afferent input to a particular cortical region.

How does the rewired cortex cope with this multitude of new inputs? Is there some kind of segregation of functions or does the visually deprived cortex become genuinely multimodal? It is difficult to answer this question since most studies only investigated one or at best two different sensory modalities or cognitive functions in the same subjects. One of the first studies that investigated the cortical representation of multiple cognitive functions in the occipital cortex of the blind suggested that different functions are indeed segregated anatomically (Amedi et al., 2003). In contrast, a more recent study reported overlap of auditory and tactile information in the occipital cortex of early blind subjects (Renier et al., 2010). However, the right middle occipital gyrus showed a preference for spatial over non-spatial processing of both auditory and tactile stimuli. More studies are needed to resolve this issue.

FINAL CONSIDERATIONS

Animal and human studies of the dark-reared brain have shed new light on many questions regarding not only the plastic rearrangements that take place in the absence of vision but also on the development of the functional architecture of the sighted brain itself. The availability of methodologies for the non-invasive functional exploration of the brain has made it possible to begin to understand the neural mechanisms that enable awareness of the surrounding world and to make sense of it. The main hypothesis that we have put forward here is that the development of consciousness in the absence of vision is made possible through the supramodal nature of functional cortical organization. The more abstract representation of the concepts of objects, space, motion, gestures, and actions – in one term, awareness of the external world – is associated with regional brain activation patterns that are essentially similar in sighted and congenitally blind individuals (Pietrini et al., 2004, 2009). The morphological and/or functional differences that exist between the sighted and the blind brain are the consequence of the cross-modal plastic reorganization that mostly affects that part of the cortex that is not multimodal in nature.

We would also like to acknowledge that the issues that we have considered are only a few among the many more that an ambitious topic such as the relation between blindness and consciousness may raise. For instance, we did not discuss the effects of congenital blindness versus blindness acquired at different ages, or the effects of monocular vision (Vecchi et al., 2006). Moreover, we have only briefly touched upon the "blind social brain," not to mention emotional life and its disturbances.

A final important thought prompted by the many different findings from studies in animal and humans is that the blind brain should not be considered as a "disabled" brain but rather as a truly "differentially able" brain.

ACKNOWLEDGMENTS

Studies reported in this paper have been supported by grants from the following foundations and granting agencies: the Lundbeck Foundation (to Ron Kupers), the European Union (Grant IST-2001-38040 and IST-2006-027141 to Pietro Pietrini), the Italian Ministry of Education, University and Research (PRIN RBNE018ET9-003, 200411841, and 2006117208 to Pietro Pietrini), Fondazione IRIS, Castagneto Carducci (Livorno, Italy to Pietro Pietrini), the Danish Medical Research Council (to Maurice Ptito), and the Harland Sanders Foundation (to Maurice Ptito). Pietro Pietrini and Emiliano Ricciardi wish to thank the Unione Italiana Ciechi for its support to the blindness research program, and the MRI Laboratory at the Fondazione "Gabriele Monasterio" Regione Toscana/CNR (Pisa, Italy).

REFERENCES

Amedi, A., Floel, A., Knecht, S., Zohary, E., and Cohen, L. G. (2004). Transcranial magnetic stimulation of the occipital pole interferes with verbal processing in blind subjects. *Nat. Neurosci.* 7, 1266–1270.

Amedi, A., Jacobson, G., Hendler, T., Malach, R., and Zohary, E. (2002). Convergence of visual and tactile shape processing in the human lateral occipital complex. *Cereb. Cortex* 12, 1202–1212.

Amedi, A., Malach, R., Hendler, T., Peled, S., and Zohary E. (2001). Visuo-haptic object-related activation in the ventral visual pathway. *Nat. Neurosci.* 4, 324–330.

Amedi, A., Raz, N., Pianka, P., Malach, R., and Zohary, E. (2003). Early 'visual' cortex activation correlates with superior verbal memory performance in the blind. *Nat. Neurosci.* 6, 758–766.

Bach-y-Rita, P., and Kercel, S. (2003). Sensory substitution and the human–machine interface. *Trends Cogn. Sci.* 7, 541–546.

Baumgaertner, A., Buccino, G., Lange, R., McNamara, A., and Binkofski, F. (2007). Polymodal conceptual processing of human biological actions in the left inferior frontal lobe. *Eur. J. Neurosci.* 25, 881–889.

Bengtsson, S. L., Nagy, Z., Skare, S., Forsman, L., Forssberg, H., and Ullén, F. (2005). Extensive piano practicing has regionally specific effects on white matter development. *Nat. Neurosci.* 8, 1148–1150.

Bicchi, A., Scilingo, E. P., Ricciardi, E., and Pietrini, P. (2008). Tactile flow explains haptic counterparts of common visual illusions. *Brain Res. Bull.* 75, 737–741.

Bittar, R. G., Ptito, M., Faubert, J., Dumoulin, S. O., and Ptito, A. (1999). Activation of the remaining hemisphere following stimulation of the blind hemifield in hemispherectomized subjects. *Neuroimage* 10, 3339–3346.

Bonino, D., Ricciardi, E., Sani, L., Gentili, C., Vanello, N., Guazzelli, M., Vecchi, T., and Pietrini, P. (2008). Tactile spatial working memory activates the dorsal extrastriate cortical pathway in congenitally blind individuals. *Arch. Ital. Biol.* 146, 133–146.

Boroojerdi, B., Bushara, K. O., Corwell, B., Immisch, I., Battaglia, F., Muellbacher, W., and Cohen, L. G. (2000). Enhanced excitability of the human visual cortex induced by short-term light deprivation. *Cereb. Cortex* 10, 529–534.

Bronchti, G., Heil, P., Sadka, R., Hess, A., Scheich, H., and Wollberg, Z. (2002). Auditory activation of "visual" cortical areas in the blind mole rat (*Spalax ehrenbergi*). *Eur. J. Neurosci.* 16, 311–329.

Bruno, M. A., Vanhaudenhuyse, A., Schnakers, C., Boly, M., Gosseries, O., Demertzi, A., Majerus, S., Moonen, G., Hustinx, R., and Laureys, S. (2010). Visual fixation in the vegetative state: an observational case series PET study. *BMC Neurol.* 10:35. doi: 10.1186/1471-2377-10-35

Buccino, G., Binkofski, F., and Riggio, L. (2004a). The mirror neuron system and action recognition. *Brain Lang.* 89, 370–376.

Buccino, G., Vogt, S., Ritzl, A., Fink, G. R., Zilles, K., Freund, H. J., and Rizzolatti, G. (2004b). Neural circuits underlying imitation learning of hand actions: an event-related fMRI study. *Neuron* 42, 323–334.

Burton, H., Diamond, J. B., and McDermott, K. B. (2003). Dissociating cortical regions activated by semantic and phonological tasks: a FMRI study in blind and sighted people. *J. Neurophysiol.* 90, 1965–1982.

Burton, H., Sinclair, R. J., and McLaren, D. G. (2004). Cortical activity to vibrotactile stimulation: an fMRI study in blind and sighted individuals. *Hum. Brain Mapp.* 23, 210–228.

Burton, H., Snyder, A. Z., Conturo, T. E., Akbudak, E., Ollinger, J. M., and Raichle, M. E. (2002). Adaptive changes in early and late blind: a fMRI study of Braille reading. *J. Neurophysiol.* 87, 589–607.

Cattaneo, Z., Vecchi, T., Cornoldi, C., Mammarella, I., Bonino, D., Ricciardi, E., and Pietrini P. (2008). Imagery and spatial processes in blindness and visual impairment. *Neurosci. Biobehav. Rev.* 32, 1346–1360.

Chabot, N., Charbonneau, V., Laramée, M. E., Tremblay, R., Boire, D., and Bronchti, G. (2008). Subcortical auditory input to the primary visual cortex in anophthalmic mice. *Neurosci. Lett.* 433, 129–134.

Chebat, D. R., Rainville, C., Kupers, R., and Ptito, M. (2007). Alterations in right posterior hippocampus in early blind individuals. *Neuroreport* 18, 329–333.

Cohen, L. G., Weeks, R. A., Sadato, N., Celnik, P., Ishii, K., and Hallett, M. (1999). Period of susceptibility for cross-modal plasticity in the blind. *Ann. Neurol.* 45, 451–460.

Cowey, A. (2010). The blindsight saga. *Exp. Brain Res.* 200, 3–24.

Cowey, A., and Stoerig, P. (1995). Blindsight in monkeys. *Nature* 373, 247–249.

Cowey, A., and Walsh, V. (2000). Magnetically induced phosphenes in sighted, blind and blindsighted observers. *Neuroreport* 11, 3269–3273.

Dehaene, S., Changeux, J. P., Naccache, L., Sackur, J., and Sergent, C. (2006). Conscious, preconscious, and subliminal processing: a testable taxonomy. *Trends Cogn. Sci.* 10, 204–211.

Desgent, S., Boire, D., and Ptito, M. (2005). Distribution of calcium binding proteins in visual and auditory cortices of hamsters. *Exp. Brain Res.* 163, 159–172.

Desgent, S., Boire, D., and Ptito, M. (2010). Altered expression of parvalbumin and calbindin in interneurons within the primary visual cortex of neonatal enucleated hamsters. *Neuroscience.* doi: 10.1016/j.neuroscience.2010.10.016. [Epub ahead of print].

Draganski, B., Gaser, C., Busch, V., Schuierer, G., Bogdahn, U., and May, A. (2004). Neuroplasticity: changes in grey matter induced by training. *Nature* 427, 311–312.

Fabbri-Destro, M., and Rizzolatti, G. (2008). Mirror neurons and mirror systems in monkeys and humans. *Physiology (Bethesda)* 23, 171–179.

Falchier, A., Clavagnier, S., Barone, P., and Kennedy, H. (2002). Anatomical evidence of multimodal integration in primate striate cortex. *J. Neurosci.* 22, 5749–5759.

Fiehler, K., Burke, M., Bien, S., Röder, B., and Rösler, F. (2009). The human dorsal action control system develops in the absence of vision. *Cereb. Cortex* 19, 1–12.

Fishman, M. C., and Michael, P. (1973). Integration of auditory information in the cat's visual cortex. *Vision Res.* 13, 1415–1419.

Fortin, M., Voss, P., Lord, C., Lassonde, M., Pruessner, J., Saint-Amour, D., Rainville, C., and Lepore, F. (2008). Wayfinding in the blind: larger hippocampal volume and supranormal spatial navigation. *Brain* 131, 2995–3005.

Frost, D. O., Boire, D., Gingras, G., and Ptito, M. (2000). Surgically created neural pathways mediate visual pattern discrimination. *Proc. Natl. Acad. Sci. U.S.A.* 97, 11068–11073.

Frost, D. O., and Metin, C. (1985). Induction of functional retinal projections to the somatosensory system. *Nature* 317, 162–164.

Fujii, T., Tanabe, H. C., Kochiyama, T., and Sadato, N. (2009). An investigation of cross-modal plasticity of effective connectivity in the blind by dynamic causal modeling of functional MRI data. *Neurosci. Res.* 65, 175–186.

Gaglianese, A., Sani, L., Roebroeck, A., Handjaras, G., Bernardi, G., Costagli, M., Ricciardi, E., and Pietrini, P. (2010). "fMRI functional and effective connectivity in humans supports a direct pathway from thalamus to hMT+," in *Presented at the 16th Annual Meeting of the Organization for Human Brain Mapping*, Barcelona.

Galati, G., Committeri, G., Spitoni, G., Aprile, T., Di Russo, F., Pitzalis, S., and Pizzamiglio, L. (2008). A selective representation of the meaning of actions in the auditory mirror system. *Neuroimage* 40, 1274–1286.

Gallese, V., Fadiga, L., Fogassi, L., and Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain* 119, 593–609.

Gaser, C., and Schlaug, G. (2003). Brain structures differ between musicians and non-musicians. *J. Neurosci.* 23, 9240–9245.

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

Gauthier, I., Tarr, M. J., Anderson, A. W., Skudlarski, P., and Gore, J. C. (1999). Activation of the middle fusiform 'face area' increases with expertise in recognizing novel objects. *Nat. Neurosci.* 2, 568–573.

Gazzola, V., Aziz-Zadeh, L., and Keysers, C. (2006). Empathy and the somatotopic auditory mirror system in humans. *Curr. Biol.* 16, 1824–1829.

Gougoux, F., Zatorre, R. J., Lassonde, M., Voss, P., and Lepore, F. (2005). A func-

tional neuroimaging study of sound localization: visual cortex activity predicts performance in early-blind individuals. *PLoS Biol.* 3, e27. doi: 10.1371/j.pbio.0030027

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

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

Hurley, S. L., and Noë, A. (2003). Neural plasticity and consciousness. *Biol. Philos.* 18, 131–168.

Hyvarinen, J., Hyvärinen, L., and Linnankoski, I. (1991). Modification of parietal association cortex and functional blindness after binocular deprivation in young monkeys. *Exp. Brain Res.* 42, 1–8.

Ishai, A. (2010). Seeing faces and objects with the "mind's eye". *Arch. Ital. Biol.* 148, 1–9.

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

Ishai, A., Ungerleider, L. G., Martin, A., Schouten, J. L., and Haxby, J. V. (1999). Distributed representation of objects in the human ventral pathway. *Proc. Natl. Acad. Sci. U.S.A.* 96, 9379–9384.

James, T. W., Humphrey, G. K., Gati, J. S., Servos, P., Menon, R. S., and Goodale, M. A. (2002). Haptic study of three-dimensional objects activates extrastriate visual areas. *Neuropsychologia* 40, 1706–1714.

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

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

Klinge, C., Eippert, F., Röder, B., and Büchel, C. (2010). Corticocortical connections mediate primary visual cortex responses to auditory stimulation in the blind. *J. Neurosci.* 30, 12798–12805.

Kohler, E., Keysers, C., Umilta, M. A., Fogassi, L., Gallese, V., and Rizzolatti, G. (2002). Hearing sounds, understanding actions: action representation in mirror neurons. *Science* 297, 846–848.

Kupers, R., Chebat, D. R., Madsen, K. H., Paulson, O. B., and Ptito, M. (2010). Neural correlates of virtual route recognition in congenital blindness. *Proc. Natl. Acad. Sci. U.S.A.* 107, 12716–12721.

Kupers, R., Christensen, R., Grey, M., and Ptito, M. (2009). Glucose metabolism of the occipital cortex in early blind subjects. *Soc. Neurosci. Abstr.* 167.26.

Kupers, R., Fumal, A., de Noordhout, A. M., Gjedde, A., Schoenen, J., and Ptito, M. (2006). Transcranial magnetic stimulation of the visual cortex induces somatotopically organized qualia in blind subjects. *Proc. Natl. Acad. Sci. U.S.A.* 103, 13256–13260.

Kupers, R., Pappens, M., Ptito, M., and Fumal, A. (2007). rTMS of the occipital cortex abolishes Braille reading and repetition priming in blind subjects. *Neurology* 68, 691–693.

Lahav, A., Saltzman, E., and Schlaug, G. (2007). Action representation of sound: audiomotor recognition network while listening to newly acquired actions. *J. Neurosci.* 27, 308–314.

Lamme, V. A. (2006). Towards a true neural stance on consciousness. *Trends Cogn. Sci.* 10, 494–501.

Laureys, S., Faymonville, M. E., Peigneux, P., Damas, P., Lambermont, B., Del Fiore, G., Degueldre, C., Aerts, J., Luxen, A., Franck, G., Lamy, M., Moonen, G., and Maquet, P. (2002). Cortical processing of noxious somatosensory stimuli in the persistent vegetative state. *Neuroimage* 17, 732–741.

Leh, S. E., Chakravarty, M. M., and Ptito, A. (2008). The connectivity of the human pulvinar: a diffusion tensor imaging tractography study. *Int. J. Biomed. Imaging*. doi 10.11555/2008/789539

Leh, S. E., Johansen-Berg, H., and Ptito, A. (2006). Unconscious vision: new insights into the neuronal correlate of blindsight using diffusion tractography. *Brain* 129, 1822–1832.

Lotze, M., Scheler, G., Tan, H. R. M., Braun, C., and Birbaumer N. (2003). The musician's brain: functional imaging of amateurs and professionals during performance and imagery. *Neuroimage* 20, 1817–1829.

Lyckman, A. W., and Sur, M. (2002). Role of afferent activity in the development of cortical specification. *Results Probl. Cell Differ.* 39, 139–156.

Macaluso, E., Frith C. D., and Driver. J. (2000). Modulation of human visual cortex by crossmodal spatial attention. *Science* 289, 1206–1208.

Matteau, I., Kupers, R., and Ptito, M. (2008). Tactile shape recognition through the tongue in the congenitally blind. *FENS. Abstr* 2008; 4:153.8.p.83

Matteau, I., Kupers, R., Ricciardi, E., Pietrini, P., and Ptito, M. (2010). Beyond visual, aural and haptic movement perception: hMT+ is activated by electrotactile motion stimulation of the tongue in sighted and in congenitally blind individuals. *Brain Res. Bull.* 82, 264–270.

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

Merabet, L. B., and Pascual-Leone, A. (2010). Neural reorganization following sensory loss: the opportunity of change. *Nat. Rev. Neurosci.* 11, 44–52.

Metin, C., and Frost, D. O. (1989). Visual responses of neurons in somatosensory cortex of hamsters with experimentally induced retinal projections to somatosensory thalamus. *Proc. Natl. Acad. Sci. U.S.A.* 86, 357–361.

Noppeney, U., Friston, K. J., Ashburner, J., Frackowiak, R., and Price, C. J. (2005). Early visual deprivation induces structural plasticity in gray and white matter. *Curr. Biol.* 15, 488–490.

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

O'Regan, J. K., and Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behav. Brain Sci.* 24, 939–973.

Owen, A. M., Schiff, N. D., and Laureys, S. (2009). "The assessment of conscious awareness in the vegetative state," in *The Neurology of Consciousness*, eds S. Laureys and G. Tononi (London: Academic Press), 163–172.

Pallas, S. L. (2001). Intrinsic and extrinsic factors that shape neocortical specification. *Trends Neurosci.* 24, 417–423.

Pallas, S. L. (2002). "Cross-modal plasticity as a tool for understanding the ontogeny and phylogeny of cerebral cortex," in *Cortical Areas: Unity and Diversity*, eds A. Schuz and R. Miller (London/New York: Taylor and Francis Inc.), 245–271.

Pan, W. J., Wu, G., Li, C. X., Lin, F., Sun, J., and Lei, H. (2007). Progressive atrophy in the optic pathway and visual cortex of early blind Chinese adults: a voxel-based morphometry magnetic resonance imaging study. *Neuroimage* 37, 212–220.

Penfield, W., and Boldrey, E. (1937). Somatic motor and sensory representation in the cerebral cortex of man as studied by electrical stimulation. *Brain* 60, 389–443.

Pietrini, P., Furey, M. L., Ricciardi, E., Gobbini, M. I., Wu, W. H., Cohen, L., and Guazzelli, M. Haxby, J. V. (2004). Beyond sensory images: object-based representation in the human ventral pathway. *Proc. Natl. Acad. Sci. U.S.A.* 101, 5658–5663.

Pietrini, P., Ptito, M., and Kupers, R. (2009). "Blindness and consciousness: new lights from the dark," in *The Neurology of Consciousness*, eds S. Laureys and G. Tononi (New York: Academic Press), 360–374.

Poirier, C., Collignon, O., Scheiber, C., Renier, L., Vanlierde, A., Tranduy, D., Veraart, C., and De Volder, A. G. (2006). Auditory motion perception activates visual motion areas in early blind subjects. *Neuroimage* 31, 279–285.

Ptito, A., and Leh, S. E. (2007). Neural substrates of blindsight after hemispherectomy. *Neuroscientist* 13, 506–518.

Ptito, M., and Desgent, S. (2006). "Sensory input-based adaptation and brain architecture," in *Lifespan Development and the Brain*, eds P. Baltes, P. Reuter-Lorenz, and F. Rösler (Cambridge University Press), 111–123.

Ptito, M., Fumal, A., de Noordhout, A. M., Schoenen, J., Gjedde, A., and Kupers, R. (2008a). TMS of the occipital cortex induces tactile sensations in the fingers of blind Braille readers. *Exp. Brain Res.* 184, 193–200.

Ptito, M., Schneider, F., Paulson, O. B., and Kupers, R. (2008b). Alterations of the visual pathways in congenital blindness. *Exp. Brain Res.* 87, 41–49.

Ptito, M., Giguère, J. F., Boire, D., Frost, D. O., and Casanova, C. (2001a). When the auditory cortex turns visual. *Prog. Brain Res.* 134, 447–458.

Ptito, M., Kupers, R., Faubert, J., and Gjedde, A. (2001b). Cortical representation of inward and outward radial motion in man. *Neuroimage* 14, 1409–1415.

Ptito, M., Moesgaard, S., Gjedde, A., and Kupers, R. (2005). Cross-modal plasticity revealed by electrotactile stimulation of the tongue in the congenitally blind. *Brain* 128, 606–614.

Radoeva, P. D., Prasad, S., Brainard, D. H., and Aguirre, G. K. (2008). Neural activity within area v1 reflects unconscious visual performance in a case of blindsight. *J. Cogn. Neurosci.* 20, 1927–1939.

Raz, N., Amedi, A., and Zohary, E. (2005). V1 activation in congenitally blind humans is associated with episodic retrieval. *Cereb. Cortex* 15, 1459–1468.

Raz, N., Striem, E., Pundak, G., Orlov, T., and Zohary, E. (2007). Superior serial memory in the blind: a case of cognitive compensatory adjustment. *Curr. Biol.* 17, 1129–1133.

Rebillard, G., Carlier, E., Rebillard, M., and Pujol, R. (1977). Enhancement of visual responses on the primary auditory cortex of the cat after an early destruction of cochlear receptors. *Brain Res.* 129, 162–164.

Renier, L. A., Anurova, I., De Volder, A. G., Carlson, S., VanMeter, J., and Rauschecker, J. P. (2010). Preserved functional specialization for spatial

processing in the middle occipital gyrus of the early blind. *Neuron* 68, 138–148.

Ricciardi, E., Basso, D., Sani, L., Bonino, D., Vecchi, T., Pietrini, P., and Miniussi, C. (2011). Functional inhibition of the human MT+ cortex affects non-visual motion perception: a repetitive transcranial magnetic stimulation study during tactile speed discrimination. *Exp. Biol. Med.* 236. (in press).

Ricciardi, E., Bonino, D., Gentili, C., Sani, L., Pietrini, P., and Vecchi, T. (2006a). Neural correlates of spatial working memory in humans: a functional magnetic resonance imaging study comparing visual and tactile processes. *Neuroscience* 139, 339–349.

Ricciardi, E., Sani, L., Gentili, C., Bonino, D., Vanello, N., Haxby, J. V., Seifritz, E., Guazzelli, M., Landini, L., Bicchi, A., Di Salle, F., and Pietrini, P. (2006b). Brain response to visual, tactile and auditory flow in sighted and blind individuals supports a supramodal functional organization in hMT+ complex. *Neuroimage* 31(Suppl. 1), 512 TH-PM.

Ricciardi, E., Bonino, D., Sani, L., Vecchi, T., Guazzelli, M., Haxby, J. V., Fadiga, L., and Pietrini, P. (2009). Do we really need vision? How blind people "see" the actions of others. *J. Neurosci.* 29, 9719–9124.

Ricciardi, E., Vanello, N., Sani, L., Gentili, C., Scilingo, E. P., Landini, L., Guazzelli, M., Bicchi, A., Haxby, J. V., and Pietrini, P. (2007). The effect of visual experience on the development of functional architecture in hMT+. *Cereb. Cortex* 17, 2933–2939.

Rizzolatti, G. (2005). The mirror neuron system and its function in humans. *Anat. Embryol. (Berl)*. 210, 419–421.

Rizzolatti, G., and Craighero, L. (2004). The mirror-neuron system. *Annu. Rev. Neurosci.* 27, 169–192.

Rizzolatti, G., and Fadiga, L. (1998). Grasping objects and grasping action meanings: the dual role of monkey rostroventral premotor cortex (area F5). *Novartis Found. Symp.* 218, 81–95; discussion 103.

Rizzolatti, G., Fadiga, L., Gallese, V., and Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Brain Res. Cogn. Brain Res.* 3, 131–141.

Rizzolatti, G., and Sinigaglia, C. (2008). Further reflections on how we interpret the actions of others. *Nature* 455, 589.

Rockland, K. S., and Ojima, H. (2003). Multisensory convergence in calcarine visual areas in macaque monkey. *Int. J. Psychophysiol.* 50,19–26.

Röder, B., Stick, O., Bien, S., Neville, H., and Rösler, F. (2002). Speech processing activates visual cortex in congenitally blind humans. *Eur. J. Neurosci.* 16, 930–936.

Röder, B., Teder-Sälejärvi, W., Sterr, A., Rösler, F., Hillyard, S. A., and Neville, H. (1999). Improved auditory spatial tuning in blind humans. *Nature* 400, 162–166.

Sadato, N., Pascual-Leone, A., Grafman, J., Ibanez, V., Deiber, M. P., Dold, G., and Hallett, M. (1996). Activation of the primary visual cortex by Braille reading in blind subjects. *Nature* 380, 526–528.

Sani, L., Ricciardi, E., Gentili, C., Vanello, N., Haxby, J. V., and Pietrini, P. (2010). Effects of visual experience on the human MT+ functional connectivity networks: an fMRI study of motion perception in sighted and congenitally blind individuals. *Front. Syst. Neurosci.* 4:159. doi: 10.3389/fnsys.2010.00159

Sathian, K., and Zangaladze, A. (2002). Feeling with the mind's eye: contribution of visual cortex to tactile perception. *Behav. Brain Res.* 135, 127–132.

Schmid, M. C., Mrowka, S. W., Turchi, J., Saunders, R. C., Wilke, M., Peters, A. J., Ye, F. Q., and Leopold, D. A. (2010). Blindsight depends on the lateral geniculate nucleus. *Nature* 466, 373–377.

Scholz, J., Klein, M. C., Behrens, T. E., and Johansen-Berg, H. (2009). Training induces changes in white-matter architecture. *Nat. Neurosci.* 12, 1370–1371.

Shimony, J. S., Burton, H., Epstein, A. A., McLaren, D. G., Sun, S. W., and Snyder, A. Z. (2006). Diffusion tensor imaging reveals white matter reorganization in early blind humans. *Cereb. Cortex* 16, 1653–1661.

Stevens, A. A., Snodgrass, M., Schwartz, D., and Weaver, K. (2007). Preparatory activity in occipital cortex in early blind humans predicts auditory perceptual performance. *J. Neurosci.* 27, 10734–10741.

Théoret, H., Boire, D., Herbin, M., and Ptito, M. (2001). Anatomical sparing in the superior colliculus of hemispherectomized monkeys. *Brain Res.* 16, 274–280.

Toldi, J., Rojik, I., and Fehér, O. (1994). Neonatal monocular enucleation-induced cross-modal effects observed in the cortex of adult rat. *Neuroscience* 62, 105–114.

Tomaiuolo, F., Ptito, M., Marzi, C. A., Paus, T., and Ptito, A. (1997). Blindsight in hemispherectomized patients as revealed by spatial summation across the vertical meridian. *Brain* 120, 795–803.

Tusa, R. J., and Ungerleider, L. G. (1985). The inferior longitudinal fasciculus: a reexamination in humans and monkeys. *Ann. Neurol.* 18, 583–591.

Ungerleider, L. G., and Mishkin, M. (1982). "Two cortical visual systems," in *Analysis of Visual Behavior*, eds D. J. Ingle, M. A. Goodale, and R. J. W. Mansfield (Cambridge, MA: MIT Press), 549–586.

Vecchi, T., Cattaneo, Z., Monegato, M., Pece, A., Cornoldi, C., and Pietrini, P. (2006). Why Cyclops could not compete with Ulysses: monocular vision and mental images. *Neuroreport* 17, 723–726.

von Melchner, L., Pallas, S. L., and Sur, M. (2000). Visual behaviour mediated by retinal projections directed to the auditory pathway. *Nature* 404, 871–876.

Wallace, M. T., Ramachandran, R., and Stein, B. E. (2004). A revised view of sensory cortical parcellation. *Proc. Natl. Acad. Sci. U.S.A.* 101, 2167–2172.

Watson, J. D., Myers, R., Frackowiak, R. S., Hajnal, J. V., Woods, R. P., Mazziota, J. C., Shipp, S., and Zeki, S. (1993). Area V5 of the human brain: evidence from a combined study using positron emission tomography and magnetic resonance imaging. *Cereb. Cortex* 3, 79–94.

Weeks, R., Horwitz, B., Aziz-Sultan, A., Tian, B., Wessinger, C. M., Cohen, L. G., Hallett, M., and Rauschecker, J. P. (2000). A positron emission tomographic study of auditory localization in the congenitally blind. *J. Neurosci.* 20, 2664–2672.

Weiskrantz, L., Warrington, E. K., Sanders, M. D., and Marshall, J. (1974). Visual capacity in the hemianopic field following a restricted cortical ablation. *Brain* 97, 709–728.

Wittenberg, G. F. Werhahn, K. J., Wassermann, E. M., Herscovitch, P., and Cohen, L. G. (2004). Functional connectivity between somatosensory and visual cortex in early blind humans. *Eur. J. Neurosci.* 20, 1923–1927.

Zangaladze, A., Epstein, C. M., Grafton, S. T., and Sathian, K. (1999). Involvement of visual cortex in tactile discrimination of orientation. *Nature* 401, 587–590.

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

Received: 23 November 2010; paper pending published: 04 January 2011; accepted: 25 January 2011; published online: 14 February 2011.

Citation: Kupers R, Pietrini P, Ricciardi E and Ptito M (2011) The nature of consciousness in the visually deprived brain. Front. Psychology ***2****:19. doi: 10.3389/fpsyg.2011.00019*

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

frontiers in PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 07 February 2011
doi: 10.3389/fpsyg.2010.00245

Neural plasticity lessons from disorders of consciousness

Athena Demertzi, Caroline Schnakers, Andrea Soddu, Marie-Aurélie Bruno, Olivia Gosseries, Audrey Vanhaudenhuyse and Steven Laureys*

Coma Science Group, Cyclotron Research Centre and Neurology Department, Sart Tilman, University and University Hospital of Liège, Liège, Belgium

Edited by:
Morten Overgaard, Aarhus University, Aarhus University Hospital, Denmark

Reviewed by:
Ryota Kanai, University College London, UK
Jessica Saenger, Heinrich-Heine-University of Düsseldorf, Germany

****Correspondence:***
Steven Laureys, Cyclotron Research Center, Allée du 6 août No. 8, Sart Tilman B30, 4000 Liège, Belgium.
e-mail: steven.laureys@ulg.ac.be

Communication and intentional behavior are supported by the brain's integrity at a structural and a functional level. When widespread loss of cerebral connectivity is brought about as a result of a severe brain injury, in many cases patients are not capable of conscious interactive behavior and are said to suffer from disorders of consciousness (e.g., coma, vegetative state/unresponsive wakefulness syndrome, minimally conscious states). This lesion paradigm has offered not only clinical insights, as how to improve diagnosis, prognosis, and treatment, but also put forward scientific opportunities to study the brain's plastic abilities. We here review interventional and observational studies performed in severely brain-injured patients with regards to recovery of consciousness. The study of the recovered conscious brain (spontaneous and/or after surgical or pharmacologic interventions), suggests a link between some specific brain areas and the capacity of the brain to sustain conscious experience, challenging at the same time the notion of fixed temporal boundaries in rehabilitative processes. Altered functional connectivity, cerebral structural reorganization as well as behavioral amelioration after invasive treatments will be discussed as the main indices for plasticity in these challenging patients. The study of patients with chronic disorders of consciousness may, thus, provide further insights not only at a clinical level (i.e., medical management and rehabilitation) but also from a scientific-theoretical perspective (i.e., the brain's plastic abilities and the pursuit of the neural correlate of consciousness).

Keywords: neural plasticity, recovery, unresponsive wakefulness syndrome, vegetative state, minimally conscious state, consciousness, functional neuroimaging, deep brain stimulation

Thanks to last decades' technological advances, the study of consciousness has been under the scrutiny of neuroscientific research. The notion that consciousness is supported by and emerged from the brain is well-documented by clinical cases of neurological patients suffering from disorders of consciousness (DOC): in coma, patients are unable to be awakened and hence show no awareness of themselves and of their environment; in the vegetative state (VS) now called unresponsive wakefulness syndrome (UWS; Laureys et al., 2010) there is a dissociation between arousal which is preserved (i.e., clinically evident by eyes opening) and awareness which is abolished (Jennett and Plum, 1972). The recently defined minimally conscious state (MCS) describes patients who show fluctuating signs of awareness but remain unable to communicate (Giacino et al., 2002). Through this lesion paradigm we have the opportunity not only to better understand the neural correlates of consciousness (Tononi and Laureys, 2009) but also to gain insight about the brain's plastic abilities (Laureys et al., 2006a). In the present review, the study of neural plasticity is approached via neurological evidence coming from neuroimaging technologies, such as structural and functional magnetic resonance imaging (MRI), and positron emission tomography (PET), during pathological states and after recovery of consciousness. Altered cerebral functional connectivity, structural reorganization as well as behavioral amelioration after invasive and non-invasive treatments will be discussed as the main indices for plasticity in this challenging population (**Figure 1**).

INTERVENTIONAL STUDIES

DEEP BRAIN STIMULATION

Deep brain stimulation (DBS) is an interventional surgical procedure which requires the implantation of microelectrodes in deep structures of the brain and the administration of low voltage electricity in these structures. Despite some sparse evidence that DBS may have some ameliorating effects on arousal in VS/UWS patients (Cohadon and Richer, 1993; Yamamoto et al., 2001), in general one cannot argue in favor of this treatment in the VS/UWS population. This is mainly due to the widespread underlying neuropathology of VS/UWS (Adams et al., 2000) which does not permit a straightforward functional re-integration after stimulating the structures of interest in these patients (Schiff and Fins, 2007).

In a more controlled experimental setting, where patients' selection was based on both their neuropathological status (i.e., specific information about the connections between the central thalamus, cerebral cortex, basal ganglia and other subcortical structures) and behavioral profile (i.e., exhibition of preserved arousal and fluctuating behavioral performance), a 38-year-old patient in a MCS more than 6 years after severe traumatic brain injury was selected for DBS treatment (Schiff et al., 2007). Up to the point of DBS treatment, the patient did not show any clinical amelioration despite a 2-year rehabilitation program. However, after applying DBS in bilateral central intralaminar thalamic nuclei (**Figure 2**), the patient showed stimulation-related improved levels of arousal, motor control, and interactive behavior as measured by neuropsychological testing during the DBS "on" periods.

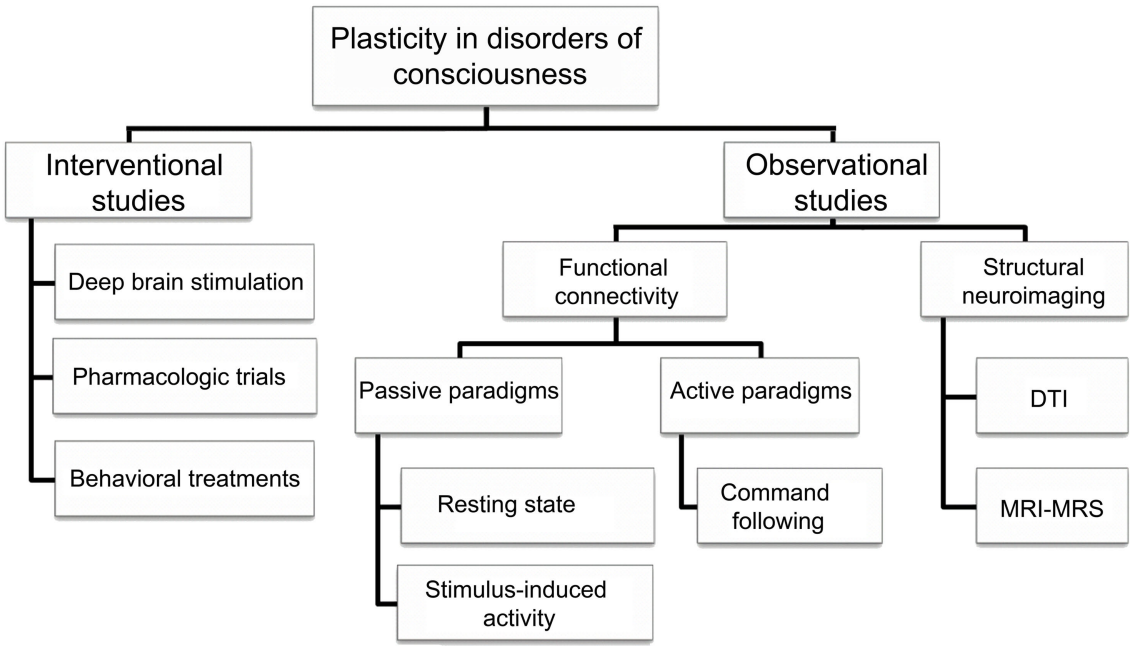


FIGURE 1 | Outline of studies dealing with plasticity issues in patients with disorders of consciousness. DTI, diffusion tensor imaging; MRI, magnetic resonance imaging; MRS, magnetic resonance spectroscopy.

The effects of DBS were attributed to the recruitable large-scale networks underlying the neuropathology of this MCS patient and were interpreted as a promotion of the patient's arousal regulation via the direct activation of the frontal cortical and basal ganglia systems, innervated by the stimulated thalamic neurons (Schiff et al., 2007; Schiff, 2010).

PHARMACOLOGIC TRIALS

Regarding the effects of pharmacologic trials in patients with DOC, generally speaking no satisfactory results exist (Laureys et al., 2006b). Small-scale pharmacologic studies indicate some exceptional respondents to either stimulant or depressant pharmacologic agents, but for whom no evidence-based recommendations can be made yet (Whyte et al., 2005; Demertzi et al., 2008).

Studies using amantadine, a mixed NMDA and dopaminergic agonist, showed a better outcome in traumatic patients with DOC (Whyte et al., 2005; Sawyer et al., 2008). In addition to behavioral amelioration, a recent PET study of chronic anoxic MCS showed a drug-related increase in fronto-parietal metabolism (Schnakers et al., 2008). Other dopaminergic agents which have been reported to lead to favorable functional outcome are levodopa and bromocriptine (Passler and Riggs, 2001). Clinical improvements has also been reported after administration of baclofen (GABA agonist administered mainly against spasticity; Taira and Hori, 2007) and zolpidem (non-benzodiazepine sedative drug that is used against insomnia in healthy people; for short review see Demertzi et al., 2008). The exact neuromodulating mechanism of these agents is not clear yet. A mesocircuit hypothesis, involving the cortico-thalamo-cortical system as well as projections of the basal ganglia to the central thalamus, has been recently proposed (Schiff, 2010). The dopaminergic agents are thought to either facilitate directly the mesio-frontal cortical neurons, which send excitatory projections to the central thalamus, or modulate the striatum leading to the restoration of the global dynamics of the cortico-thalamic system. On the other hand, zolpidem effects may be explained by a direct action at the level of the global pallidus interna which sends inhibitory projections to the central thalamus; this inhibitory effect could substitute for the normal inhibition of the globus pallidus from the striatum and hence lead to a more stabilized central thalamic activity (Schiff, 2010).

NON-PHARMACOLOGIC INTERVENTIONS

Non-pharmacologic interventions for DOC patients here refer to sensory stimulation techniques and physical therapy, which mainly aim at both preventing complications (i.e., contractures or pressure sore preventions) and/or at enhancing recovery. Sensory stimulation can refer to two types of approaches: multisensory stimulation or sensory regulation (Tolle and Reimer, 2003). The first expresses the principles of behaviorism and holds that enhanced environmental stimulation of the sensory systems is hoped to enhance synaptic re-innervations, whereas the second is based on the principles of information processing and focuses on the enhancement of selective attention by regulating the environment. Concerning physical therapy, there is some evidence that early (Oh and Seo, 2003) and increased intervention (Shiel et al., 2001) leads to better outcome.

PERSPECTIVES

The existing therapeutic nihilism in the field of DOC is currently getting challenged by recent data supporting that some DOC patients could benefit from some rehabilitative interventions (surgical, pharmacologic, or behavioral) reviewed above. Larger-scale

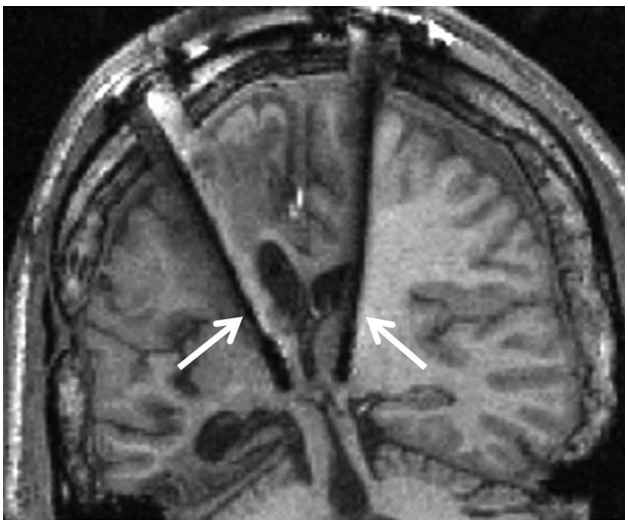


FIGURE 2 | Clinical setup of electrodes implementation in the central thalami bilaterally (white arrows) during deep brain stimulation. Adapted from Schiff et al. (2007).

studies with higher number of patients of various pathologies are ongoing, in order to better comprehend the underlying neuromodulatory effects of DBS and the induced neuroplastic changes in severely injured brains. Currently, the beneficial effects of the pharmacologic and non-pharmacologic approaches described above are not evidence-based and hence are not generally accepted by the medical community (for a systematic review see Lombardi et al., 2002). No unique hypothesis or theoretical framework (Laureys, 2005; Tononi and Laureys, 2009) can at present combine the temporal dynamics and pathophysiological mechanisms of all the aforementioned interventions (e.g., Pistoia et al., 2010) and many questions remain as to the precise mechanisms differentiating spontaneous from therapy-induced cerebral plasticity.

In the therapeutic management of patients with DOC, no "standards of care" do yet exist, mainly due to the limitation of their scientific evidence coming from small-scale studies under suboptimal or uncontrolled settings. Thus, no evidence-based recommendations can be made for a particular treatment option (Demertzi et al., 2008).

OBSERVATIONAL STUDIES

FUNCTIONAL CONNECTIVITY STUDIES

Passive paradigms

Functional neuroimaging does not only allow one to examine the functional segregation (i.e., localizing a function to a cerebral area) but also the functional integration (i.e., assessing the interaction between functionally segregated areas). Current analytical tools permit to assess the functional or effective connectivity between distant cerebral areas in functional imaging. Such analyses explain the activity in one cortical area in terms of an interaction between the influence of another area and some experimental condition (i.e., comparing data obtained during unconsciousness and after recovery). Functional connectivity is defined as the temporal correlation of a neurophysiological index (hemodynamic or metabolic) measured in different remote brain areas, whereas effective connectivity is defined as the influence one neural system exerts over another (Friston, 2002). Based on their experimental design, functional connectivity studies can be reduced to two main categories: passive (i.e., resting state, stimulus-induced activity) and active (i.e., mental task command following) paradigms (Bruno et al., 2010).

Using "resting state" fluorodeoxyglucose PET imaging, decreased global metabolic levels have been identified in VS/UWS patients, with no significant global metabolic resumption after recovery of consciousness. However, "functional disconnections" were identified in a large fronto-parietal network which exhibited regional metabolic restoration in long-range cortico-cortical (between latero-frontal and midline-posterior areas; Laureys et al., 1999) and cortico-thalamo-cortical (between non-specific thalamic nuclei and midline-posterior cortices) after recovery of consciousness from chronic VS/UWS (Laureys et al., 2000b; **Figure 3**). It is hence suggested that fronto-parietal network connectivity is critical in sustaining conscious awareness (Baars et al., 2003; Laureys et al., 2004a), as is also supported by evidence from studies on sensory perception in normal volunteers (Dehaene et al., 2006; Boly et al., 2007).

Positron emission tomography and fMRI studies have identified a "default mode network," defined as a set of areas, encompassing posterior cingulate/precuneus, anterior cingulate/mesio-frontal cortex, and temporo-parietal junctions, which show more activity at rest than during attention-demanding tasks. Recent studies have shown that it is possible to reliably identify this network in the absence of any task, by resting state fMRI connectivity analyses in healthy volunteers (Boly et al., 2008b, 2009). This "default mode network" is considered to be involved in self-related processes (Mason et al., 2007; Buckner et al., 2008; Vanhaudenhuyse et al., 2011) but the functional significance of these spontaneous brain activity fluctuations in pathological states remain only partially understood. It has recently been shown that default mode connectivity decreases during propofol general anesthesia (Boveroux et al., 2010), sleep (Gould et al., 1999), and hypnotic state (McGeown et al., 2009; Demertzi et al., in press). In pathological impaired consciousness, resting state connectivity was shown to disappear in brain death (Boly et al., 2009) and to show a non-linear disintegration in pseudocoma or locked-in syndrome as compared to minimally conscious or relative to unconscious states (VS/UWS or coma; Vanhaudenhuyse et al., 2009; **Figure 4**).

Apart from resting state acquisitions, valuable information is gathered by studies observing the cerebral responses to external sensory stimulation. Using PET, stimulus-induced somatosensory (Laureys et al., 2002; Boly et al., 2008a) and auditory (Laureys et al., 2000a; Boly et al., 2004) activation protocols in VS/UWS patients have identified a cerebral response restricted to primary sensory cortices, whereas MCS patients demonstrated a stronger functional connectivity between sensory and fronto-parietal associative areas in these patients. These findings indicate that the presence of isolated neuronal groups that work in a module-like fashion, are not functionally sufficient for the conscious perception of the world and the generation of conscious behavior (Schiff et al., 2002). Additionally, stimuli with emotional valence like infant cries (Laureys et al., 2004b) or the patient's own name

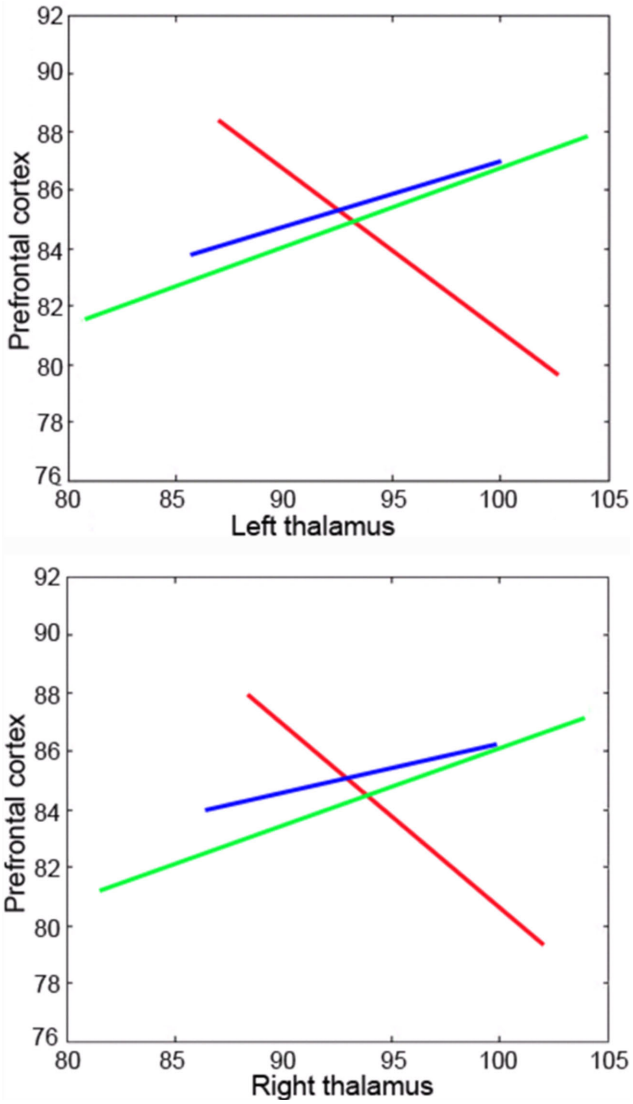


FIGURE 3 | Functional connectivity between the thalami and prefrontal cortex is abolished during pathological unconsciousness (UWS/VS; red regression line) and resumes a near-normal modulation after recovery of consciousness (blue line). The green line illustrates connectivity in healthy controls. Regression lines represent cross-correlation of normalized regional cerebral blood flow as measured by H_2O^{15}-PET. Adapted from Laureys et al. (2000b).

(Di et al., 2007) induced a widespread near-normal activation in MCS. The latter fMRI study also showed to be informative of patients' prognosis and recovery as confirmed by a recent study by Coleman et al. (2007).

Active paradigms

"Active paradigms" in neuroimaging studies, aiming to show command following, constitute a more direct proof to demonstrate conscious awareness, independent of motor activity. If a patient systematically follows a specific mental command, then this subject is expected to activate certain brain areas in a consistent manner and only then one can infer that this subject is conscious. Using this approach, a collaborative study between the Cambridge and Liège imaging centers, we showed in a clinically VS/UWS patient fMRI evidence of obeying to simple commands (i.e., "imagine walking around in your house" or "imaging playing tennis") in specific brain areas (i.e., parahippocampal and supplementary motor areas, respectively) indistinguishable from that observed in healthy controls (Owen et al., 2006). Such activation could not be attributed to automatic recruitment of these areas of interest (Soddu et al., 2009) and thus the patient was considered to be conscious. Of note is that 6 months later, when the patient was clinically re-examined, she recovered visual pursuit of a mirror (Vanhaudenhuyse et al., 2008), indicating her transition to a MCS. The residual brain activity detected via neuroimaging technologies could not be initially identified in the patient's bedside, suggesting that neuronal activation was taking place in absence of any behavioral output.

STRUCTURAL IMAGING

Structural connectivity refers to a network of physical or structural (axonal) connections which bines sets of neuronal populations. In clinical cases, information about the structural architecture of the brain can provide insights about recovery and neural plasticity in anoxic or traumatic brain injury. In chronic DOC, patients will progressively develop diffuse brain atrophy. In these cases, classical morphological MRI may not be a reliable indicator of the severity of the axonal injury and hence of the level of consciousness (Tshibanda et al., 2009). However, tools with higher sensitivity have been introduced which hold promise for studying plasticity in patients with DOC, such as diffusion tensor imaging (DTI) and magnetic resonance spectroscopy (MRS; Tshibanda et al., 2010).

Diffusion tensor imaging assesses the architectural organization of white matter fibers and hence can detect *in vivo* diffuse axonal injury (Arfanakis et al., 2002). In an exceptional case of late recovery from traumatic brain injury, Voss et al. (2006) used DTI to document an increased fractional anisotropy (thought to reflect fiber density) in large, bilateral regions of medial parieto-occipital areas of the white matter paralleling his clinical recovery of speech and motor function 19 years after the acute insult. These findings were contingent to an increased regional metabolism in these areas when measured with PET, similar to the partially restored cortical regions observed in patients who recover consciousness after being in a chronic VS/UWS (Laureys et al., 2000b, 2006a; **Figure 5**). This multimodal posteriomedial associative area has been previously suggested to be part of the human awareness network (Vogt and Laureys, 2005).

Magnetic resonance spectroscopy is another non-invasive technique that can provide *in vivo* quantification of certain biochemical markers such as N-acetylaspartate (heralding information about neuronal density and viability), choline (reflecting cell membrane turnover), and creatine (reflecting cell aerobic energy metabolism; Tshibanda et al., 2009). When information from this technique was combined with morphological MRI in traumatic brain injury, patients could be separated in prognostic subgroups based on the Glasgow Outcome Scale which was not possible when the different imaging techniques were applied separately (Carpentier et al., 2006).

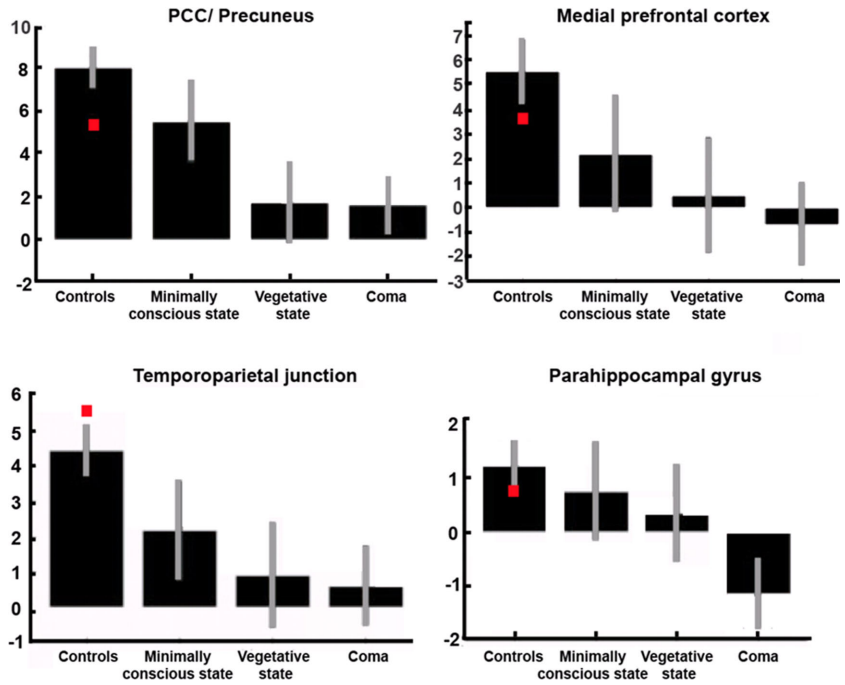


FIGURE 4 | The most representative nodes of the "default mode network" show a decrease in functional connectivity as we move from normal consciousness and locked-in syndrome (red squares) to minimally conscious or unconscious states. Graphs represent connectivity strength (mean *z* scores with 90% confidence intervals). PCC, posterior cingulate cortex. Adapted from Vanhaudenhuyse et al. (2009).

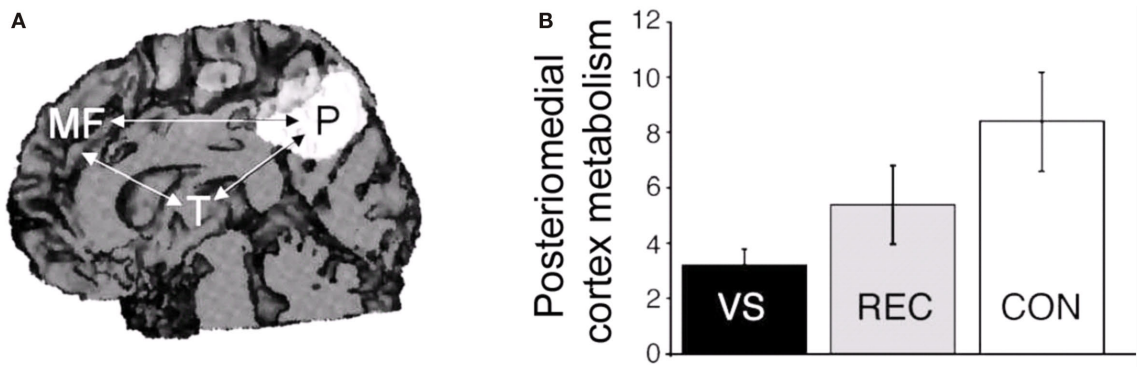


FIGURE 5 | (A) Restoration of glucose metabolism in posteriomedial cortices after recovery from the VS/UWS (area in white; arrows indicate the functional disconnections observed in VS/UWS patients). **(B)** In the same posteriomedial area, lower metabolism was observed in VS/UWS patients (black bar) as compared to those who recovered consciousness (gray bar) and to healthy controls (white bar), suggesting its critical role in the mediation of conscious awareness. Scale represents cerebral metabolic rates for glucose uptake (in mg/100 g/min). Adapted from Laureys et al. (2006a).

PERSPECTIVES

The use of resting state fMRI to study functional recovery and neural plasticity in DOC patients and its clinical routine use as a diagnostic and prognostic tool needs a controlled methodology and inclusion of a larger number of patients as is currently being tackled by multicentric collaborations. Nevertheless, clinicians should be aware of the many limitations and pitfalls intrinsic to "resting state" analyses – especially the challenge to disentangle genuine neural activity from artifactual movement-related fMRI signal in studies on severe brain injury (Soddu et al., in press). Although clinical assessment presently remains the gold standard in diagnosing this challenging population (Majerus et al., 2005), neuroimaging instruments in some exceptional cases of motor-deprived non-communicating DOC patients may be used as a means to establish a reliable communication code (Monti et al., 2010b). The challenge now will be to validate these novel technologies and to define the ethical and legal frameworks redefining cognitive competence in these patients with very limited and technology-dependent communication (Fins et al., 2008). Structural MRI coupled to spectroscopic and DTI techniques are currently being validated as prognostic markers in acute

and chronic DOC (e.g., Lescot et al., 2009; Tshibanda et al., 2010). These studies also will improve our understanding of residual neural plasticity in the recovery of consciousness.

CONCLUSION

Most recoveries of consciousness, with or without recovery of social or professional integration, take place within the first 3 months after non-traumatic and after 12 months after traumatic cerebral accidents and survival beyond 10 years remains unusual – albeit depending on the level of medical and nursing care (The Multi-Society Task Force on PVS, 1994; Monti et al., 2010a). However, clinical cases of both late spontaneous recoveries (e.g., Voss et al., 2006; Estraneo et al., 2010) or after invasive interventional treatments (e.g., Schiff et al., 2007) challenge the dogma of temporally fixed periods for possible neuronal plasticity. It is important to stress that the cellular mechanisms underlying recovery of consciousness after severe brain damage remain speculative. Our understanding of possible neurogenesis (known to occur predominantly in associative fronto-parietal cortices in non-human primates; Gould et al., 1999), axonal sprouting and neurite outgrowth, or even apoptosis in this patient population remains very limited. The residual cerebral plasticity during vegetative/unresponsive and MCS patients has been largely overlooked by the medical community and deserves further investigation. We believe that the challenge is now to identify the conditions in which and the mechanisms by which some patients may recover consciousness by use of the latest MRI and PET neuroimaging tools. The absence of large, controlled randomized interventional studies in patients with chronic DOC account for the present lack of evidence-based guidelines and tendency for therapeutic nihilism and can be related to the continuing societal, political, legal, and ethical debates in this field. The study of patients with chronic DOC may hence provide further insights in the medical management and rehabilitation of these patients at the clinical level, as well as increasing our understanding of the brain's long overlooked plastic abilities and the scientific quest for the neural correlates of human consciousness.

ACKNOWLEDGMENTS

This research was funded by the Belgian National Funds for Scientific Research (FNRS), the European Commission (DISCOS Marie-Curie Actions, COST, DECODER), the James McDonnell Foundation, the Mind Science Foundation, the French Speaking Community Concerted Research Action (ARC-06/11-340), the Fondation Médicale Reine Elisabeth, and the University and University Hospital of Liège.

REFERENCES

Adams, J. H., Graham, D. I., and Jennett, B. (2000). The neuropathology of the vegetative state after an acute brain insult. *Brain* 123(Pt 7), 1327–1338.

Arfanakis, K., Haughton, V. M., Carew, J. D., Rogers, B. P., Dempsey, R. J., and Meyerand, M. E. (2002). Diffusion tensor MR imaging in diffuse axonal injury. *Am. J. Neuroradiol.* 23, 794–802.

Baars, B. J., Ramsoy, T. Z., and Laureys, S. (2003). Brain, conscious experience and the observing self. *Trends Neurosci.* 26, 671–675.

Boly, M., Balteau, E., Schnakers, C., Degueldre, C., Moonen, G., Luxen, A., Phillips, C., Peigneux, P., Maquet, P., and Laureys, S. (2007). Baseline brain activity fluctuations predict somatosensory perception in humans. *Proc. Natl. Acad. Sci. U.S.A.* 104, 12187–12192.

Boly, M., Faymonville, M. E., Peigneux, P., Lambermont, B., Damas, P., Del Fiore, G., Degueldre, C., Franck, G., Luxen, A., Lamy, M., Moonen, G., Maquet, P., and Laureys, S. (2004). Auditory processing in severely brain injured patients: differences between the minimally conscious state and the persistent vegetative state. *Arch. Neurol.* 61, 233–238.

Boly, M., Faymonville, M. E., Schnakers, C., Peigneux, P., Lambermont, B., Phillips, C., Lancellotti, P., Luxen, A., Lamy, M., Moonen, G., Maquet, P., and Laureys, S. (2008a). Perception of pain in the minimally conscious state with PET activation: an observational study. *Lancet Neurol.* 7, 1013–1020.

Boly, M., Phillips, C., Tshibanda, L., Vanhaudenhuyse, A., Schabus, M., Dang-Vu, T. T., Moonen, G., Hustinx, R., Maquet, P., and Laureys, S. (2008b). Intrinsic brain activity in altered states of consciousness: how conscious is the default mode of brain function? *Ann. N.Y. Acad. Sci.* 1129, 119–129.

Boly, M., Tshibanda, L., Vanhaudenhuyse, A., Noirhomme, Q., Schnakers, C., Ledoux, D., Boveroux, P., Garweg, C., Lambermont, B., Phillips, C., Luxen, A., Moonen, G., Bassetti, C., Maquet, P., and Laureys, S. (2009). Functional connectivity in the default network during resting state is preserved in a vegetative but not in a brain dead patient. *Hum. Brain Mapp.* 30, 2393–2400.

Boveroux, P., Vanhaudenhuyse, A., Bruno, M. A., Noirhomme, Q., Lauwick, S., Luxen, A., Degueldre, C., Perlbarg, V., Plenevaux, A., Schnakers, C., Phillips, C., Brichant, J. F., Bohnomme, V., Maquet, P., Greicius, M. D., Laureys, S., and Boly, M. (2010). Breakdown of within- and between-network resting state functional magnetic resonance imaging connectivity during propofol-induced loss of consciousness. *Anesthesiology* 113, 1038–1053.

Bruno, M. A., Soddu, A., Demertzi, A., Laureys, S., Gosseries, O., Schnackers, C., Boly, M., Noirhomme, Q., Thonnard, M., Chatelle, C., and Vanhaudenhuyse, A. (2010). Disorders of consciousness: moving from passive to resting state and active paradigms. *Cogn. Neurosci.* 1, 193–203.

Buckner, R. L., Andrews-Hanna, J. R., and Schacter, D. L. (2008). The brain's default network: anatomy, function, and relevance to disease. *Ann. N.Y. Acad. Sci.* 1124, 1–38.

Carpentier, A., Galanaud, D., Puybasset, L., Muller, J. C., Lescot, T., Boch, A. L., Riedl, V., Cornu, P., Coriat, P., Dormont, D., and van Effenterre, R. (2006). Early morphologic and spectroscopic magnetic resonance in severe traumatic brain injuries can detect "invisible brain stem damage" and predict "vegetative states". *J. Neurotrauma* 23, 674–685.

Cohadon, F., and Richer, E. (1993). [Deep cerebral stimulation in patients with post-traumatic vegetative state. 25 cases]. *Neurochirurgie* 39, 281–292.

Coleman, M. R., Rodd, J. M., Davis, M. H., Johnsrude, I. S., Menon, D. K., Pickard, J. D., and Owen, A. M. (2007). Do vegetative patients retain aspects of language comprehension? Evidence from fMRI. *Brain* 130, 2494–2507.

Dehaene, S., Changeux, J. P., Naccache, L., Sackur, J., and Sergent, C. (2006). Conscious, preconscious, and subliminal processing: a testable taxonomy. *Trends Cogn. Sci.* 10, 204–211.

Demertzi, A., Soddu, A., Faymonville, M. E., Bahri, M. A., Gosseries, O., Vanhaudenhuyse, A., Phillips, C., Maquet, P., Noirhomme, Q., Luxen, A., and Laureys, S. (in press). Hypnotic modulation of resting state fMRI default mode and extrinsic network connectivity. *Prog. Brain Res.*

Demertzi, A., Vanhaudenhuyse, A., Bruno, M. A., Schnakers, C., Boly, M., Boveroux, P., Maquet, P., Moonen, G., and Laureys, S. (2008). Is there anybody in there? Detecting awareness in disorders of consciousness. *Expert Rev. Neurother.* 8, 1719–1730.

Di, H. B., Yu, S. M., Weng, X. C., Laureys, S., Yu, D., Li, J. Q., Qin, P. M., Zhu, Y. H., Zhang, S. Z., and Chen, Y. Z. (2007). Cerebral response to patient's own name in the vegetative and minimally conscious states. *Neurology* 68, 895–899.

Estraneo, A., Moretta, P., Loreto, V., Lanzillo, B., Santoro, L., and Trojano, L. (2010). Late recovery after traumatic, anoxic, or hemorrhagic long-lasting vegetative state. *Neurology* 75, 239–245.

Fins, J. J., Illes, J., Bernat, J. L., Hirsch, J., Laureys, S., and Murphy, E. (2008). Neuroimaging and disorders of consciousness: envisioning an ethical research agenda. *Am. J. Bioeth.* 8, 3–12.

Friston, K. (2002). Beyond phrenology: what can neuroimaging tell us about distributed circuitry? *Annu. Rev. Neurosci.* 25, 221–250.

Giacino, J. T., Ashwal, S., Childs, N., Cranford, R., Jennett, B., Katz, D. I., Kelly, J. P., Rosenberg, J. H., Whyte, J., Zafonte, R. D., and Zasler, N. D. (2002). The minimally conscious state: definition and diagnostic criteria. *Neurology* 58, 349–353.

Gould, E., Reeves, A. J., Graziano, M. S., and Gross, C. G. (1999). Neurogenesis in the neocortex of adult primates. *Science* 286, 548–552.

Jennett, B., and Plum, F. (1972). Persistent vegetative state after brain damage. A syndrome in search of a name. *Lancet* 1, 734–737.

Laureys, S. (2005). The neural correlate of (un)awareness: lessons from the vegetative state. *Trends Cogn. Sci.* 9, 556–559.

Laureys, S., Celesia, G. G., Cohadon, F., Lavrijsen, J., Leon-Carrion, J., Sannita, W. G., Sazbon, L., Schmutzhard, E., von Wild, K. R., Zeman, A., and Dolce, G. (2010). Unresponsive wakefulness syndrome: a new name for the vegetative state or apallic syndrome. *BMC Med.* 8, 68. doi: 10.1186/1741-7015-8-68

Laureys, S., Faymonville, M. E., Degueldre, C., Fiore, G. D., Damas, P., Lambermont, B., Janssens, N., Aerts, J., Franck, G., Luxen, A., Moonen, G., Lamy, M., and Maquet, P. (2000a). Auditory processing in the vegetative state. *Brain* 123, 1589–1601.

Laureys, S., Faymonville, M. E., Luxen, A., Lamy, M., Franck, G., and Maquet, P. (2000b). Restoration of thalamocortical connectivity after recovery from persistent vegetative state. *Lancet* 355, 1790–1791.

Laureys, S., Faymonville, M. E., Peigneux, P., Damas, P., Lambermont, B., Del Fiore, G., Degueldre, C., Aerts, J., Luxen, A., Franck, G., Lamy, M., Moonen, G., and Maquet, P. (2002). Cortical processing of noxious somatosensory stimuli in the persistent vegetative state. *Neuroimage* 17, 732–741.

Laureys, S., Boly, M., and Maquet, P. (2006a). Tracking the recovery of consciousness from coma. *J. Clin. Invest.* 116, 1823–1825.

Laureys, S., Giacino, J. T., Schiff, N. D., Schabus, M., and Owen, A. M. (2006b). How should functional imaging of patients with disorders of consciousness contribute to their clinical rehabilitation needs? *Curr. Opin. Neurol.* 19, 520–527.

Laureys, S., Goldman, S., Phillips, C., Van Bogaert, P., Aerts, J., Luxen, A., Franck, G., and Maquet, P. (1999). Impaired effective cortical connectivity in vegetative state: preliminary investigation using PET. *Neuroimage* 9, 377–382.

Laureys, S., Owen, A. M., and Schiff, N. D. (2004a). Brain function in coma, vegetative state, and related disorders. *Lancet Neurol.* 3, 537–546.

Laureys, S., Perrin, F., Faymonville, M. E., Schnakers, C., Boly, M., Bartsch, V., Majerus, S., Moonen, G., and Maquet, P. (2004b). Cerebral processing in the minimally conscious state. *Neurology* 63, 916–918.

Lescot, T., Galanaud, D., and Puybasset, L. (2009). Exploring altered consciousness states by magnetic resonance imaging in brain injury. *Ann. N.Y. Acad. Sci.* 1157, 71–80.

Lombardi, F., Taricco, M., De Tanti, A., Telaro, E., and Liberati, A. (2002). Sensory stimulation of brain-injured individuals in coma or vegetative state: results of a Cochrane systematic review. *Clin. Rehabil.* 16, 464–472.

Majerus, S., Gill-Thwaites, H., Andrews, K., and Laureys, S. (2005). Behavioral evaluation of consciousness in severe brain damage. *Prog. Brain Res.* 150, 397–413.

Mason, M. F., Norton, M. I., Van Horn, J. D., Wegner, D. M., Grafton, S. T., and Macrae, C. N. (2007). Wandering minds: the default network and stimulus-independent thought. *Science* 315, 393–395.

McGeown, W. J., Mazzoni, G., Venneri, A., and Kirsch, I. (2009). Hypnotic induction decreases anterior default mode activity. *Conscious Cogn.* 18, 848–555.

Monti, M. M., Laureys, S., and Owen, A. M. (2010a). The vegetative state. *BMJ* 341, c3765.

Monti, M. M., Vanhaudenhuyse, A., Coleman, M. R., Boly, M., Pickard, J. D., Tshibanda, L., Owen, A., and Laureys, S. (2010b). Willful modulation of brain activity in disorders of consciousness. *N. Engl. J. Med.* 362, 579–589.

Oh, H., and Seo, W. (2003). Sensory stimulation programme to improve recovery in comatose patients. *J. Clin. Nurs.* 12, 394–404.

Owen, A. M., Coleman, M. R., Boly, M., Davis, M. H., Laureys, S., and Pickard, J. D. (2006). Detecting awareness in the vegetative state. *Science* 313, 1402.

Passler, M. A., and Riggs, R. V. (2001). Positive outcomes in traumatic brain injury-vegetative state: patients treated with bromocriptine. *Arch. Phys. Med. Rehabil.* 82, 311–315.

Pistoia, F., Mura, E., Govoni, S., Fini, M., and Sara, M. (2010). Awakenings and awareness recovery in disorders of consciousness: is there a role for drugs? *CNS Drugs* 24, 625–638.

Sawyer, E., Mauro, L. S., and Ohlinger, M. J. (2008). Amantadine enhancement of arousal and cognition after traumatic brain injury. *Ann. Pharmacother.* 42, 247–252.

Schiff, N. D. (2010). Recovery of consciousness after brain injury: a mesocircuit hypothesis. *Trends Neurosci.* 33, 1–9.

Schiff, N. D., and Fins, J. J. (2007). Deep brain stimulation and cognition: moving from animal to patient. *Curr. Opin. Neurol.* 20, 638–642.

Schiff, N. D., Giacino, J. T., Kalmar, K., Victor, J. D., Baker, K., Gerber, M., Fritz, B., Eisenberg, B., O'Connor, J., Kobylarz, E. J., Farris, S., Machado, A., McCagg, C., Plum, F., Fins, J. J., and Rezai, A. R. (2007). Behavioural improvements with thalamic stimulation after severe traumatic brain injury. *Nature* 448, 600–603.

Schiff, N. D., Ribary, U., Moreno, D. R., Beattie, B., Kronberg, E., Blasberg, R., Giacino, J. T., McCagg, C., Fins, J. J., Llinas, R., and Plum, F. (2002). Residual cerebral activity and behavioural fragments can remain in the persistently vegetative brain. *Brain* 125, 1210–1234.

Schnakers, C., Hustinx, R., Vandewalle, G., Majerus, S., Moonen, G., Boly, M., Vanhaudenhuyse, A., and Laureys, S. (2008). Measuring the effect of amantadine in chronic anoxic minimally conscious state. *J. Neurol. Neurosurg. Psychiatry* 79, 225–227.

Shiel, A., Burn, J. P., Henry, D., Clark, J., Wilson, B. A., Burnett, M. E., and McLellan, D. L. (2001). The effects of increased rehabilitation therapy after brain injury: results of a prospective controlled trial. *Clin. Rehabil.* 15, 501–514.

Soddu, A., Boly, M., Nir, Y., Noirhomme, Q., Vanhaudenhuyse, A., Demertzi, A., Arzi, A., Ovadia, S., Stanziano, M., Papa, M., Laureys, S., and Malach, R. (2009). Reaching across the abyss: recent advances in functional magnetic resonance imaging and their potential relevance to disorders of consciousness. *Prog. Brain Res.* 177, 261–274.

Soddu, A., Vanhaudenhuyse, A., Bahri, M., A., Bruno, M. A., Boly, M., Demertzi, A., Tshibanda, J., Phillips, C., Stanziano, M., Ovadia-Caro, S., Nir, Y., Maquet, P., Papa, M., Malach, R., Laureys, S., and Noirhomme, Q. (in press). Identifying the default mode component in spatial ICA of patients with disorders of consciousness. *Hum. Brain Mapp.*

Taira, T., and Hori, T. (2007). Intrathecal baclofen in the treatment of post-stroke central pain, dystonia, and persistent vegetative state. *Acta Neurochir. Suppl.* 97, 227–229.

The Multi-Society Task Force on PVS. (1994). Medical aspects of the persistent vegetative state (2). *N. Engl. J. Med.* 330, 1572–1579.

Tolle, P., and Reimer, M. (2003). Do we need stimulation programs as a part of nursing care for patients in "persistent vegetative state"? A conceptual analysis. *Axone* 25, 20–26.

Tononi, G., and Laureys, S. (2009). "The neurology of consciousness: an overview," in *The Neurology of Consciousness: Cognitive Neuroscience and Neuropathology,* eds S. Laureys and G. Tononi (Oxford, UK: Academic Press), 375–412.

Tshibanda, L., Vanhaudenhuyse, A., Boly, M., Soddu, A., Bruno, M. A., Moonen, G., Laureys, S., and Noirhomme, Q. (2010). Neuroimaging after coma. *Neuroradiology* 52, 15–24.

Tshibanda, L., Vanhaudenhuyse, A., Galanaud, D., Boly, M., Laureys, S., and Puybasset, L. (2009). Magnetic resonance spectroscopy and diffusion tensor imaging in coma survivors: promises and pitfalls. *Prog. Brain Res.* 177, 215–229.

Vanhaudenhuyse, A., Demertzi, A., Schabus, M., Noirhomme, Q., Bredart, S., Boly, M., Phillips, C., Soddu, A., Luxen, A., Moonen, G., and Laureys, S. (2011). Two distinct neuronal networks mediate the awareness of environment and of self. *J. Cogn. Neurosci.* 23, 570–578.

Vanhaudenhuyse, A., Noirhomme, Q., Tshibanda, L. J., Bruno, M. A., Boveroux, P., Schnakers, C., Soddu, A., Perlbarg, V., Ledoux, D., Brichant, J. F., Moonen, G., Maquet, P., Greicius, M. D., Laureys, S., and Boly, M. (2009). Default network connectivity reflects the level of consciousness in non-communicative brain-damaged patients. *Brain* 133, 161–171.

Vanhaudenhuyse, A., Schnakers, C., Bredart, S., and Laureys, S. (2008). Assessment of visual pursuit in post-comatose states: use a mirror. *J. Neurol. Neurosurg. Psychiatry* 79, 223.

Vogt, B. A., and Laureys, S. (2005). Posterior cingulate, precuneal and retrosplenial cortices: cytology and components of the neural network correlates of consciousness. *Prog. Brain Res.* 150, 205–217.

Voss, H. U., Uluc, A. M., Dyke, J. P., Watts, R., Kobylarz, E. J., McCandliss, B. D., Heier, L. A., Beattie, B. J., Hamacher, K. A., Vallabhajosula, S., Goldsmith, S. J., Ballon, D., Giacino, J. T., and Schiff, N. D. (2006). Possible axonal regrowth in late recovery from the minimally conscious state. *J. Clin. Invest.* 116, 2005–2011.

Whyte, J., Katz, D., Long, D., DiPasquale, M. C., Polansky, M., Kalmar, K., Giacino, J. T., Childs, N., Mercer, W., Novak, P., Maurer, P., and Eifert, B. (2005). Predictors of outcome in prolonged posttraumatic disorders of consciousness and assessment of medication effects: a multicenter study. *Arch. Phys. Med. Rehabil.* 86, 453–462.

Yamamoto, T., Katayama, Y., Oshima, H., Fukaya, C., Kawamata, T., and Tsubokawa, T. (2001). Deep brain stimulation therapy for a persistent vegetative state. *Acta Neurochir. Suppl.* 79, 79–82.

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

Received: 16 October 2010; paper pending published: 19 November 2010; accepted: 24 December 2010; published online: 07 February 2011.

Citation: Demertzi A, Schnakers C, Soddu A, Bruno M-A, Gosseries O, Vanhaudenhuyse A and Laureys S (2010) Neural plasticity lessons from disorders of consciousness. Front. Psychology ***1****:245. doi: 10.3389/fpsyg.2010.00245*

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

frontiers in
PSYCHOLOGY

HYPOTHESIS AND THEORY ARTICLE
published: 06 December 2011
doi: 10.3389/fpsyg.2011.00337

Sea slugs, subliminal pictures, and vegetative state patients: boundaries of consciousness in classical conditioning

Tristan A. Bekinschtein[1]*, Moos Peeters[1], Diego Shalom[2] and Mariano Sigman[2]

[1] *Cognition and Brain Sciences Unit, Medical Research Council, Cambridge, UK*
[2] *Integrative Neuroscience Lab, Physics Department, University of Buenos Aires, Buenos Aires, Argentina*

Edited by:
Mads Jensen, Aarhus University Hospital, Denmark

Reviewed by:
Anil K. Seth, University of Sussex, UK
Melanie Wilke, California Institute of Technology, USA
Giorgio Marchetti, www.mind-consciousness-language.com research net, Italy

****Correspondence:***
Tristan A. Bekinschtein, Cognition and Brain Sciences Unit, Medical Research Council, 15 Chaucer Road, CB2 7EF Cambridge, UK.
e-mail: tristan.bekinschtein@mrc-cbu.cam.ac.uk

Classical (trace) conditioning is a specific variant of associative learning in which a neutral stimulus leads to the subsequent prediction of an emotionally charged or noxious stimulus after a temporal gap. When conditioning is concurrent with a distraction task, only participants who can report the relationship (the contingency) between stimuli explicitly show associative learning. This suggests that consciousness is a prerequisite for trace conditioning. We review and question three main controversies concerning this view. Firstly, virtually all animals, even invertebrate sea slugs, show this type of learning; secondly, unconsciously perceived stimuli may elicit trace conditioning; and thirdly, some vegetative state patients show trace learning. We discuss and analyze these seemingly contradictory arguments to find the theoretical boundaries of consciousness in classical conditioning. We conclude that trace conditioning remains one of the best measures to test conscious processing in the absence of explicit reports.

Keywords: *Aplysia*, subliminal, vegetative state, consciousness, learning, trace conditioning

INTRODUCTION

Learning comes in many forms. Some forms, such as declarative memory, are explicit memories of objects, places, or events. Others forms, such as non-declarative memory, are implicit, like habits, skills, or priming, and do not require conscious awareness (Gilbert et al., 2001; Squire, 2004). While this has been an established taxonomy of learning, the boundaries are often vague and difficult to demarcate. For instance, in associative learning there is a fine-grained distinction between declarative associative learning, which is dependent on the conscious association of events, and non-declarative unconscious associative learning, which occurs without awareness of the link between the related meaningful events (Shanks, 1995).

In an associative learning task, when two stimuli are systematically presented in a temporal sequence, a new relationship between these two items is learned. In classical trace learning, a neutral conditioning stimulus (CS) precedes and, therefore, causes the subsequent prediction of an emotionally charged or noxious stimulus (unconditioned stimulus, US; **Figure 1**). During a differential eye-blink trace conditioning task, a specific tone (CS+) warns of a puff of air to the eye, whereas another tone (CS−) does not. The presence of an anticipatory eye-blink response, which is a conditioned response (CR) to the CS+ tone, is highly correlated with participants' verbal report of the relationship between the stimuli presented; this is known as the contingency (Dawson and Reardon, 1973; Clark and Squire, 1998). In contrast, during a delay conditioning task, for which there is no gap between the neutral stimulus (CS) and the puff of air to the eye, blink responses (CR) are elicited despite the lack of awareness of the contingency (Clark and Squire, 1998, 1999; **Figure 1**). Furthermore, trace, but not delay, conditioning is strongly influenced by an expectancy of the US (Clark et al., 2001).

Consequently, delay conditioning has been considered a hallmark of non-declarative learning and is systematically used as an associative learning task in vertebrates and invertebrates (Lavond et al., 1993). In contrast, human trace learning is dependent on conscious awareness of the contingency between stimuli (Christian and Thompson, 2003). Moreover it is this fact that has made human trace learning a potential "*Turing-test*" of consciousness (Koch, 2004).

Though this "turing-test" potential has been influential in behavioral neuroscience, three arguments could pose a serious challenge to this notion: (1) trace conditioning can be learnt by almost every animal, even invertebrate sea slugs (Glanzman, 1995); (2) trace conditioning can be elicited using subliminal stimuli (Esteves et al., 1994); and (3) clinically defined unconscious patients might learn trace conditioning (Bekinschtein et al., 2009). It is in addressing these three arguments that we will characterize the theoretical boundaries of consciousness of learning.

IS CONDITIONING DIFFERENT BETWEEN SEA SLUGS AND HUMANS?

The evolution of wings can be used as an analogy for the evolution of conditioning. During the evolution of animals on this planet, wings appeared at least three times from three different ancestors. The first appearance was probably 390–320 million years ago. Primitive wingless insects known as Bristletails used long antennae-like filaments at the ends of their bodies to glide down to tree trunks from forest canopies (Yanoviak et al., 2009). The second appearance, around 150 million years ago, involved the

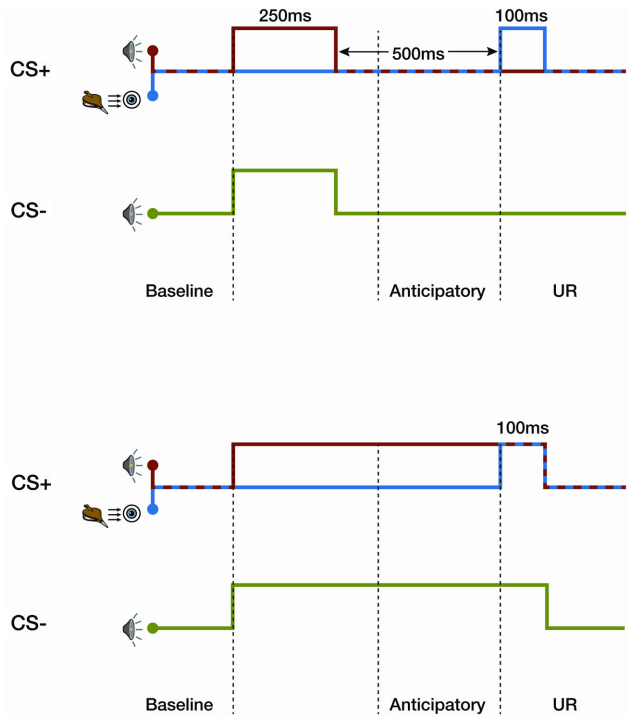


FIGURE 1 | Differential trace (upper graph) and delay conditioning (lower graph) for the eye-blink task. In trace conditioning there is a silent period between a tone and a puff of air to the eye. The CS+ tone is presented for 250 ms, followed by a 500-ms silent gap, before a 100-ms puff of air to the eye (aversive stimulus). A second neutral stimulus tone, CS−, is presented alone. In delay conditioning, the tone and puff of air co-occur. The tone lasts for 850 ms, covering the silent gap and co-terminating with the puff of air.

development of wings on dinosaurs. Some dinosaurs started to evolve lighter skeletons; their wrists changed and feathers grew to form wings (Lewin, 1983; Sullivan et al., 2010). Though dinosaurs evolved wings, which gave birth to modern birds their capacity to fly has a different origin and ancestry than flying insects. The third appearance of wings developed in bats (Simmons et al., 2008). This is yet another origin of this feature. Though the three different wings serve a similar function, that is to fly, they were not derived from a common ancestor. Moreover, they are not controlled by the same machinery, and they do not obey the same rules.

Is the same true of conditioning? Has a conditioning mechanism appeared only once in evolution? While the well preserved molecular machinery may suggest this, findings at the systems level are disparate and may suggest a different conclusion (Barco et al., 2006). Molecular mechanisms underlying acquisition and consolidation of memory are in fact preserved in most species. However, learning relies on different network mechanisms in humans and sea slugs (Takehara et al., 2003). The plasticity of arrays of neurons and the ubiquitous evolutionary pressure of associative learning makes it extremely difficult to disentangle whether conditioning emerged only once during the course of evolution. It could be that conditioning emerged several times from several ancestral sources just as wings did. We are only left with the certainty that associative learning can be instantiated in simply a few neurons, as in sea slugs, or in millions of neurons, as in mammals.

Humans show the highest degree of behavioral flexibility among animals, probably through a flexible network that relies on a frontoparietal hub. This flexibility, however, comes at a cost that results in slow, limited computational capacity. Trace conditioning does not seem to escape from this rule (Zylberberg et al., 2010). Arbitrary associations in humans are instantiated in this routing system that appears to be intrinsically related to consciousness (Zylberberg et al., 2011). On the contrary, in the sea slug these temporary associations may rely on direct connections (Grol et al., 2006). This may also happen in humans with highly compatible sensory–motor relations which bypass the central routing system (Grol et al., 2006). Trace conditioning in sea slugs, where consciousness is not required, does not necessarily imply that consciousness does not play a role in human trace conditioning. Alone, it does not imply more than, for example, a bacteria in need of oxygen breathing under water makes implications about human respiration.

Trace conditioning it is also instantiated in neuroanatomically different systems. The sea slug *Aplysia Californica* only needs a few neurons to perform trace conditioning, while rabbits require hippocampal, frontal cortex, and cerebellar networks (Christian and Thompson, 2003). In contrast, rabbits seem to be more flexible than gastropods in their learning abilities, but not more than insects, i.e., fruit flies, (Heisenberg et al., 2001). In humans, as in rodents, trace conditioning is dependent of the hippocampus (James et al., 1987; Moyer et al., 1990; Clark and Squire, 1998) while delay conditioning can be elicited without hippocampi (Ivkovich and Stanton, 2001; Woodruff-Pak and Disterhoft, 2008). Delay conditioning seems to rely primarily on a functional cerebellum (McCormick and Thompson, 1984; Mauk and Thompson, 1987; Gerwig et al., 2007). Functional imaging studies have shown that the hippocampus is activated by both trace and delay conditioning, but it is significantly more activated by trace (Cheng et al., 2008). These studies of neuroanatomy and the function of associative conditioning may indicate declarative memory's high dependence on the hippocampi and neocortex, as opposed to non-declarative memory's need for these structures, which is minimal (Eichenbaum and Cohen, 2000).

Further evidence of a causal link between awareness of the contingency and trace conditioning comes from human and non-primate mammalian data. In humans the variability in trace conditioning responses and learning seems to be linked to attention (Lovibond and Shanks, 2002; Carter et al., 2003), suggesting a strong modulation by central cognitive processes. In healthy volunteers, increasing attentional load parametrically modulates the degree of trace learning (Carter et al., 2003). This elegant experiment used the classic *n*-back task to engage attentional resources and working memory. The engagement of these dramatically decreases the anticipatory responses to the US in trace conditioning. The effect was less prominent for delay conditioning. Interestingly, the interference paradigm used in humans to decrease conditioning has been replicated using mice (Han et al., 2003), demonstrating a similar functional frontotemporal network supporting successful trace conditioning in a non-primate

mammalian species. In this study mice heard a tone and received a foot shock immediately after (delay conditioning) or, alternatively, the mice received a silent gap between tone and shock (trace conditioning). Two additional groups of mice experienced two lights flashing as interference for the tone–shock conditions. Learning was impaired by light interference in trace but not in delay conditioning. In mice the light flashing acted as "cognitive" interference, just as working memory load might cause interference in trace conditioning with humans. In both populations, i.e., in humans and mice, interference effects are found in trace but not delay conditioning. The fact that attention is a key component of conscious learning and also that trace conditioning is affected by awareness may indicate that trace conditioning is a type of conscious learning.

IS TRACE CONDITIONING POSSIBLE WHEN STIMULI ARE UNCONSCIOUSLY PERCEIVED?

There is general agreement that exposure to a contingency between conditioned stimulus (CS) and US will create an associative process called conditioning (Dickinson, 1980; Rescorla, 1988). The difference between human and non-human animal models is that the former can easily produce a verbal or motor report of the relationship between the CS and US. If humans can form an internal representation of the contingency and verbalize it or make a voluntary response, then this behavior is taken as evidence of conscious awareness (Lovibond and Shanks, 2002).

What if the CS is masked, or more generally, not accessible to verbal reports? Can a verbal report of the contingency still be present? This question is central to current theoretical discussions of the role of conscious processing in trace conditioning. Is conscious awareness of the stimuli needed in order for trace conditioning to occur? There are primarily two models that account for awareness in trace conditioning. The single-process model, asserts that a sole propositional learning process mediates expression CR and the expectancy of US (full network mapping; Lovibond and Shanks, 2002). The dual-process model, however, claims that these behavioral responses, both CR expression and US expectancy, are expressions of two independent learning processes (partial network mapping; Perruchet, 1985; Morris et al., 1999; Perruchet et al., 2006).

There is empirical evidence championing each model. Some studies find evidence in support of the single-process model (Daum et al., 1991, 1992; Manns et al., 2000a,b; Weike et al., 2007). While the series of experiments performed by Perruchet, Destrebecqz, Cleeremans, and colleagues (Destrebecqz and Cleeremans, 2001; Destrebecqz et al., 2005, 2010; Perruchet et al., 2006) and experiments performed by others (Ohman and Soares, 1993, 1994, 1998; Ohman et al., 1995; Weidemann et al., 2009) strongly support a two-way learning process.

The discussion is centered on the verbal reports of awareness of the contingencies (in Questionnaires, motor evaluation or subjective ratings) and the measures of CR. There is either complete agreement or disagreement between the two. This discrepancy raises two relevant methodological aspects about how to measure the awareness. Firstly, post-training questionnaires may elicit metacognitive process. Therefore the participant may verbally report the contingencies, not because they noticed them during the learning phase, but because they were forced to think about the contingencies after conditioning had finished. In this case it is difficult to know when conscious learning of the relationship between the stimuli occurred. It could have been during learning or it could have been when that participant was prompted about the relationship between stimuli. Secondly, if awareness is measured online during the experiment, as to avoid a *post hoc* metacognitive process, then participants may then take note of the contingencies of which they were previously unaware. This may lead to previously unaware participants becoming explicitly aware of the contingencies. The measurement of both verbal reports of contingencies and the CR is critical.

In a series of studies, Ohman et al. (1995) and Ohman and Soares (1993, 1994, 1998) argue that conditioning of electrodermal responses to electric shocks can occur with masked, unconsciously perceived, stimuli. It has been found that unconsciously perceived stimuli only elicited a CR when fear-relevant stimuli, such as spiders and snakes (Ohman and Soares, 1993) or angry faces (Esteves et al., 1994), were presented as CS+. In support of Ohman's claims, one recent study showed that sensitivity to masking conditions was related to the CR of a masked CS but not an unmasked CS (Cornwell et al., 2007). In this study, sensitivity to the masked condition was a marker of unconscious processing, i.e., if participants were not aware of the masked item then this was taken to indicate that items were processed at an unconscious level. The depth of unconscious processing of the CS was linked to the intensity of the CR. Weak perception of the CS through masking may not elicit conscious recognition of it, but the CS may still be above an identification threshold. However, there are criticisms of these findings.

One criticism is that the measure used to assess perceptual awareness of the CS may not be sufficiently sensitive to identify participants with residual awareness of stimulus features (Pessoa, 2005; Graziano and Sigman, 2009). Additionally, some fear relevance effects in backward masking conditioning, as observed by Ohman, could be due to selective sensitization rather than unconscious associative processes. Some methodological concerns have also been raised concerning the extent to which participants were truly unaware of the stimuli (Lovibond and Shanks, 2002). Ohman replied to these criticisms thoroughly, using two main arguments. Firstly, not all verbal discriminative responses indicate awareness; discrimination of stimuli above chance levels does not necessarily imply conscious awareness (Merikle and Daneman, 2000; Wiens and Ohman, 2002). Secondly, criticisms have assumed that awareness is a conscious experience. Therefore a measure of awareness must involve a measure of a subjective state.

Another highly debated experiment was performed by Núñez and de Vicente (2004), they have also showed that CR can be elicited when masked words are paired with a mild shock, and that this response is, as it was in the studies by Ohman, related to the participants' detection threshold. However, the results of this study are somewhat difficult to interpret as a higher proportion of participants in the unconscious masked condition produced a CR than in the conscious group. This study evaluated masked words paired with electric shocks; they used either a detection threshold, i.e., *Was the stimulus a word or a blank?,* or an identification threshold, i.e., *Was it word1, word2 or not a*

word? When participants failed to detect a stimulus in the tachitoscope, half exhibited a CR above learning criterion (four out of eight). Yet, when participants were above detection threshold ("conscious") only 11% (2 of 18) showed learning. There was a higher instance of conditioning, using a detection threshold, when stimuli were are presented unconsciously and when CR was measured using autonomic nervous system signals like skin conductance. On the contrary, when an identification threshold was applied, i.e., subjects had to differentiate between two words or two non-words, only 10% (1 of 10 participants) of participants in the unconscious condition exhibited CR, but 58% (7 out of 12) of participants in the conscious group exhibited CR. These contradictory results point to two different learning systems, the unconscious system, which bypasses central processes and consciousness-related workspaces. It also possibly directly links the early visual system with the autonomic nervous system and the conscious associative system, where the activation of the frontoparietal cortices may influence the autonomous nervous system giving rise to a different signal and a different type of learning. It is indeed the case that the variability, speed, and regularity of the CR was higher in the conscious identification group as compared to the unconscious detection learners (Núñez and de Vicente, 2004).

There are, however, caveats to this study's design that must be considered when discussing the conclusions. Firstly, in the identification threshold condition, words were repeated and could have therefore been deprived of their meaning. When a word is repeated it becomes easy to rely on low level features, such as the letter array, to determine the word's identity without the need to access its meaning. The participants' decisions in this condition may not have been based on anything more than a surface features. Another methodological problem of this study is that the variability of the perceptual threshold was high and several subjects were therefore excluded leading to a small sample size. The final subgroups of learners were 2/18 and 4/8 for conscious and unconscious detection of a word/blank, respectively, and 7/12 and 1/10 subjects for conscious and unconscious identification of the word, respectively. The low number of participants complicates the statistical analysis and make conclusions difficult to extrapolate to other cases. Thirdly, there is another possible explanation of "unconscious conditioning". When an electric shock is used as the aversive stimulus, as opposed to a puff of air to the eye or a loud tone, it may induce a general increase in arousal. This may, change detection thresholds and act as a confounding factor, leading to difficulty in interpreting the results.

In short, there seems to be consistent evidence showing that trace conditioning can also be elicited by unconscious stimuli with a strong emotional content, which is not accessible for verbal report. This is reminiscent of the sea slug. There may be some specific residual forms of trace learning that can be mediated by the most likely candidates: unconscious processing and emotional stimuli. However, evidence of subliminal abstract, non-emotional stimuli eliciting trace conditioning is not conclusive. Careful and well-designed experiments are needed to robustly deliminate the boundaries of consciousness thresholds of stimuli in a given task. Their ability to elicit trace learning above and below this threshold should certainly be included in the agenda of highly relevant experiments in the next years.

DO UNCONSCIOUS PATIENTS SHOW CONDITIONING?

If trace conditioning is taken as an indirect test of awareness, then a clinically defined unconscious patient, i.e., those in a vegetative state (VS), that shows CRs (learning) should be considered conscious. Conversely, if a clinically defined unconscious patient shows anticipatory responses in trace conditioning, then this learning may not necessarily be indicative of conscious awareness (Bekinschtein et al., 2009). VS patients are considered to be unconscious. They lack any behavior consistent with conscious awareness, such as, they do not follow someone with their eyes or head across the room, and they do not gesticulate or react to signs, words or commands. However, we used a classic differential trace conditioning eye-blink paradigm in 13 VS patients to investigate whether these unconscious patients might show learning (Bekinschtein et al., 2009). Moreover, we found a subset of these VS patients did show learning (see **Figure 2**). Two VS could appropriately produce CR to CS+ and not to CS−, and another four patients showed non-discriminatory anticipatory responses, i.e., producing the CR to both stimuli types. Patients that demonstrated learning eventually recovered by regaining awareness, as opposed to those patients that did not show learning. It is our belief that the patients that showed learning were partially conscious at the time of testing, but were unable to produce overt voluntary responses. In a recent trace conditioning study, using aversive noises and pleasant fanfares measuring skin conductance response, Scott et al. (2011), showed that only participants attending to the stimuli and able to catch the rule of the experiment showed CRs. If this paradigm is applied in the same disorders of consciousness patients group as the trace eye-blink conditioning, it may show convergence in conscious-dependent learning, and helping to better define the learning–consciousness relationship.

It is crucial to consider alternate hypotheses related to trace conditioning and the limited capacity of consciousness. Limited capacity implies that to learn trace conditioning, progressive gain of awareness of the contingency is required by maintaining a free global workspace clear of other contents (Dehaene and Naccache, 2001). The use of this limited space for other events or processes may weaken the link that establishes the association between the CS and the US. This in turn would prevent the associative relationship from being established.

In our original study the control group for the unconscious patients were anesthetized participants. These sedated healthy participants showed no learning of the contingency, or CR, for eye-blink trace conditioning. The sedative (propofol) that anesthetized these participants ensured that they were indeed unconscious, and the levels of drug caused a significant decrease in absolute cerebral blood flow. In particular the propofol-related variations in the thalamic blood flow appeared to be linked to the midbrain reticular formation, thus suggests a close functional relationship between the two brain structures while unconscious (Fiset et al., 1999). These deeply sedated participants showed true unconsciousness through both diminished awareness and arousal (wakefulness). Moreover, the participants under low doses of propofol had a

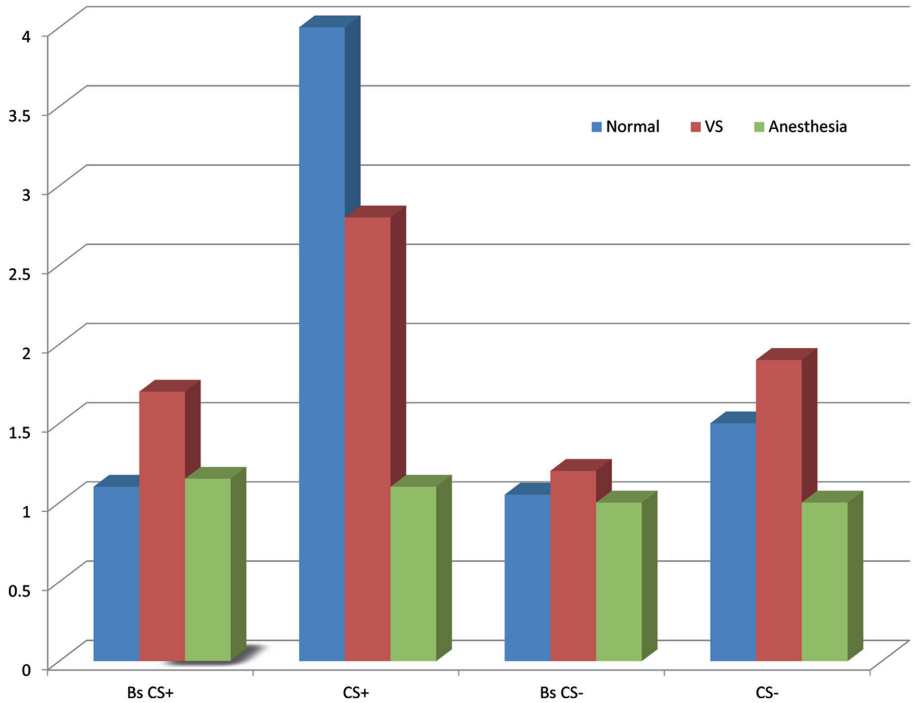


FIGURE 2 | Anticipatory learning in normal and anesthetized participants, and in vegetative state patients during trace differential conditioning. Bars show mean muscle activity for baseline of the CS+(Bs CS+), CS+ and baseline of the CS−(Bs CS−), CS−. Learning is in arbitrary units. Sedated subjects show no learning; vegetative state patients, as a group, show less muscle activity between CS+ and its baseline, and to CS−.

completely disorganized system for global integrated processing (Davis et al., 2007; Stamatakis et al., 2010). On the contrary, unconscious VS patients show dissociated wakefulness and awareness; they are regarded as being awake but not aware (Jennett and Plum, 1972). It could be that some VS patients retain partial functionality of the networks that support the acquisition of trace conditioning, but this is not enough to produce volitional movements.

CONCLUSION

Sea slugs can learn trace conditioning, but they do not show a form of learning that reveals the flexibility typically displayed in conscious forms of learning (Van den Bussche et al., 2008; Heinemann et al., 2009; see Dehaene and Changeux, 2011 for a review). Instead, these mollusks use the minimal numbers of systems necessary for successful associative learning. As far as we know sea slugs are not conscious, that is they are not conscious in the same way as humans are. Hence the mere observation of trace conditioning learning in this species does not provide evidence that trace conditioning is not an adequate signature of consciousness.

Quite the opposite is true; humans are extremely sensitive to context and are continuously interpreting all incoming stimuli in multiple ways. This capacity for over-interpretation may help in social interaction, goal-oriented behavior and possibly changes the way that trace conditioning is encoded and processed. The expression of trace learning in humans through a display of non-stereotyped representations of the contingency relies on an entirely distinct neural architecture. It also seems to rely on overt report in addition to CR.

Several studies have used masked or subliminally presented stimuli in order to make CS impossible to report (unconsciously perceived). This type of design has been used to explicitly determine whether conscious perception of the stimuli is necessary to achieve trace learning. Unfortunately, the results are inconclusive about whether trace learning, generally, is achievable through subliminal stimuli or under some specific circumstances. A robust conclusion concerning these studies involves the CS. When it is not neutral, but of negative valence, the masked stimuli seems to be processed up to the point of forming an association with the US, despite not being reported, detected, or discriminated.

The fact that clinically defined unconscious patients show trace learning suggest that they may have partial capacity for conscious awareness, as trace learning is dependant on some form of conscious awareness of the contingency. This result is further strengthened by data from sedated participants that show drug induced unconsciousness does not produce anticipatory responses above the baseline (Bekinschtein et al., 2009).

Trace conditioning remains very much linked to awareness of the contingency between CS and US. We believe it is a combination of the timing between stimuli and the variations on the CS that will allow us to better frame, over the next few years, the boundaries between this basic form of learning and conscious awareness.

The aforementioned analyses leads us to propose here three conditions which should be met in order for trace conditioning to be used as a test for conscious awareness: (1) a relatively stable, sustained attention to the stimuli, (2) a low central processing load to avoid interference and (3) a well defined stimuli, perceptually

discernible and close in presentation time. If (1) is not met, then attention will deviate from the stimuli, and the creation of the association between CS and US will be disrupted. This would lead to sparse and inconsistent demonstration of the CR (Armony and Dolan, 2002). If (2) is not met then working memory capacity will be saturated by another task. Learning will then decrease, and as a result there will be low awareness of the contingencies. In this respect, trace conditioning seems to require the central resources of the attentional system in order for the association to be established, and this seems to be paired with awareness (Carter et al., 2003). If (3) is not met, then the stimuli will be ambiguous and the trace that is established will not be long-lasting or robust. This will lead to partial learning only and high variance in CR (Lovibond and Shanks, 2002; Sehlmeyer et al., 2009).

We would like to raise one final important point that has not yet been emphasized. It is of paramount importance if trace conditioning is to be used in practical terms. Clark and Squire (1998) established that in order to observe an anticipatory eye-blink response participants must be aware of the contingency at a level that is sufficient for a verbal report. The contrary, however, is not true (both in the original Clark and Squire data and in our data making more than 200 participants combined): Some participants show verbal reports of the contingency, but do not show the CR of an anticipatory associative response. This simply may be because participants react in different ways to the anticipation of the air puff when they are aware that it is imminent. The most frequent reaction is a contraction of the muscle prior to the air puff, however, some participants may control this spontaneous response by relaxing and, inhibiting their blinks or by simply doing nothing. Hence, an improved statement is needed. While the absence of learning does not provide information concerning the participants' degree of awareness; the contrary is true. The presence of trace learning may be taken as evidence in favor of conscious awareness.

ACKNOWLEDGMENTS

Tristan A. Bekinschtein is supported by a Wellcome Trust Biomedical Research Career Development Fellowship. We thank Simon Strangeways for his help with the figures and Elisa Pye for critically reading the manuscript.

REFERENCES

Armony, J. L., and Dolan, R. J. (2002). Modulation of spatial attention by fear-conditioned stimuli: an event-related fMRI study. *Neuropsychologia* 40, 817–826.

Barco, A., Bailey, C. H., and Kandel, E. R. (2006). Common molecular mechanisms in explicit and implicit memory. *J. Neurochem.* 97, 1520–1533.

Bekinschtein, T. A., Shalom, D. R., Forcato, C., Herrera, M., Coleman, M. R., Manes, F. F., and Sigman, M. (2009). Classical conditioning in the vegetative and minimally conscious state. *Nat. Neurosci.* 12, 1343–1351.

Carter, R. M., Hofstotter, C., Tsuchiya, N., and Koch, C. (2003). Working memory and fear conditioning. *Proc. Natl. Acad. Sci.* 100, 1399–1404.

Cheng, D. T., Disterhoft, J. F., Power, J. M., Ellis, D. A., and Desmond, J. E. (2008). Neural substrates underlying human delay and trace eyeblink conditioning. *Proc. Natl. Acad. Sci. U.S.A.* 105, 8108–8113.

Christian, K. M., and Thompson, R. F. (2003). Neural substrates of eyeblink conditioning: acquisition and retention. *Learn. Mem.* 10, 427–455.

Clark, R. E., Manns, J. R., and Squire, L. R. (2001). Trace and delay eyeblink conditioning: contrasting phenomena of declarative and nondeclarative memory. *Psychol. Sci.* 12, 304–308.

Clark, R. E., and Squire, L. R. (1998). Classical conditioning and brain systems: the role of awareness. *Science* 280, 77–81.

Clark, R. E., and Squire, L. R. (1999). Human eyeblink classical conditioning: effects of manipulating awareness of the stimulus contingencies. *Psychol. Sci.* 10, 14–18.

Cornwell, B. R., Echiverri, A. M., and Grillon, C. (2007). Sensitivity to masked conditioned stimuli predicts conditioned response magnitude under masked conditions. *Psychophysiology* 44, 403–406.

Daum, I., Channon, S., and Gray, J. A. (1992). Classical conditioning after temporal lobe lesions in man: sparing of simple discrimination and extinction. *Behav. Brain Res.* 52, 159–165.

Daum, I., Channon, S., Polkey, C. E., and Gray, J. A. (1991). Classical conditioning after temporal lobe lesions in man: impairment in conditional discrimination. *Behav. Neurosci.* 105, 396–408.

Davis, M. H., Coleman, M. R., Absalom, A. R., Rodd, J. M., Johnsrude, I. S., Matta, B. F., Owen, A. M., and Menon, D. K. (2007). Dissociating speech perception and comprehension at reduced levels of awareness. *Proc. Natl. Acad. Sci. U.S.A.* 104, 16032–16037.

Dawson, M. E., and Reardon, P. (1973). Construct validity of recall and recognition postconditioning measures of awareness. *J. Exp. Psychol.* 98, 308–315.

Dehaene, S., and Changeux, J. P. (2011). Experimental and theoretical approaches to conscious processing. *Neuron* 70, 200–227.

Dehaene, S., and Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition* 79, 1–37.

Destrebecqz, A., and Cleeremans, A. (2001). Can sequence learning be implicit? New evidence with the process dissociation procedure. *Psychon. Bull. Rev.* 8, 343–350.

Destrebecqz, A., Peigneux, P., Laureys, S., Degueldre, C., Del Fiore, G., Aerts, J., Luxen, A., Van Der Linden, M., Cleeremans, A., and Maquet, P. (2005). The neural correlates of implicit and explicit sequence learning: interacting networks revealed by the process dissociation procedure. *Learn. Mem.* 12, 480–490.

Destrebecqz, A., Perruchet, P., Cleeremans, A., Laureys, S., Maquet, P., and Peigneux, P. (2010). The influence of temporal factors on automatic priming and conscious expectancy in a simple reaction time task. *Q. J. Exp. Psychol.* 63, 291–309.

Dickinson, A. (1980). *Contemporary Animal Learning Theory*. Cambridge: Cambridge University Press.

Eichenbaum, H., and Cohen, N. J. (2000). *From Conditioning to Conscious Recollection: Memory Systems of the Brain*. New York: Oxford University Press.

Esteves, F., Parra, C., Dimberg, U., and Ohman, A. (1994). Nonconscious associative learning: Pavlovian conditioning of skin conductance responses to masked fear-relevant facial stimuli. *Psychophysiology* 31, 375–385.

Fiset, P., Paus, T., Daloze, T., Plourde, G., Meuret, P., Bonhomme, V., Hajj-Ali, N., Backman, S. B., and Evans, A. C. (1999). Brain mechanisms of propofol-induced loss of consciousness in humans: a positron emission tomographic study. *J. Neurosci.* 19, 5506–5513.

Gerwig, M., Kolb, F. P., and Timmann, D. (2007). The involvement of the human cerebellum in eyeblink conditioning. *Cerebellum* 6, 38–57.

Gilbert, C. D., Sigman, M., and Crist, R. E. (2001). The neural basis of perceptual learning. *Neuron* 31, 681–697.

Glanzman, D. L. (1995). The cellular basis of classical conditioning in *Aplysia californica* – it's less simple than you think. *Trends Neurosci.* 18, 30–36.

Graziano, M., and Sigman, M. (2009). The spatial and temporal construction of confidence in the visual scene. *PLoS ONE* 4, 4909. doi:10.1371/journal.pone.0004909

Grol, M. J., de Lange, F. P., Verstraten, F. A., Passingham, R. E., and Toni, I. (2006). Cerebral changes during performance of overlearned arbitrary visuomotor associations. *J. Neurosci.* 26, 117–125.

Han, C. J., O'Tuathaigh, C. M., van Trigt, L., Quinn, J. J., Fanselow, M. S., Mongeau, R., Koch, C., and Anderson, D. J. (2003). Trace but not delay fear conditioning requires attention and the anterior cingulate cortex. *Proc. Natl. Acad. Sci. U.S.A.* 100, 13087–13092.

Heinemann, A., Kunde, W., and Kiesel, A. (2009). Context-specific prime-congruency effects: on the role of conscious stimulus representations for cognitive control. *Conscious. Cogn.* 18, 966–976.

Heisenberg, M., Wolf, R., and Brembs, B. (2001). Flexibility in a single

behavioral variable of *Drosophila*. *Learn. Mem.* 8, 1–10.

Ivkovich, D., and Stanton, M. E. (2001). Effects of early hippocampal lesions on trace, delay, and long-delay eyeblink conditioning in developing rats. *Neurobiol. Learn. Mem.* 76, 426–446.

James, G. O., Hardiman, M. J., and Yeo, C. H. (1987). Hippocampal lesions and trace conditioning in the rabbit. *Behav. Brain Res.* 23, 109–116.

Jennett, B., and Plum, F. (1972). Persistent vegetative state after brain damage: a syndrome in search of a name. *Lancet* 1, 734–737.

Koch, C. (2004). *The Quest for Consciousness*. Englewood, CO: Roberts and Publisher.

Lavond, D. G., Kim, J. J., and Thompson, R. F. (1993). Mammalian brain substrates of aversive classical conditioning. *Annu. Rev. Psychol.* 44, 317–342.

Lewin, R. (1983). "How did vertebrates take to the air?" *Science* 221, 38–39.

Lovibond, P. F., and Shanks, D. R. (2002). The role of awareness in Pavlovian conditioning: empirical evidence and theoretical implications. *J. Exp. Psychol. Anim. Behav. Process.* 28, 3–26.

Manns, J. R., Clark, R. E., and Squire, L. R. (2000a). Awareness predicts the magnitude of single-cue trace eyeblink conditioning. *Hippocampus* 10, 181–186.

Manns, J. R., Clark, R. E., and Squire, L. R. (2000b). Parallel acquisition of awareness and trace eyeblink. *Learn. Mem.* 7, 267–272.

Mauk, M. D., and Thompson, R. F. (1987). Retention of classically conditioned eyelid responses following acute decerebration. *Brain Res.* 403, 89–95.

McCormick, D. A., and Thompson, R. F. (1984). Neuronal responses of the rabbit cerebellum during acquisition and performance of a classically conditioned nictitating membrane-eyelid response. *J. Neurosci.* 4, 2811–2822.

Merikle, P. M., and Daneman, M. (2000). "Conscious vs. unconscious perception," in *The New Cognitive Neurosciences*, 2nd Edn, ed. M. S. Gazzaniga (Cambridge, MA: MIT Press), 1295–1303.

Morris, J. S., Ohman, A., and Dolan, R. J. (1999). A subcortical pathway to the right amygdala mediating "unseen" fear. *Proc. Natl. Acad. Sci. U.S.A.* 96, 1680–1685.

Moyer, J. R., Deyo, R. A., and Disterhoft, J. F. (1990). Hippocampectomy disrupts associative learning of the trace conditioned eye-blink response in rabbits. *Behav. Neurosci.* 104, 243–252.

Núñez, J. P., and de Vicente, F. (2004). Unconscious learning. Conditioning to subliminal stimuli. *Span. J. Psychol.* 7, 13–28.

Ohman, A., Esteves, F., and Soares, J. J. F. (1995). Preparedness and preattentive associative learning: electrodermal conditioning to masked stimuli. *J. Psychophysiol.* 9, 99–108.

Ohman, A., and Soares, J. J. F. (1993). On the automatic nature of phobic fear: conditioned electrodermal responses to masked fear-relevant stimuli. *J. Abnorm. Psychol.* 102, 121–132.

Ohman, A., and Soares, J. J. F. (1994). "Unconscious anxiety": phobic responses to masked stimuli. *J. Abnorm. Psychol.* 103, 231–240.

Ohman, A., and Soares, J. J. F. (1998). Emotional conditioning to masked stimuli: expectancies for aversive outcomes following nonrecognized fear-relevant stimuli. *J. Exp. Psychol. Gen.* 127, 69–82.

Perruchet, P. (1985). A pitfall for the expectancy theory of human eyelid conditioning. *Pavlov. J. Biol. Sci.* 20, 163–170.

Perruchet, P., Cleeremans, A., and Destrebecqz, A. (2006). Dissociating the effects of automatic activation and explicit expectancy on reaction times in a simple associative learning task. *J. Exp. Psychol. Learn. Mem. Cogn.* 32, 955–965.

Pessoa, L. (2005). To what extent are emotional visual stimuli processed without attention and awareness? *Curr. Opin. Neurobiol.* 15, 188–196.

Rescorla, R. A. (1988). Pavlovian conditioning: it's not what you think it is. *Am. Psychol.* 43, 151–160.

Scott, R. B., Minati, L., Dienes, Z., Critchley, H. D., and Seth, A. K. (2011). Detecting conscious awareness from involuntary autonomic responses. *Conscious. Cogn.* 20, 936–942.

Sehlmeyer, C., Schöning, S., Zwitserlood, P., Pfleiderer, B., Kircher, T., Arolt, V., and Konrad, C. (2009). Human fear conditioning and extinction in neuroimaging: a systematic review. *PLoS One* 4(6):e5865.

Shanks, D. R. (1995). *The Psychology of Associative Learning*. Oxford: Oxford University Press.

Simmons, N. B., Seymour, K. L., Habersetzer, J., and Gunnell, G. F. (2008). Primitive early Eocene bat from Wyoming and the evolution of flight and echolocation. *Nature* 451, 818–821.

Squire, L. R. (2004). Memory systems of the brain: a brief history and current perspective. *Neurobiol. Learn. Mem.* 82, 171–177.

Stamatakis, E. A., Adapa, R. M., Absalom, A. R., and Menon, D. K. (2010). Changes in resting neural connectivity during propofol sedation. *PLoS ONE* 5, e14224. doi:10.1371/journal.pone.0014224

Sullivan, C., Hone, D., Xu, X., and Zhang, F. (2010). The asymmetry of the carpal joint and the evolution of wing folding in maniraptoran theropod dinosaurs. *Proc. Biol. Sci.* 277, 2027–2033.

Takehara, K., Kawahara, S., and Kirino, Y. (2003). Time-dependent reorganization of the brain components underlying memory retention in trace eyeblink conditioning. *J. Neurosci.* 23, 9897–9905.

Van den Bussche, E., Segers, G., and Reynvoet, B. (2008). Conscious and unconscious proportion effects in masked priming. *Conscious. Cogn.* 17, 1345–1358.

Weidemann, G., Tangen, J. M., Lovibond, P. F., and Mitchell, C. J. (2009). Is Perruchet's dissociation between eyeblink conditioned responding and outcome expectancy evidence for two learning systems? *J. Exp. Psychol. Anim. Behav. Process.* 35, 169–176.

Weike, A. I., Schupp, H. T., and Hamm, A. O. (2007). Fear acquisition requires awareness in trace but not delay conditioning. *Psychophysiology* 44, 170–180.

Wiens, S., and Ohman, A. (2002). Unawareness is more than a chance event: comment on Lovibond and Shanks. *J. Exp. Psychol. Anim. Behav. Process.* 28, 27–31.

Woodruff-Pak, D. S., and Disterhoft, J. F. (2008). Where is the trace in trace conditioning? *Trends Neurosci.* 31, 105–112.

Yanoviak, S. P., Kaspari, M., and Dudley, R. (2009). Gliding hexapods and the origins of insect aerial behaviour. *Biol. Lett.* 5, 510–512.

Zylberberg, A., Dehaene, S., Roelfsema, P. R., and Sigman, M. (2011). The human Turing machine: a neural framework for mental programs. *Trends Cogn. Sci. (Regul. Ed.)* 15, 293–300.

Zylberberg, A., Fernández Slezak, D., Roelfsema, P. R., Dehaene, S., and Sigman, M. (2010). The brain's router: a cortical network model of serial processing in the primate brain. *PLoS Comput. Biol.* 6, e1000765. doi:10.1371/journal.pcbi.1000765

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

Received: 30 September 2010; accepted: 31 October 2011; published online: 06 December 2011.

Citation: Bekinschtein TA, Peeters M, Shalom D and Sigman M (2011) Sea slugs, subliminal pictures, and vegetative state patients: boundaries of consciousness in classical conditioning. Front. Psychology ***2****:337. doi: 10.3389/fpsyg.2011.00337*

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

frontiers in PSYCHOLOGY

HYPOTHESIS AND THEORY ARTICLE
published: 26 January 2011
doi: 10.3389/fpsyg.2011.00007

Reorganization of the injured brain: implications for studies of the neural substrate of cognition

Jesper Mogensen*

The Unit for Cognitive Neuroscience, Department of Psychology, University of Copenhagen, Copenhagen, Denmark

Edited by:
Morten Overgaard, Aarhus University, Denmark

Reviewed by:
Gerald Zernig, Medical University Innsbruck, Austria

****Correspondence:***
Jesper Mogensen, Department of Psychology, The Unit for Cognitive Neuroscience, University of Copenhagen, Oester Farimagsgade 2A, DK-1353 Copenhagen K, Denmark.
e-mail: jesper.mogensen@psy.ku.dk

In the search for a neural substrate of cognitive processes, a frequently utilized method is the scrutiny of post-traumatic symptoms exhibited by individuals suffering focal injury to the brain. For instance, the presence or absence of conscious awareness within a particular domain may, combined with knowledge of which regions of the brain have been injured, provide important data in the search for neural correlates of consciousness. Like all studies addressing the consequences of brain injury, however, such research has to face the fact that in most cases, post-traumatic impairments are accompanied by a "functional recovery" during which symptoms are reduced or eliminated. The apparent contradiction between localization and recovery, respectively, of functions constitutes a problem to almost all aspects of cognitive neuroscience. Several lines of investigation indicate that although the brain remains highly plastic throughout life, the post-traumatic plasticity does not recreate a copy of the neural mechanisms lost to injury. Instead, the uninjured parts of the brain are functionally reorganized in a manner which – in spite of not recreating the basic information processing lost to injury – is able to allow a more or less complete return of the surface phenomena (including manifestations of consciousness) originally impaired by the trauma. A novel model [the Reorganization of Elementary Functions-model] of these processes is presented – and some of its implications discussed relative to studies of the neural substrates of cognition and consciousness.

Keywords: neural plasticity, reorganization, brain injury, localization of function, recovery, recovery of function, consciousness, neural organization

INTRODUCTION

Research into the relationship between on the one hand brain injury [be it vascular or traumatic (TBI)] and on the other hand various cognitive processes remains a crucial part of cognitive neuroscience. Some of the consciousness-related phenomena studied in brain injured patients are hemispatial neglect and blindsight. There is a constantly growing body of clinical as well as experimental data on various types of neglect and blindsight in focally lesioned patients (e.g., Natsoulas, 1997; Marcel, 1998; Rossetti et al., 1998; Kentridge et al., 1999; Danckert and Goodale, 2000; Schindler et al., 2006; Bartolomeo, 2007; Silvanto et al., 2008; Funk et al., 2010). Even within other areas than visual perception lesion-induced dissociations between conscious and non-conscious processes are being addressed in attempts to yield information regarding the neural processes mediating subjective consciousness (e.g., Lane et al., 1997). Studies focusing on the manifestations of consciousness in brain injured patients fall within the framework of research addressing "localization of functions" within the brain. Other cognitive domains exhibit a constantly growing body of such studies, e.g., language (e.g., Thomas et al., 1997; Thulborn et al., 1999; Ansaldo et al., 2002; Ansaldo and Arguin, 2003; Perani et al., 2003; Baumgaertner et al., 2005; Meinzer et al., 2008; Specht et al., 2009; Szaflarski et al., 2010). Associating various functions with specific brain regions is one of the main research traditions of cognitive neuroscience. But it is also an endeavor facing both methodological and theoretical challenges (e.g., Mogensen and Malá, 2009) – not the least the apparent contradiction between the localization and post-traumatic recovery, respectively, of various "functions." Focusing on the recently developed Reorganization of Elementary Functions (REF)-model (Mogensen and Malá, 2009) the present communication discusses these issues with respect to the understanding of the neural substrate and post-traumatic recovery of cognitive functions in general. A subsequent paper (Overgaard and Mogensen, 2011) will deal more specifically with the issue of what is required in terms of theoretical conceptualization as well as experimental documentation if post-traumatic functional recovery is to be taken as documentation of multiple realizations of neural substrates of consciousness.

LOCALIZATION AND/OR RECOVERY

There can be little doubt that the brain is regionally specialized. Various brain structures and substructures perform apparently unique types of information processing and consequently participate differentially in the mediation of various types of behavior and cognition. In other and more commonly used words: the functions of the brain are regionally localized. The main support for this assumption grows out of two research traditions. One of these is the study of post-traumatic impairments in brain injured individuals – patients or experimental animals. In such studies, the profile of post-injury symptoms is compared to the localization and extent of brain injury. The logics of these lesion experiments (see e.g., Coltheart, 2001; Selnes, 2001; controlled experiments in animal models and clinical studies in brain injured patients) is that in the absence of a brain structure, the symptoms must reflect

the absence of functional contributions from the affected neural machinery – and that with the proper analysis and comparisons across symptoms, conclusions can be drawn regarding the information processing of the missing brain circuitry. A more recent contribution to this type of study is the use of transcranial magnetic stimulation (e.g., Pascual-Leone et al., 1999, 2000; Walsh and Cowey, 2000) in which the functional integrity of a part of the brain is temporarily disturbed – allowing an analysis of the consequences of what can be seen as a "temporary lesion." The other main source of support for functional localization within the brain is the studies utilizing various types of neuroimaging techniques. By studying the regional pattern of brain activation during the performance of various tests, one can – provided adequate baseline measures are utilized for the "subtraction" from the test condition – provide information about whether or not a particular brain structure changes its level of activity (often reflected directly or indirectly as a change in metabolism and blood supply) in association with the performance of a particular task or stimulation. Numerous contradictions exist within these branches of the neuroscientific literature, but there is an overall agreement that although often poorly understood, there is a regional functional specialization within the brain – a "functional localization" (e.g., Monakow, 1914; Coltheart, 2001; Selnes, 2001; Kringelbach and Rolls, 2004).

As mentioned above, a crucial aspect of the interpretation of results from various types of "lesion experiments" is that in the absence of a brain structure – and consequently the functional contributions from that circuitry – the behavior and conscious manifestations of the individual must reflect the lack of whatever functional contributions were provided by the now missing part of the brain. But if it is assumed that the lost circuitry is post-traumatically never regained (an issue to which I will return later), one should expect the impairments after brain injury to be chronic. Nevertheless, it is a well-established fact that a post-traumatic functional recovery does occur.

In patients as well as in animal models of brain injury, most lesion-associated impairments post-traumatically undergo some level of "recovery" – at least within cognitive domains not closely linked to the direct in- and output pathways. Less or even no recovery may be seen within some sensory and motoric domains. But for almost all "higher" cognitive functions, trauma-related impairments are followed by an apparent return toward the proficiency seen pre-traumatically (e.g., Ramachandran and Blakeslee, 1998; Carney et al., 1999; Buller and Hardcastle, 2000; Panksepp and Panksepp, 2000; León-Carrión and Machuca-Murga, 2001; Mogensen et al., 2004, 2007; Mogensen and Malá, 2009; Rohling et al., 2009). In most cases, such a functional recovery is associated with more or less formalized and institutionalized rehabilitative training, but in the absence of such training, "spontaneous" recovery is also seen (e.g., León-Carrión and Machuca-Murga, 2001). It must, however, be remembered that even in the absence of a formalized post-traumatic training program, practically all brain injured organisms (patients and experimental animals alike) are subjected to the informal training of daily life activities. Even the most basic daily activities and communicative efforts constitute challenges and tasks which the brain injured individual must attempt to meet and master. Consequently, the absence of an externally imposed training regime does not allow a claim that the potential occurrence of a functional recovery is independent of interactions with the environment. Some instances in which a "spontaneous" recovery may be of a more automatic nature can, however, be seen in cases where a lesion-induced "penumbra" gradually disappears. The "penumbra" phenomenon can briefly be described as a situation in which injury within one part of the brain causes other brain areas to receive a reduced level of blood supply. While being sufficient for the survival of neurons within that penumbra region, the reduced blood supply does not allow a normal level of functionality. Consequently, the observable symptoms are not only associated with the trauma *per se* but also with the impaired neural activities within the penumbra. Penumbras, however, mostly disappear spontaneously and allow a return to normal levels of functional performance within that region of the brain (e.g., Choi et al., 2007).

Although sometimes incomplete, this post-traumatic functional recovery may both clinically and in animal models turn out to be "complete" – defined as the acquisition of a post-traumatic proficiency equal to that seen in the absence of any brain injury (e.g., Mogensen et al., 2004). In case of animal models, this is even seen under circumstances ensuring the complete removal of the brain structure in question as well as a well-established pre-traumatic functional baseline.

If the relatively few instances in which post-traumatic symptoms can be associated with penumbras or similar phenomena are excluded, what remains is the apparent contradiction between the two phenomena of "functional localization" and "recovery of function."

A radical and maybe tempting way to utilize lesion experiments without having to deal with the contradiction between localization and recovery of functions is to accept only the post-traumatic symptoms as relevant to arguments regarding functional specialization, in case those symptoms turn out to be "chronic" – never to demonstrate any functional recovery. Such a radical position has been advocated by Olton (1978). It has to be realized, however, that if only those instances in which post-traumatic impairments persist chronically are to be considered when conclusions are drawn from lesion experiments, the vast majority of such clinical and animal model derived data would have to be discarded.

But the arguments against dismissing all but the chronic post-traumatic symptoms for consideration when the principles of localization are considered, are not only of such a practical nature (although disregarding almost the entire mass of post-traumatic data is in itself not an insignificant obstacle!). The fact that a functional recovery can actually take place in spite of the continued presence of a lesion, which originally had such an impact on the information processing of the brain that significant symptoms occurred, is in itself a highly relevant phenomenon. It indicates dynamic changes, which must be an essential part of the functional organization of the brain. If these phenomena are not considered in neuroscientific studies and the construction of various models, the result will be an incomplete and lacking understanding of the functionally dynamic brain.

But in which ways, then, can one imagine the apparent return (potentially to a normal level of proficiency) of the behavioral and cognitive abilities of a brain injured organism? One possibility is that post-traumatically the brain is able to reconstruct – within the region of injury or elsewhere – a circuitry, which can accomplish

an information processing similar to what has been lost to injury. Although it is well known that the actual site of injury mostly turn into scar tissue, this does not in itself preclude the possibility of a recreation somewhere in the brain of circuitry fulfilling the same information processing demands as the lost structure.

IS THE BASIC CIRCUITRY OF THE BRAIN RECREATED POST-TRAUMATICALLY?

In order to evaluate the likeliness that a reconstruction of the lost circuitry can occur in the injured brain, it is relevant to compare the types of plasticity available in the injured adult brain to the plastic processes, which contribute to the original construction of such networks during development.

When the adult brain is injured, the trauma itself induces a range of changes in gene expression – both in the tissue immediately surrounding the site of injury and in more remote parts of the brain. Some such changes are likely primarily to contribute to the detrimental effects of the injury – for instance by promoting metabolic dysfunction, inflammatory responses, etc. (e.g., Delfs et al., 1995; Hermann et al., 1999; Harris et al., 2001; Rao et al., 2003). Other changes in gene expression are likely to be supportive of rehabilitation – promoting plasticity and reorganization (e.g., Witte, 1998; Frost et al., 2003; Kleim et al., 2003; Nudo, 2003). One such process may be the increased occurrence of long-term potentiation (LTP) like synaptic plasticity after brain injury (e.g., Hagemann et al., 1998).

Relative to the issue of recreation of circuitry similar to what has been lost to injury, it might be especially relevant that injury to the brain potentiates the ongoing neurogenesis (e.g., Magavi et al., 2000; Scharff et al., 2000; Arvidsson et al., 2002; Nakatomi et al., 2002; Chen et al., 2004). A reason for this to be important is that during maturation neurons undergo a number of changes – reducing their similarity to the developing neurons, which originally formed various circuits (e.g., Fawcett et al., 1989; Chen et al., 1995; Goldberg et al., 2002). In contrast, the newly formed neurons produced by adult neurogenesis are unlikely to have similar limitations.

If the newly created neurons in the injured brain are to reach a specific destination and contribute to recreation of a circuit, they will have to migrate in an appropriate manner. After injury, mature astrocytes are able to transform themselves into radial glial cells similar to those guiding neural migration during development (e.g., Rakic, 1971, 1985). Such radial glial cells are able to guide the migration of immature neurons even in the adult brain (e.g., Leavitt et al., 1999). Further optimism regarding the potentials of the injured adult brain may come from the fact that most substances, which played a guiding role during the original outgrowth of dendrites and axons (e.g., Keynes and Cook, 1992; Brose et al., 1999; Chen et al., 2000; Hiramoto et al., 2000; Polleux et al., 2000) are also present in the adult brain (e.g., Koeberle and Bahr, 2004). This optimism may, however, be tempered by the observation that the distribution of these substances undergoes major changes during the maturation of the brain – making it questionable whether they in an injured adult brain can play similar roles to those of development (e.g., Harel and Strittmatter, 2006). The most important factors preventing immature neurons in the adult nervous system from recreating the injured circuitry, may, however, be associated with glial cells and myelin (e.g., Berry, 1982; Schwab and Thoenen, 1985; Schäfer et al., 2008). Especially important may be the astrocyte-produced chondroitin sulfate proteoglycans (CSPGs), which play an important role in terminating the developmentally "critical" periods (e.g., Pizzorusso et al., 2002; Berardi et al., 2004; McGee et al., 2005). While consolidating the plastic processes occurring during critical periods, these substances may also play a role, which prevents an adult recreation of the circuit lost to injury (e.g., Del Rio and Soriano, 2007; Schäfer et al., 2008). An improved functional recovery (potentially associated with recreation of lost circuitry) has been found when the CSPGs are pharmacologically inhibited locally (e.g., Del Rio and Soriano, 2007). While such a local inhibition may have therapeutic potentials in the future, the results also demonstrate that without an external intervention, the CSPGs are likely to prevent or at least reduce the possibility of a post-traumatic re-establishment of the circuitry and thereby information processing lost to trauma.

THE MECHANISMS OF POST-TRAUMATIC FUNCTIONAL RECOVERY

While it is important in these ways to establish whether the post-traumatically available neuroplasticity seems capable of processes, which can create a copy of what has been lost – another and at least as important approach is to scrutinize the neural and cognitive processes accompanying the actual functional recovery. An extensive animal model-based research program (e.g., Mogensen et al., 2002, 2003, 2004, 2005, 2007) has performed such an analysis and some of the results have been reviewed by Mogensen and Malá (2009). What has emerged is a pattern of principles regarding the mechanisms mediating post-traumatic recovery. Three general principles are especially important – describing the situation after a successful post-traumatic rehabilitation:

1. Modification of degree of contribution to task mediation by individual brain structures
   - Some structures exhibit an increased or decreased level of contribution to task mediation.
2. Task dependent and dissimilar neural substrates
   - After a given lesion, the functional recovery of various cognitive tasks is mediated by unique and dissimilar neural substrates.
3. Application of new cognitive strategies
   - The fully post-traumatically recovered individuals solve the task by applying new strategies that are dissimilar to those applied pre-traumatically.

These three principles – like the above consideration of the types of plasticity available in the developing and injured, mature brain, respectively – indicate that a recreation of the lost circuitry is unlikely. Both the second and third of the above principles show that the neural mechanisms mediating post-traumatic functional recovery do not include a copy of what has been lost to trauma. If any part of the injured brain at the end of rehabilitation training contained a circuitry similar to what was pre-traumatically available, it would be expected that all cognitive domains affected by the lesion would post-traumatically receive equal contribution to functional recovery from the brain region within which the circuitry had been (re)created. This possibility, however, is contradicted by

Principle 2. Additionally, Principle 3 contradicts the post-traumatic re-establishment of information processing identical to what was available pre-traumatically: if post-traumatic processes had re-established the information processing of the injured structure, one would expect not only task solution of a proficiency similar to that seen preoperatively, but also that such a task solution would employ similar strategies to those of the pre-traumatic situation.

But if the post-traumatic functional recovery is not mediated via mechanisms recreating what has been lost to injury, how can a sometimes even complete level of proficiency be re-established within traumatically impaired cognitive domains?

The recently proposed REF model of Mogensen and Malá (2009) is an attempt to describe neural and cognitive mechanisms, which – in spite of the absence of a recreation of the lost circuitry – can account for a potentially full proficiency of post-traumatic cognitive recovery.

At the most basic level of the REF-model are the information processing modules named Elementary Functions (EFs). The EFs are truly localized in the sense that they are mediated by local circuitry within a structure of the brain. Each traditionally defined neural structure (e.g., the hippocampus or the dorsolateral prefrontal cortex) contains the neural substrates of a huge number of EFs. When a region of the brain is lost to injury, all EFs mediated by the lost tissue are – according to the REF-model – irreversibly lost, too. The information processing of an individual EF is of a highly basic and "modular" type. An EF does not in itself have any of the "functions" traditionally described by psychology. Psychologically defined functions such as object discrimination, explicit memory, or allocentric spatial orientation belong at a different level of analysis (to be described shortly). The "function" of an EF may more easily be described in mathematical terms – rather than in the vocabulary of for instance cognitive psychology. In contrast, the "functions" and cognitive domains normally described by psychology are in the REF-model represented at the third and highest level – the level of surface phenomena. The surface phenomena of the REF-model consist of observable behavior (for instance the performance of a task by a patient or an experimental animal) as well as conscious manifestations such as the subjective experience of recognizing a familiar face or planning a course of action in order to solve a problem. It is at the level of these surface phenomena that post-traumatic functional recovery is normally defined and evaluated. Tests conducted in order to determine what is impaired by brain injury address the observable behavior as well as whatever representations can be obtained of the subjective experiences of the patient. And it is at the same level that the more or less complete recovery of these impaired dimensions of cognition and consciousness are determined – using similar methods to those utilized during diagnosis. An essential component of the REF-model is the layer of analysis introduced between the basic layer of the EFs and the uppermost layer of the surface phenomena. This intermediate level consists of the processes named algorithmic strategies (ASs).

Each AS consists of numerous interacting EFs. Most – if not all – ASs are established as the result of experience and learning. Most EFs are simultaneously part of several or many ASs. While the individual EFs are strictly localized within a subregion of the brain, the neural substrate of an AS consists of the neural substrates of all its constituent EFs as well as the interconnections between the neural substrates of these EFs. This makes the neural substrate of an AS highly distributed and in most cases components of the neural substrate of an AS will be found within a number of brain structures. The information processing of an AS is the mechanism enabling a specific surface phenomenon. For instance, a specific type of solution of a task is obtained by activation of a particular AS. Most surface phenomena can be realized via the activity of multiple ASs. Unless special analytical techniques are employed, it may at the surface level not be possible to discriminate between behavioral or conscious phenomena reflecting two related but different ASs. Whenever brain injury destroys the neural substrate of one or more of the constituent EFs within an AS, that AS is lost. Consequently, the surface phenomena associated with the activity of the lost AS are also lost and post-traumatically an impairment is registered.

According to the REF-model, the mechanisms enabling a post-traumatic functional recovery are special cases of a more general mechanism, which in the intact brain has evolved as a crucial aspect of learning and problem solving. When an individual encounters a situation calling for a task solution for which there is no established procedure available, a mechanism is initiated during which various existing ASs are "tested out." A "selector/evaluator" mechanism – which resembles (without being identical to) the "Supervisory Attentional System" (SAS) of Norman and Shallice (1986) – sequentially activates existing ASs. When activated, an AS results in the associated surface phenomenon and, in turn, the quality of the resultant behavior or mental manifestation is evaluated. In case the desired result is obtained, activation of that AS will in the future be associated with the situation in question. This mechanism resembles – without being identical to – the mechanism of "hypotheses" evaluation described by Krechevsky (1932, 1933). In case activation of an existing AS can obtain the desired result, the neural plasticity associated with this entire process is restricted to modifications within the selector/evaluator mechanism – plasticity ensuring a future association between the situation in question and activation of the successfully selected AS.

If, however, a situation requiring the solution of a problem cannot be solved by activation of any existing AS, a novel AS will have to be established. The creation of a novel AS involves a reorganization of the functional interaction between EFs. This is the "Reorganization of Elementary Functions" process, which has given name to the REF-model. Such a reorganization (see Mogensen and Malá, 2009) utilizes a type of process resembling the backpropagation algorithm (e.g., Rumelhart and McClelland, 1986; Werbos, 1994). Such mechanisms – constantly utilizing the feedback of the environment – form a novel AS by combining a set of EFs, which previously did not constitute an interacting entity. Whenever such a REF-process is required in order to successfully obtain a task solution, the required neuroplasticity includes modified connections between the neural substrates of the constituent EFs. Additionally, the complete process of eventual activation of the novel AS and its association with the situation in question is also associated with neuroplasticity within the evaluator/selector mechanism. A schematic representation of these processes is given in **Figure 1**.

These processes of selection and potentially even *de novo* establishment of ASs when a novel situation is encountered are according to the REF-model essential mechanisms in the mediation of normal problem solving. What is special about brain injury is that many

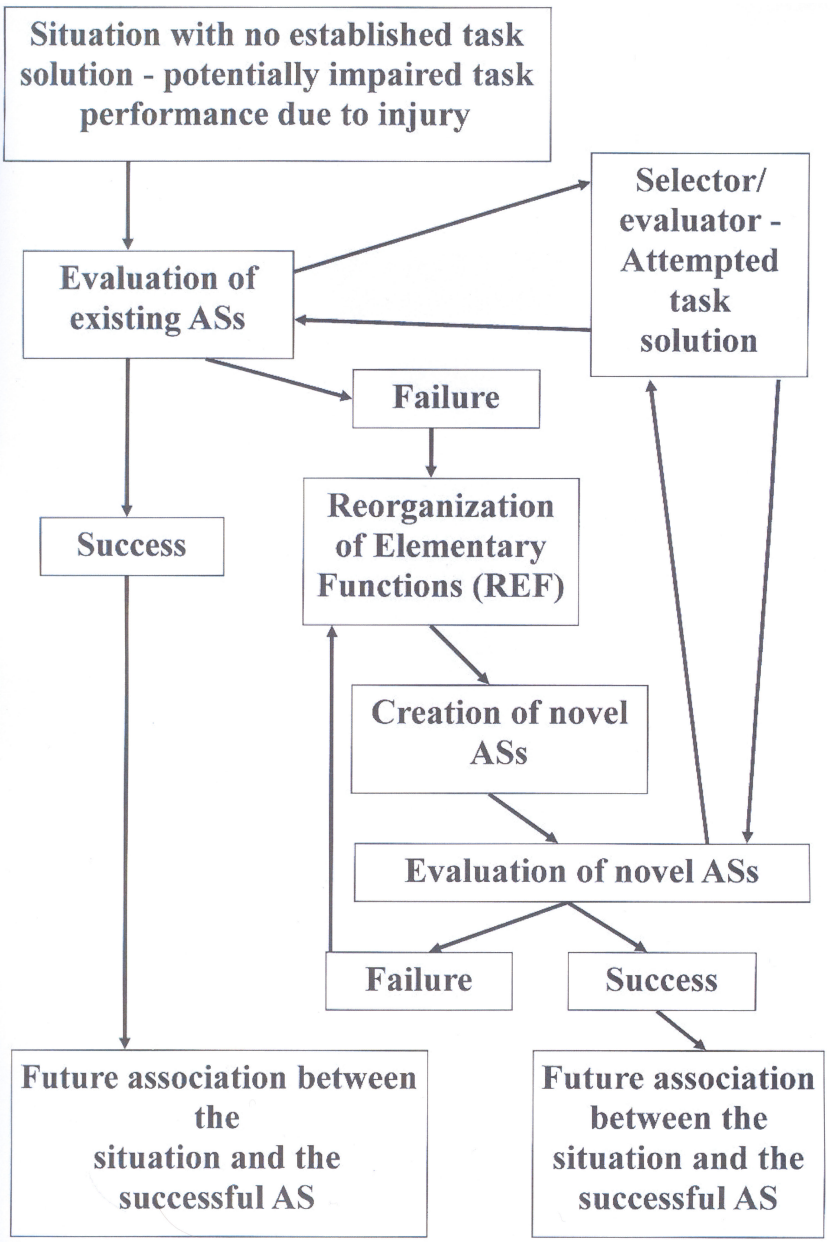


FIGURE 1 | Flow diagram depicting the sequence of events, which according to the REF-model leads to a successful development of a task solution – potentially a successful functional recovery after brain injury. These processes are always associated with plastic modifications within the selector/evaluator mechanisms. Additional plasticity modifying the connections between the neural substrates of EFs is only expected in case an actual Reorganization of Elementary Functions (REF) process is required. For further details: see the present text as well as Figures 3, 4, and 5 in Mogensen and Malá (2009).

situations, for which there used to be an established mechanism of task solution, will after the traumatic event have a status similar to novel situations. The AS, which would normally be activated and ensure an efficient task solution, is no longer available since some of its constituent EFs have been lost to injury. Consequently, there is no immediately available mechanism of task solution. Which leads to the above-described mechanisms of initial search for an appropriate AS. Since multiple ASs might pre-traumatically have been able to allow an efficient task solution, some of these might post-traumatically still be available. If this is the case, the search for an available and appropriate AS will be successful. Alternatively, the REF-process (including backpropagation-based reorganization of interconnectivity between the neural substrates of EFs) is required in order to obtain a satisfactory solution to the task. Viewed in this way, the processes allowing a post-traumatic functionally recovery to take place have mostly if not exclusively evolved as mechanisms mediating problem solving in the intact brain.

In one study addressing the types of neural and cognitive reorganizational processes described by the REF-model (Mogensen et al., 2004) the mechanisms of post-traumatic functional recovery

of a water maze based allocentric place learning task of the mapping type was addressed in rats subjected to various types of focal brain injury. Lesions of the hippocampus provoked a major functional impairment, which, however, disappeared completely during an approximately 1 month long post-traumatic training period – leaving the animals capable of a task proficiency indistinguishable from that of intact rats. This recovery process depended on contributions from the prefrontal cortex. Relying on partly prefrontal cortical mechanisms, the animals could thus achieve a fully proficient surface phenomenon – task performance. It turned out, however, that even rats subjected to hippocampal lesions as well as removal of the prefrontal cortex were able to achieve an equal – that is: normal – proficiency of task performance within a similar period of training. In the absence of both hippocampal and prefrontal contributions to the mediation of task performance, the neural substrate of task solution appeared to depend upon neural mechanisms within the parietal association cortex. In terms of behavioral parameters such as the time and swim distance required to reach the hidden target location, functional recovery mediated by mechanisms within the prefrontal cortex and parietal association cortex, respectively, were of equal proficiency. In the terminology of the REF-model this means that equally proficient ASs can be constructed on the basis of populations of EFs including either prefrontally based EFs or EFs mediated by the parietal association cortex, respectively. Although equally proficient, the ASs relying on prefrontal and parietal mechanisms, respectively, differed with respect to the cognitive mechanisms (as opposed to proficiency) of task solution (as would be expected from the REF-model): While the recovered task solution in animals relying on prefrontal task mediation included cognitive representations of the goal position, such a "knowledge" of the spatial location of the goal appeared to be absent in even fully recovered animals relying on mechanisms within the parietal association cortex (Mogensen et al., 2004).

The situation in which a "complete post-traumatic functional recovery" is encountered as described at the level of surface phenomena, the actual situation is that the post-traumatically selected and potentially established AS is able to allow such a proficiency of task solution that – using standard methods of analysis – the result cannot be distinguished from the pre-traumatic situation. Although each AS is, in fact, associated with a particular way of solving the task, the observable behavior or subjective experience at the level of surface phenomena may in all of these cases be similar enough to be (in a sense wrongly) identified as "the same." And it is this "sameness" of in reality dissimilar surface phenomena that is a crucial aspect of the apparent contradiction between localization and post-traumatic recovery of functions. What is truly localized are the EFs, and when the neural substrate of these basic information processing entities is lost, there is no recovery of the "function" associated with that information processing. In contrast, functional recovery is identified at the level of surface phenomena where highly detailed and special analytical techniques are required in order to discriminate between the phenomena associated with activation of various ASs.

The dynamic reorganizations associated with the REF-process only affect the input/output relationships of the EFs. The actual information processing conducted by the individual EF remains unchanged. It does, however, contribute this information processing within a novel context – within the newly established AS. The situation in which an information processing module continues to perform its previous operations but on a novel input, bears a certain resemblance to some of the plastic processes found in uninjured brains. The somatosensory cortex, which after the amputation of a hand has become "vacant," continues its functional activities – but now operating on information regarding the face or arm (e.g., Yang et al., 1994; Weiss et al., 2000; Karl et al., 2001). Such relative shifts within the somatosensory representations can also be seen after intensive training restricted to part of the body (e.g., Merzenich and Jenkins, 1993; Elbert et al., 1995; Xerri et al., 1996; Münte et al., 2002). These plastic processes within the somatosensory system are not unique. Within the auditory system tonotopic representations undergo plastic changes due to changes in input or experience (e.g., Robertson and Irvine, 1989; Scheich, 1991; Recanzone et al., 1993; Irvine, 2007; Thai-Van et al., 2007). Another obvious parallel is the situation in which a cortical area specialized in analysis of figure orientation within the visual domain can become engaged in apparently similar or at least related information processing on somatosensory information in the blind (Ptito et al., 2005).

Given the highly specialized information processing units of the REF-model – the EFs – this model falls within what is called "Massive Modularity" by for instance Barrett and Kurzban (2006). It should, however, be noted that the modularity of the REF-model is far from identical to the kind of modularity described by Fodor (1983). That radical type of modularity has, however, subsequently been denounced by Fodor (2000) himself.

As emphasized above and by Mogensen and Malá (2009), according to the REF-model apparently the same surface phenomenon may be achieved by activation of a variety of ASs. In the terminology of Price and Friston (Price and Friston, 2002; Friston and Price, 2003) such a situation represents a degeneracy relative to the manifestation of the surface phenomena. In cases where multiple ASs give rise to surface phenomena that cannot be distinguished from each other and that surface phenomenon is then characterized as one "function," such a "function" is degenerate. It should, however, be stressed that according to the REF-model, such a degeneracy is the result of considering multiple surface phenomena, which might by a more or less superficial examination be indistinguishable from each other, "the same" – in spite of the fact that these surface phenomena are in reality different (although perhaps only marginally so). This analysis seems to be in agreement with a number of the examples given by Price and Friston (2002). Degeneracy is by Price and Friston (2002) illustrated by examples in which successful solution of a task (in the terminology of the REF-model: surface phenomenon) can be achieved via activity in separate and potentially not overlapping neural systems. It is recognized that – for instance in case of a linguistic task – the separate systems which are individually able to achieve a successful task solution do perform dissimilar types of information processing and thereby mediate dissimilar types of cognitive analysis. Such systems would in the REF-model be identified as separate ASs. While degeneracy might, thus, apply to the neural and cognitive mechanisms of a particular surface phenomenon, there is according to the REF-model no degeneracy with respect to the substrate of EFs. An EF and its neural substrate are unique and if lost due to injury not replaced.

IMPLICATION OF THE REORGANIZATIONS OF THE INJURED BRAIN

When studying brain injured individuals it is important to realize that the result of the recovery process is not – at the neural and more basic cognitive levels – a return to the pre-traumatic situation. Instead it constitutes a novel state of affairs, which has been constructed in an interaction with the environment during the period of rehabilitation. The ASs which form the basis of the post-traumatically observable surface phenomena have been selected and potentially constructed via the interactions between the injured individual and the broadly defined environment. This means that a post-traumatic recovery process can be very situational specific. While the symptoms within a cognitive domain can appear to have disappeared completely when tested in one situation and under certain circumstances, symptoms within the same cognitive domain in the same individual may be evident and for that matter show no signs of recovery when tested under different circumstances.

As has been emphasized elsewhere (e.g., Mogensen, 2003, 2011; Mogensen, in preparation; Overgaard and Mogensen, 2011; Wilms and Mogensen, in preparation) animal models have frequently demonstrated that while one variant (/setup) of a cognitive test is able to reveal a striking level of post-traumatic impairment, another test (/setup), which characterized according to the formal demands of that cognitive test must be considered identical, shows no sign of post-traumatic impairment.

Across all studied species, animals subjected to lesions within the prefrontal cortex or the associated structures such as the prefrontal part of the neostriatum show an impaired performance of the task known as spatial delayed alternation (e.g., Mogensen, 2003; Mogensen et al., 2007, 2008). Nevertheless, even this phenomenon is more situational specific than one would have expected. Mogensen et al. (1987) tested rats subjected to lesions of the prefrontal part of the neostriatum (a lesion which also renders the prefrontal cortex inoperable due to undercutting) in two variants of the spatial delayed alternation task. While the variant administered in a T-maze clearly revealed the expected symptoms, an operant chamber-based version of the task (in spite of fulfilling all the procedural/cognitive demands normally made on such a test setup) showed no sign of post-traumatic impairment. A somewhat related demonstration of the importance of the procedures employed in animal models can be found in a study by Lepore et al. (1985). Cats were tested for their ability to transfer visual discrimination-relevant information from one hemisphere to the other in the absence of the corpus callosum – the main pathway between the two hemispheres. After having acquired the visual discrimination task based on one hemisphere exclusively (information was provided via one eye only in animals with a split optic chiasm), cats subjected to lesions of the corpus callosum were tested for their ability to perform the task with only the contralateral eye open – a situation in which the task performance had to be based on the hemisphere contralateral to the one originally trained. The animals were tested for this ability in two experimental setups: a "Lashley-type" jumping-stand and a traditional (maze-like) two-choice discrimination box. When tested in the discrimination box, the cats did not demonstrate any ability to transfer information between the hemispheres subcallosally, while such an ability was clearly revealed when the test was performed in the jumping-stand.

While animal model-based studies are often able to more clearly demonstrate the procedure and setup-related differences in the degree or even presence of post-traumatic symptoms and post-traumatic recovery, clinical data are frequently of a more anecdotal nature. There are, however, studies in which such phenomena have been scrutinized in patients under controlled circumstances. One example is the results of Wilms and Malá (2010). The background of this study is a fascinating method of rehabilitative training of brain injured patients suffering hemispatial neglect (e.g., Rossetti et al., 1998): the prism adaptation therapy (PAT; e.g., Rossetti et al., 1998; Frassinetti et al., 2002). In this method, the patients are trained in a task requiring them to point (without being able visually to follow their arm during the pointing movement) to targets defined by the therapist – and doing so when wearing prism goggles which diverts the visual field 10° to the right (the patients are exhibiting a hemispatial neglect of the left hemispace). Normally, the feedback provided to the patient is the sight of the pointing finger at the moment when the pointing movement has been terminated. In most cases, the patient will gradually adapt to the perceptual shift and eventually show an after-effect in the form of a relative shift of the pointing movement – even after the removal of the goggles. In other words, the procedure constitutes an at least partial therapeutic intervention regarding the neglect of the left hemispace. It has been demonstrated that an essential element of the procedure is the feedback regarding the precision of the pointing movements during the training period (e.g., Frassinetti et al., 2002; Serino et al., 2006, 2007; Sarri et al., 2008). In the study by Wilms and Malá (2010) this traditional version of the PAT-procedure was included – and compared directly to a procedure in which the patients pointed to a touch-sensitive computer screen and feedback was provided graphically on the screen rather than via the direct sight of the pointing finger. Surprisingly, in both patients and normal subjects the two procedures differed significantly – the version in which an icon on the computer screen provided the feedback did not lead to any demonstrable after-effect.

Clinically, this specificity of the post-traumatic recovery poses a significant problem with respect to obtaining a therapeutic outcome, which generalizes to for instance the every-day-situations of the patient at home or at the work place. But with respect to studies addressing neural organization and reorganization – and for that matter the neural substrate of consciousness – this specificity may present both problems and promises.

The problems caused by the rather specific reorganizations provoked by a particular post-traumatic training process are primarily related to the (lack of) generality of the conclusions, which can be drawn from studies of a "functionally recovered" patient or animal. The pattern of neural mechanisms (e.g., regional activations seen in neuroimaging studies) and cognitive mechanisms allowing a more or less successful task solution cannot be seen as a more global indication of which parts of the brain are able to "take over" from those lost to injury. Instead, the observable pattern is the result of a specific process, which in principle is only designed to solve the manifestation of the task, which has been trained.

But exactly this specificity may also become a window through which much more detailed information can be obtained. If certain demands are fulfilled, a novel – more difficult and refined – but also more promising type of "localization" research may

be conducted. Scrutinizing the post-traumatic situation in both patients and experimental animals, such research may elucidate central aspects of the organization and post-traumatic reorganization of the brain – thereby providing a better insight into the neural mechanisms of cognition and consciousness.

And which demands are then to be met by such studies of post-traumatic symptomatology and recovery?

An essential aspect will be a more refined conceptualization of what constitutes a "function" and what is actually obtained during post-traumatic functional recovery. Replacing the old contradiction between localization and recovery of function with the more detailed concepts of for instance the EFs and ASs of the REF-model can provide a better framework for conceptualization of the results obtained.

But not only conceptually will progress have to be made. Also the methodology of such studies needs refinement. Realizing (like what is indicated by the REF-model) that rehabilitative training obtains a relatively task-specific reorganization rather than a recreation of what has been lost to injury, will have to provoke a different type of research strategy. A recovery process must be systematically addressed across various manifestations of what appear to be the same cognitive task as well as across cognitive domains. For each task it should also (as far as possible) be attempted to include studies utilizing various types of "organic" as well as behavioral/cognitive "challenges" – as has been suggested by Mogensen and Malá (2009) and Mogensen (2011). Such challenge procedures can provide more detailed insights into the pattern of neural activities mediating a recovered task performance as well as the cognitive strategies allowing a particular surface phenomenon to be achieved.

The plastic nature of the brain – and especially the regionally injured brain – can provide many surprises and frustrations to the neuroscientist trying to understand this most complicated product of nature. Apparently similar tasks provide dissimilar results post-traumatically, and patients who apparently "just" need a precise feedback regarding their pointing errors during rehabilitative training of hemispatial neglect respond differentially in case of a graphic representation on a screen or the actual sight of their finger. These and many more frustrations can in the short term become obstacles to clinical utilization of methods as well as the possibility to reach more global conclusions. But like the apparent contradiction between localization and recovery of function they are also the types of data, which we should cherish and on which we should focus our attention. Because here the dynamic brain is offering us windows through which we may eventually understand the neural mechanisms of cognition – and maybe consciousness.

ACKNOWLEDGMENT

The present study was supported by a grant from the Danish Research Council.

REFERENCES

Ansaldo, A. I., and Arguin, M. (2003). The recovery from aphasia depends on both the left and right hemispheres: three longitudinal case studies on the dynamics of language function after aphasia. *Brain Lang.* 87, 177–178.

Ansaldo, A. I., Arguin, M., and Lecours, A. R. (2002). The contribution of the right cerebral hemisphere to the recovery from aphasia: a single longitudinal case study. *Brain Lang.* 82, 206–222.

Arvidsson, A., Collin, T., Kirik, D., Kokaia, Z., and Lindvall, O. (2002). Neuronal replacement from endogenous precursors in the adult brain after stroke. *Nat. Med.* 8, 963–970.

Barrett, H. C., and Kurzban, R. (2006). Modularity in cognition: framing the debate. *Psychol. Rev.* 113, 628–647.

Bartolomeo, P. (2007). Visual neglect. *Curr. Opin. Neurol.* 20, 381–386.

Baumgaertner, A., Schraknepper, V., and Saur, D. (2005). Differential recovery of aphasia and apraxia of speech in an adolescent after infarction of the left frontal lobe: longitudinal behavioral and fMRI data. *Brain Lang.* 95, 211–212.

Berardi, N., Pizzorusso, T., and Maffei, L. (2004). Extracellular matrix and visual cortical plasticity: freeing the synapse. *Neuron* 44, 905–908.

Berry, M. (1982). Post-injury myelin-breakdown products inhibit axonal growth: an hypothesis to explain the failure of axonal regeneration in the mammalian central nervous system. *Bibl. Anat.* 23, 1–11.

Brose, K., Bland, K. S., Wang, K. H., Arnott, D., Henzel, W., Goodman, C. S., Tessier-Lavigne, M., and Kidd, T. (1999). Slit proteins bind Robo receptors and have an evolutionarily conserved role in repulsive axon guidance. *Cell* 96, 795–806.

Buller, D. J., and Hardcastle, V. G. (2000). Evolutionary psychology, meet developmental neurobiology: against promiscuous modularity. *Brain Mind* 1, 307–325.

Carney, N., Chesnut, R. M., Maynard, H., Mann, N. C., Patterson, P., and Helfand, M. (1999). Effect of cognitive rehabilitation on outcomes for persons with traumatic brain injury: a systematic review. *J. Head Trauma Rehabil.* 14, 277–307.

Chen, D. F., Jhaveri, S., and Schneider, G. E. (1995). Intrinsic changes in developing retinal neurons result in regenerative failure of their axons. *Proc. Natl. Acad. Sci. U.S.A.* 92, 7287–7291.

Chen, J., Magavi, S. S., and Macklis, J. D. (2004). Neurogenesis of corticospinal motor neurons extending spinal projections in adult mice. *Proc. Natl. Acad. Sci. U.S.A.* 101, 16357–16362.

Chen, M. S., Huber, A. B., Van Der Haar, M. E., Frank, M., Schnell, L., Spillmann, A. A., Christ, F., and Schwab, M. E. (2000). Nogo-A is a myelin-associated neurite outgrowth inhibitor and an antigen for monoclonal antibody IN-1. *Nature* 403, 434–439.

Choi, J. Y., Lee, K. H., Na, D. L., Byun, H. S., Lee, S. J., Kim, H., Kwon, M., Lee, K.-H., and Kim, B.-T. (2007). Subcortical aphasia after striatocapsular infarction: quantitative analysis of brain perfusion SPECT using statistical parametric mapping and a statistical probabilistic anatomic map. *J. Nucl. Med.* 48, 194–200.

Coltheart, M. (2001). "Assumptions and methods in cognitive neuropsychology," in *The Handbook of Cognitive Neuropsychology*, ed. B. Rapp (Philadelphia, PA: Psychology Press), 3–21.

Danckert, J., and Goodale, M. A. (2000). Blindsight: a conscious route to unconscious vision. *Curr. Biol.* 10, R64–R67.

Del Rio, J. A., and Soriano, E. (2007). Overcoming chondroitin sulphate proteoglycan inhibition of axon growth in the injured brain: lessons from chondroitinase ABC. *Curr. Pharmacol. Des.* 13, 2485–2492.

Delfs, J. M., Ciaramitaro, V. M., Parry, T. J., and Chesselet, M.-F. (1995). Subthalamic nucleus lesions: widespread effects on changes in gene expression induced by nigrostriatal dopamine depletion in rats. *J. Neurosci.* 15, 6562–6575.

Elbert, T., Pantev, C., Weinbruch, C., Rockstroh, B., and Taub, E. (1995). Increased cortical representation of the fingers of the left hand in string players. *Science* 270, 305–307.

Fawcett, J. W., Housden, E., Smith-Thomas, L., and Meyer, R. L. (1989). The growth of axons in three-dimensional astrocyte cultures. *Dev. Biol.* 135, 449–458.

Fodor, J. (1983). *The Modularity of Mind*. Cambridge, MA: MIT Press.

Fodor, J. (2000). *The Mind Doesn't Work that Way: The Scope and Limits of Computational Psychology*. Cambridge, MA: MIT Press.

Frassinetti, F., Angeli, V., Meneghello, F., Avanzi, S., and Ladavas, E. (2002). Long-lasting amelioration of visuospatial neglect by prism adaptation. *Brain* 125, 608–623.

Friston, K. J., and Price, C. J. (2003). Degeneracy and redundancy in cognitive anatomy. *Trends Cogn. Sci.* 7, 151–152.

Frost, S. B., Barbay, S., Friel, K. M., Plautz, E. J., and Nudo, R. J. (2003). Reorganization of remote cortical regions after ischemic brain injury: a potential substrate for stroke recovery. *J. Neurophysiol.* 89, 3205–3214.

Funk, J., Finke, K., Müller, H. J., Preger, R., and Kerkhoff, G. (2010). Systematic biases in the tactile perception of the subjective vertical in patients with unilateral neglect and the influence of upright vs. supine posture. *Neuropsychologia* 48, 298–308.

Goldberg, J. L., Klassen, M. P., Hua, Y., and Barres, B. A. (2002). Amacrine-signaled loss of intrinsic axon growth

ability by retinal ganglion cells. *Science* 296, 1860–1864.

Hagemann, G., Redecker, C., Neumann-Haefelin, T., Freund, H.-J., and Witte, O. W. (1998). Increased long-term potentiation in the surround of experimentally induced focal cortical infarction. *Ann. Neurol.* 44, 255–258.

Harel, N. Y., and Strittmatter, S. M. (2006). Can regenerating axons recapitulate developmental guidance during recovery from spinal cord injury? *Nat. Rev. Neurosci.* 7, 603–616.

Harris, L. K., Black, R. T., Golden, K. M., Reeves, T. M., Povlishock, J. T., and Phillips, L. L. (2001). Traumatic brain injury-induced changes in gene expression and functional activity of mitochondrial cytochrome *c* oxidase. *J. Neurotrauma* 18, 993–1009.

Hermann, D. M., Mies, G., and Hossmann, K.-A. (1999). Biochemical changes and gene expression following traumatic brain injury: role of spreading depression. *Restor. Neurol. Neurosci.* 14, 103–108.

Hiramoto, M., Hiromi, Y., Giniger, E., and Hotta, Y. (2000). The *Drosophila* Netrin receptor Frazzled guides axons by controlling Netrin distribution. *Nature* 406, 886–889.

Irvine, D. R. F. (2007). Auditory cortical plasticity: does it provide evidence for cognitive processing in the auditory cortex? *Hear. Res.* 229, 158–170.

Karl, A., Birbaumer, N., Lutzenberger, W., Cohen, L. G., and Flor, H. (2001). Reorganization of motor and somatosensory cortex in upper extremity amputees with phantom limb pain. *J. Neurosci.* 15, 3609–3618.

Kentridge, R. W., Heywood, C. A., and Weiskrantz, L. (1999). Attention without awareness in blindsight. *Proc. R. Soc. Lond. B Biol. Sci.* 266, 1805–1811.

Keynes, R. J., and Cook, G. M. (1992). Repellent cues in axon guidance. *Curr. Opin. Neurobiol.* 2, 55–59.

Kleim, J. A., Jones, T. A., and Schallert, T. (2003). Motor enrichment and the induction of plasticity before or after brain injury? *Neurochem. Res.* 28, 1737–1769.

Koeberle, P. D., and Bahr, M. (2004). Growth and guidance cues for regenerating axons: where have they gone? *J. Neurobiol.* 59, 162–180.

Krechevsky, I. (1932). "Hypotheses" in rats. *Psychol. Rev.* 39, 516–532.

Krechevsky, I. (1933). Hereditary nature of "hypotheses". *J. Comp. Psychol.* 16, 99–116.

Kringelbach, M. L., and Rolls, E. T. (2004). The functional neuroanatomy of the human orbitofrontal cortex: evidence from neuroimaging and neuropsychology. *Prog. Neurobiol.* 7, 341–372.

Lane, R. D., Ahern, G. L., Schwartz, G. E., and Kaszniak, A. W. (1997). Is alexithymia the emotional equivalent of blindsight? *Biol. Psychiatry* 42, 834–844.

Leavitt, B. R., Hernit-Grant, C. S., and Macklis, J. D. (1999). Mature astrocytes transform into transitional radial glia within adult mouse neocortex that supports directed migration of transplanted immature neurons. *Exp. Neurol.* 157, 43–57.

León-Carrión, J., and Machuca-Murga, F. (2001). Spontaneous recovery of cognitive functions after severe brain injury: when are neurocognitive sequelae established? *Rev. Esp. Neuropsicol.* 3, 58–67.

Lepore, F., Ptito, M., Provencal, C., Bedard, S., and Guillemot, J.-P. (1985). Interhemispheric transfer of visual training in the split-brain cat: effects of the experimental set-up. *Can. J. Psychol.* 37, 527.

Magavi, S. S., Leavitt, B. R., and Macklis, J. D. (2000). Induction of neurogenesis in the neocortex of adult mice. *Nature* 405, 951–955.

Marcel, A. J. (1998). Blindsight and shape perception: deficit of visual consciousness or of visual function? *Brain* 121, 1565–1588.

McGee, A. W., Yang, Y., Fischer, Q. S., Daw, N. W., and Strittmatter, S. M. (2005). Experience-driven plasticity of visual cortex limited by myelin and Nogo receptor. *Science* 309, 2222–2226.

Meinzer, M., Flaisch, T., Breitenstein, C., Wienbruch, C., Elbert, T., and Rockstroh, B. (2008). Functional re-recruitment of dysfunctional brain areas predicts language recovery in chronic aphasia. *Neuroimage* 39, 2038–2046.

Merzenich, M. M., and Jenkins, W. M. (1993). Reorganization of cortical representations of the hand following alterations of skin inputs induced by nerve injury, skin island transfers, and experience. *J. Hand Therapy* 6, 89–104.

Mogensen, J. (2003). "Animal models in neuroscience," in *Handbook of Laboratory Animal Science*, 2nd Edn, Vol. II. *Animal Models*, eds J. Hau and G. L. van Hoosier (Boca Raton, FL: CRC Press LLC), 95–109.

Mogensen, J. (2011). "Animal models in neuroscience," in *Handbook of Laboratory Animal Science*, 3rd Edn, Vol. II. *Animal Models*, eds J. Hau and S. Schapiro (Boca Raton, FL: CRC Press LLC) (in press).

Mogensen, J., Boyd, M. H., Nielsen, M. D., Kristensen, R. S., and Malá, H. (2008). Erythropoietin improves spatial delayed alternation in a T-maze in rats subjected to ablation of the prefrontal cortex. *Brain Res. Bull.* 77, 1–7.

Mogensen, J., Christensen, L. H., Johansson, A., Wörtwein, G., Bang, L. E., and Holm, S. (2002). Place learning in scopolamine treated rats: the roles of distal cues and catecholaminergic mediation. *Neurobiol. Learn. Mem.* 78, 139–166.

Mogensen, J., Hjortkjær, J., Ibervang, K. L., Stedal, K., and Malá, H. (2007). Prefrontal cortex and hippocampus in posttraumatic functional recovery: spatial delayed alternation by rats subjected to transection of the fimbria–fornix and/or ablation of the prefrontal cortex. *Brain Res. Bull.* 73, 86–95.

Mogensen, J., Iversen, I. H., and Divac, I. (1987). Neostriatal lesions impaired rats' delayed alternation performance in a T-maze but not in a two-key operant chamber. *Acta Neurobiol. Exp.* 47, 45–54.

Mogensen, J., Lauritsen, K. T., Elvertorp, S., Hasman, A., Moustgaard, A., and Wörtwein, G. (2004). Place learning and object recognition by rats subjected to transection of the fimbria–fornix and/or ablation of the prefrontal cortex. *Brain Res. Bull.* 63, 217–236.

Mogensen, J., and Malá, H. (2009). Post-traumatic functional recovery and reorganization in animal models. A theoretical and methodological challenge. *Scand. J. Psychol.* 50, 561–573.

Mogensen, J., Moustgaard, A., Khan, U., Wörtwein, G., and Nielsen, K. S. (2005). Egocentric spatial orientation in a water maze by rats subjected to transection of the fimbria–fornix and/or ablation of the prefrontal cortex. *Brain Res. Bull.* 65, 41–58.

Mogensen, J., Wörtwein, G., Plenge, P., and Mellerup, E. T. (2003). Serotonin, locomotion, exploration, and place recall in the rat. *Pharmacol. Biochem. Behav.* 75, 381–395.

Monakow, C. V. (1914). *Die Lokalisation im Grosshirn und der Abbau der Funktion durch Kortikale Herde*. Wiesbaden: Bergmann.

Münte, T. F., Altenmüller, E., and Jäncke, L. (2002). The musician's brain as a model of neuroplasticity. *Nat. Rev. Neurosci.* 3, 473–478.

Nakatomi, H., Kuriu, T., Okabe, S., Yamamoto, S.-C., Hatano, O., Kawahara, N., Tamura, A., Kirino, T., and Nakafuku, M. (2002). Regeneration of hippocampal pyramidal neurons after ischemic brain injury by recruitment of endogenous neural progenitors. *Cell* 110, 429–441.

Natsoulas, T. (1997). Blindsight and consciousness. *Am. J. Psychol.* 110, 1–33.

Norman, D. A., and Shallice, T. (1986). "Attention to action: willed and automatic control of behavior," in *Consciousness and Self-Regulation*, Vol. 4, eds R. J. Davidson, G. E. Schwartz, and D. Shapiro (New York: Plenum Press), 1–18.

Nudo, R. J. (2003). Adaptive plasticity in motor cortex: implications for rehabilitation after brain injury. *J. Rehab. Med.* 41(Suppl.), 7–10.

Olton, D. S. (1978). "The function of septo-hippocampal connections in spatially organized behaviour," in *Functions of the Septo-Hippocampal System*, Ciba Foundation Symposium 58 (New York: Elsevier), 327–342.

Overgaard, M., and Mogensen, J. (2011). A framework for the study of multiple realizations: the importance of levels of analysis. *Front. Conscious. Res.* (in press).

Panksepp, J., and Panksepp, J. B. (2000). The seven sins of evolutionary psychology. *Evol. Cogn.* 6, 108–131.

Pascual-Leone, A., Bartres-Faz, D., and Keenan, J. P. (1999). Transcranial magnetic stimulation: studying the brain–behaviour relationship by induction of "virtual lesions". *Philos. Trans. R. Soc. Lond. B* 354, 1229–1238.

Pascual-Leone, A., Walsh, V., and Rothwell, J. (2000). Transcranial magnetic stimulation in cognitive neuroscience – virtual lesion, chronometry, and functional connectivity. *Curr. Opin. Neurobiol.* 10, 232–237.

Perani, D., Cappa, S. F., Tettamanti, M., Rosa, M., Scifo, P., Miozzo, A., Basso, A., and Fazio, F. (2003). A fMRI study of word retrieval in aphasia. *Brain Lang.* 85, 357–368.

Pizzorusso, T., Medini, P., Berardi, N., Chierzi, S., Fawcett, J. W., and Maffei, L. (2002). Reactivation of ocular dominance plasticity in the adult visual cortex. *Science* 298, 1248–1251.

Polleux, F., Morrow, T., and Ghosh, A. (2000). Semaphorin 3A is a chemoattractant for cortical apical dendrites. *Nature* 404, 567–573.

Price, C. J., and Friston, K. J. (2002). Degeneracy and cognitive anatomy. *Trends Cogn. Sci.* 6, 416–421.

Ptito, M., Moesgaard, S. M., Gjedde, A., and Kupers, R. (2005). Cross-modal plasticity revealed by electrotactile stimulation of the tongue in the congenitally blind. *Brain* 128, 606–614.

Rakic, P. (1971). Guidance of neurons migrating to the fetal monkey neocortex. *Brain Res.* 33, 471–476.

Rakic, P. (1985). "Mechanisms of neuronal migration in developing cerebellar cortex," in *Molecular Basis of Neural Development*, eds G. M. Edelman, W. M. Cowan, and E. Gull (New York: Wiley), 139–160.

Ramachandran, V. S., and Blakeslee, S. (1998). *Phantoms in the Brain: Probing the Mysteries of the Human Mind*. New York: William Morrow.

Rao, V. L. R., Dhodda, V. K., Song, G., Bowen, K. K., and Dempsey, R. J. (2003). Traumatic brain injury-induced acute gene expression changes in rat cerebral cortex identified by GeneChip analysis. *J. Neurosci. Res.* 71, 208–219.

Recanzone, G. H., Schreiner, C. E., and Merzenich, M. M. (1993). Plasticity in the frequency representation of primary auditory cortex following discrimination training in adult owl monkeys. *J. Neurosci.* 13, 87–103.

Robertson, D., and Irvine, D. R. F. (1989). Plasticity of frequency organization in auditory cortex of guinea pigs with partial unilateral deafness. *J. Comp. Neurol.* 282, 456–471.

Rohling, M. L., Faust, M. E., Beverly, B., and Demakis, G. (2009). Effectiveness of cognitive rehabilitation following acquired brain injury: a meta-analytic re-examination of Cicerone et al.'s (2000, 2005) systematic reviews. *Neuropsychology* 23, 20–39.

Rossetti, Y., Rode, G., Pisella, L., Farne, A., Li, L., Boisson, D., and Perenin, M.-T. (1998). Prism adaptation to a rightward optical deviation rehabilitates left hemispatial neglect. *Nature* 395, 166–169.

Rumelhart, D., and McClelland, J. (1986). *Parallel Distributed Processing*. Cambridge, MA: MIT Press.

Sarri, M., Greenwood, R., Kalra, L., Papps, B., Husain, M., and Driver, J. (2008). Prism adaptation aftereffects in stroke patients with spatial neglect: pathological effects on subjective straight ahead but not visual open-loop pointing. *Neuropsychologia* 46, 1069–1080.

Schäfer, R., Dehn, D., Burbach, G. J., and Deller, T. (2008). Differential regulation of chondroitin sulfate proteoglycan mRNAs in the denervated rat fascia dentata after unilateral entorhinal cortex lesion. *Neurosci. Lett.* 439, 61–69.

Scharff, C., Kirn, J. R., Grossman, M., Macklis, J. D., and Nottebohm, G. (2000). Targeted neuronal death affects neuronal replacement and vocal behavior in adult songbirds. *Neuron* 25, 481–492.

Scheich, H. (1991). Auditory cortex: comparative aspects of maps and plasticity. *Curr. Opin. Neurobiol.* 1, 236–247.

Schindler, I., Clavagnier, S., Karnath, H. O., Derex, L., and Perenin, M.-T. (2006). A common basis for visual and tactile exploration deficits in spatial neglect? *Neuropsychologia* 44, 1444–1451.

Schwab, M. E., and Thoenen, H. (1985). Dissociated neurons regenerate into sciatic but not optic nerve explants in culture irrespective of neurotrophic factors. *J. Neurosci.* 5, 2415–2423.

Selnes, O. A. (2001). "A historical overview of contributions from the study of deficits," in *The Handbook of Cognitive Neuropsychology*, ed. B. Rapp (Philadelphia, PA: Psychology Press) 23–41.

Serino, A., Angeli, V., Frassinetti, F., and Ladavas, E. (2006). Mechanisms underlying neglect recovery after prism adaptation. *Neuropsychologia* 44, 1068–1078.

Serino, A., Bonifazi, S., Pierfederici, L., and Ladavas, E. (2007). Neglect treatment by prism adaptation: what recovers and for how long. *Neuropsychol. Rehabil.* 17, 657–687.

Silvanto, J., Cowey, A., and Walsh, V. (2008). Inducing conscious perception of colour in blindsight. *Curr. Biol.* 18, R950–R951.

Specht, K., Zahn, R., Willmes, K., Weis, S., Holtel, C., Krause, B. J., Herzog, H., and Huber, W. (2009). Joint independent component analysis of structural and functional images reveals complex patterns of functional reorganisation in stroke aphasia. *Neuroimage* 47, 2057–2063.

Szaflarski, J. P., Eaton, K., Ball, A. L., Banks, C., Vannest, J., Allendorfer, J. B., Page, S., and Holland, S. K. (2010). Poststroke aphasia recovery assessed with functional magnetic resonance imaging and a picture identification task. J. *Stroke Cerebrovasc. Dis.* 1–10. (published online August 2010).

Thai-Van, H., Micheyl, C., Norena, A., Veuillet, E., Gabriel, D., and Collet, L. (2007). Enhanced frequency discrimination in hearing-impaired individuals: a review of perceptual correlates of central neural plasticity induced by cochlear damage. *Hear. Res.* 233, 14–22.

Thomas, C., Altenmüller, E., Marckmann, G., Kahrs, J., and Dichgans, J. (1997). Language processing in aphasia: changes in lateralization patterns during recovery reflect cerebral plasticity in adults. *Electroencephalogr. Clin. Neurophysiol.* 102, 86–97.

Thulborn, K. R., Carpenter, P. A., and Just, M. A. (1999). Plasticity of language-related brain function during recovery from stroke. *Stroke* 30, 749–754.

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

Weiss, T., Miltner, W. H. R., Huonker, R., Friedel, R., Schmidt, I., and Taub, E. (2000). Rapid functional plasticity of the somatosensory cortex after finger amputation. *Exp. Brain Res.* 134, 199–203.

Werbos, P. J. (1994). *The Roots of Backpropagation: From Ordered Derivatives to Neural Networks and Political Forecasting*. New York: John Wiley & Sons.

Wilms, I., and Malá, H. (2010). Indirect versus direct feedback in computer-based Prism Adaptation Therapy. *Neuropsychol. Rehabil.* 20, 830–853.

Witte, O. W. (1998). Lesion-induced plasticity as a potential mechanism for recovery and rehabilitative training. *Curr. Opin. Neurol.* 11, 655–662.

Xerri, C., Coq, J., Merzenich, M., and Jenkins, W. (1996). Experience-induced plasticity of cutaneous maps in the primary somatosensory cortex of adult monkeys and rats. *J. Physiol.* 90, 277–287.

Yang, T. T., Gallen, C. C., Ramachandran, V. S., Cobb, S., Schwartz, B. J., and Bloom, F. E. (1994). Noninvasive detection of cerebral plasticity in adult human somatosensory cortex. *Neuroreport* 5, 701–704.

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

Received: 03 October 2010; accepted: 05 January 2011; published online: 26 January 2011.

Citation: Mogensen J (2011) Reorganization of the injured brain: implications for studies of the neural substrate of cognition. Front. Psychology ***2****:7. doi: 10.3389/fpsyg.2011.00007*

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

frontiers in PSYCHOLOGY

HYPOTHESIS AND THEORY ARTICLE
published: 27 April 2011
doi: 10.3389/fpsyg.2011.00079

A framework for the study of multiple realizations: the importance of levels of analysis

***Morten Overgaard*[1,2]* *and Jesper Mogensen*[3]**

[1] *CNRU, Department of Psychology and Communication, Aalborg University, Aalborg, Denmark*
[2] *Cognitive Neuroscience Research Unit, Hammel Neurorehabilitation and Research Center, MindLab, Aarhus University, Aarhus, Denmark*
[3] *The Unit for Cognitive Neuroscience, Department of Psychology, University of Copenhagen, Copenhagen, Denmark*

Edited by:
Mads Jensen, Aarhus University Hospital, Denmark

Reviewed by:
Alexander Maier, National Institute for Mental Health, USA
John Bickle, Mississippi State University, USA

***Correspondence:**
Morten Overgaard, Cognitive Neuroscience Research Unit, Hammel Neurorehabilitation and Research Center, MindLab, Aarhus University, Noerrebrogade 44, building 10G, DK-8000 Aarhus, Denmark.
e-mail: mortover@rm.dk

The brain may undergo functional reorganizations. Selective loss of sensory input or training within a restricted part of a modality cause "shifts" within for instance somatotopic or tonotopic maps. Cross-modal plasticity occurs when input within a modality is absent – e.g., in the congenitally blind. Reorganizations are also found in functional recovery after brain injury. Focusing on such reorganizations, it may be studied whether a cognitive or conscious process can exclusively be mediated by one neural substrate – or may be associated with multiple neural representations. This is typically known as the problem of multiple realization – an essentially empirical issue with wide theoretical implications. This issue may appear to have a simple solution. When, for instance, the symptoms associated with brain injury disappear and the recovery is associated with increased activities within spared regions of the brain, it is tempting to conclude that the processes originally associated with the injured part of the brain are now mediated by an alternative neural substrate. Such a conclusion is, however, not a simple matter. Without a more thorough analysis, it cannot be concluded that a functional recovery of for instance language or attention is necessarily associated with a novel representation of the processes lost to injury. Alternatively, for instance, the recovery may reflect that apparently similar surface phenomena are obtained via dissimilar cognitive mechanisms. In this paper we propose a theoretical framework, which we believe can guide the design and interpretations of studies of post-traumatic recovery. It is essential to distinguish between a number of levels of analysis – including a differentiation between the surface phenomena and the underlying information processing – when addressing, for instance, whether a pre-traumatic and post-traumatically recovered cognitive or conscious process are actually the same. We propose a (somewhat preliminary) system of levels of analysis, which can be applied to such studies.

Keywords: neural substrate, cognition, consciousness, plasticity, brain injury, reorganization, levels of analysis, neural correlate of consciousness

INTRODUCTION

One of the primary concerns in the attempt to isolate some neural correlate of consciousness (NCC) is the question whether conscious content relates to brain processes in a fixed 1:1 pattern, or whether the same content may be realized in different brain processes. The answer to this question will have important theoretical consequences: Should the latter view be correct, it would seem futile to reduce or identify conscious content with one specific brain process. Instead, such a finding would support theories that allow a "one-to-many" relationship.

The experimental literature is already rich with examples showing that individual connections in the brain are constantly being modified, largely dependent upon how they are used (e.g., Keller-Peck et al., 2001; Lee et al., 2006, 2008). Decades ago, it was shown that if the cortical organization for a sensory system is deprived of input, it will become activated at a later time in response to other, usually adjacent inputs (e.g., Yang et al., 1994; Weiss et al., 2000; Karl et al., 2001). Such reorganization seems to be found with any chosen method of analysis, be it changes in individual cells, networks of cells, or entire "modules."

However well-established the phenomenon of brain plasticity may be, the question whether we should conceive of mind–brain relations as 1:1 or 1:X remains unanswered. We believe this lack of success is mainly due to the lack of a proper conceptual and methodological framework to approach the issue: It seems difficult to decide whether two occurrences of a mental state or a brain state are identical or not. This decision must, however, be made in order to answer the over-all question about the relationship between the two kinds of states.

This paper aims to take the first steps toward such a framework, outlining an approach to analyze the problem, and giving examples of the issues being considered.

BACKGROUND

There is a constantly growing number of studies demonstrating that when a particular sensory input is lacking (due to peripheral loss – but without injury to the brain) a series of short- and long-term reorganizations take place within various levels of the brain systems representing the affected modality. Such reorganizations occur, for instance, within the somatotopic representations

of various body-parts (e.g., Yang et al., 1994; Weiss et al., 2000; Karl et al., 2001) and the tonotopic representations of the auditory system (e.g., Robertson and Irvine, 1989; Scheich, 1991; Irvine, 2007; Thai-Van et al., 2007).

Individuals, who lost one hand to amputation (thereby depriving the somatosensory cortex of the contralateral hemisphere of normal input to the region specialized for processing hand-related information), have been studied extensively (e.g., Yang et al., 1994; Weiss et al., 2000; Karl et al., 2001). In such persons, the region of the primary somatosensory cortex, which used to represent the now amputated hand, does not remain inactive. Instead, the neighboring somatosensory regions (representing the arm and face, respectively) encroach on the now "vacant" area in such a way that the "input-lacking" cortical hand area is eventually fully taken over by inputs from arm and face, respectively. Plastic reorganizations of the somatosensory cortical map do, however, not only occur under conditions of a lacking input from parts of the body. Even changed experiences in terms of intensive training restricted to part of the body may be associated with a relative shift within somatosensory representations (e.g., Merzenich and Jenkins, 1993; Elbert et al., 1995; Xerri et al., 1996; Münte et al., 2002).

Showing that these plastic processes within the somatosensory system are not unique, the tonotopic representations of the auditory system appear also to undergo plastic changes due to changes in input or experience. Partial loss of input (e.g., due to restricted cochlear lesions) deprives the tonotopic representation of input within a restricted spectrum of frequencies. Upon such a manipulation, the representations of the neighboring frequencies expand to fill the "vacant" region of for instance the auditory cortex (e.g., Robertson and Irvine, 1989; Scheich, 1991; Irvine, 2007; Thai-Van et al., 2007). Within the frequencies, which have – due to such mechanisms – gained an increased neural representation, the quality of frequency discrimination appear to be enhanced (Thai-Van et al., 2007). Learning experiences specific to a restricted tonal representation are also able to induce an increased representation of the frequencies in question (e.g., Scheich, 1991; Recanzone et al., 1993; Irvine, 2007) – and this increased representation within the tonotopic representation seems to be correlated with the quality of performance of the auditory discrimination task (e.g., Recanzone et al., 1993).

While the above examples deal with plastic reorganizations within one modality, there are also examples demonstrating the even more radical phenomenon of cross-modal plasticity. For instance, such a phenomenon has been found in the context of attempts to provide the blind with a "visual prosthesis." The idea of such research is to provide a – presently crude but nevertheless potentially useful – representation of the visual world via a somatosensory input. Captured by a camera, the image of the environment is transformed into a crude – "low-resolution" – representation in the form of a matrix with "active" and "inactive" pixels. This matrix is presented to the blind subject in the form of somatosensory stimulation. The matrix is brought into contact with a part of the body surface and "active" pixels are designated by for instance vibration or mild electric stimulation (while "inactive" pixels lack such a stimulation). In a number of these experiments the inputs have been simple geometric patterns – for instance letters of various orientations (e.g., Bach-y-Rita et al., 1969, 1998; Kaczmarek et al., 1991; Ptito et al., 2005). It has been demonstrated by Ptito et al. (2005) that both blind and normally sighted (but blindfolded) individuals can learn to utilize this device in solving discrimination tasks in which they are required to determine the orientation of letters. While the ability to acquire the task was similar in both groups, the neural regions activated during task performance turned out to differ. In both groups somatosensory cortical regions were activated but in the blind group (in contrast to the sighted individuals) additional activations were found in cortical regions normally associated with the processing of visual information (Ptito et al., 2005). In contrast, the sighted but blindfolded subjects demonstrated a cortical activation pattern similar to what would be expected when performing an ordinary tactile discrimination task (Ptito et al., 2005). It may be especially relevant that in the blind subjects, successful task solution was associated with activation of cortical regions, which in sighted individuals participates in mediation of visual tasks in which the spatial orientation of figures is to be determined. It should be noted that Ptito et al. (2005) studied individuals who were "early blind."

The most commonly studied aspect of cross-modal plasticity in the blind is the involvement of visual cortical areas in the performance of Braille reading (e.g., Gizewski et al., 2003; Ptito et al., 2008). There have been some doubts regarding the specificity of the involvement of visual areas in tactile discrimination tasks related to Braille reading (Gizewski et al., 2003). However, in a study directly addressing this issue, Gizewski et al. (2003) found that in early blind subjects somatosensory activation of visual areas was exclusively related to the performance of tactile discrimination tasks (including Braille reading) – but not elicited by electrical stimulation of the hand used in the performance these tasks.

The main focus of studies such as those described above is the ability of the brain to reorganize in ways that allow a structure or substructure to become functionally engaged within a domain, which is more or less dissimilar to the type of information processing, originally managed by the circuitry in question. Regarding shifts of somatotopic or tonotopic maps, the plasticity-induced novelty concerns different aspects within the same sensory modality, while – in cross-modal plasticity – other cases even demonstrate a shift to processing sensory information of another modality than the one to which the brain region is supposedly originally "programmed."

Normally, such results are taken to indicate that a given brain structure is able to subserve multiple "functions" – at least within the restrictions indicated by various studies (cross-modal changes may, for instance, most easily or maybe exclusively be achieved in case of early deprivation of a sensory modality, e.g., Ptito et al., 2005).

When seen from another angle, the above-described examples can also be seen as indications that the same "function" can be mediated by multiple neural substrates. This may not commonly be the focus of such studies, but when, for instance, an area previously specialized in processing information regarding a hand is able also to mediate face-related information processing, this may also be framed in terms of a "relocation" of face-related information processing. More clear-cut cases of an apparent "relocation" of functional mediation are, however, to be found within the literature dealing with the neural substrate of post-traumatic functional recovery.

It must, however, be remembered that theoretically an apparent "relocalization" of functional mediation (as a consequence of either a modified input to an intact brain or regional brain injury) might

be accompanied by a "rewiring" of the local circuitry within the novel substrate of functional mediation. If the neural plasticity mediating the "relocalization" not only changes the input/output properties of the novel structure mediating a function, but also rewires the internal networks of this brain region, it could be argued that the entire process includes the creation of an actual copy of the neural network mediating the function within its original location. If this can be the case, a 1:1 relationship between function and neural processing could be preserved while the site of the mediating circuitry has just shifted to a novel structure.

Thus, it is important to consider how likely it is to find a (re) establishment of the basic circuitry of brain regions in the adult brain. This issue is pursued in more detail by Mogensen (2011a,c). During maturation, neurons undergo a number of changes making them less similar to the developing neurons, which originally formed the local circuits of the brain (e.g., Fawcett et al., 1989; Chen et al., 1995; Goldberg et al., 2002). These restrictions may, however, be somewhat outweighed by the ongoing neurogenesis in the adult brain – a neurogenesis which is further potentiated when the brain is hit by injury (e.g., Magavi et al., 2000; Scharff et al., 2000; Arvidsson et al., 2002; Nakatomi et al., 2002; Chen et al., 2004). These newly formed neurons may in the injured brain even receive some support in reaching relevant brain regions – upon injury, mature astrocytes are able to transform themselves into radial glial cells similar to those guiding neural migration (and thereby the arrival of newly formed neurons in the relevant brain structures) during development (e.g., Rakic, 1971, 1985; Leavitt et al., 1999). There may, however, be at least one crucial factor preventing the – uninjured or injured – adult brain from recreation of networks similar to those formed during development: from the final stages of the original ontogenic development (and thereby formation of the basic circuitry of the brain) a number of factors associated with glial cells and myelin (e.g., Berry, 1982; Schwab and Thoenen, 1985; Schäfer et al., 2008) appear to prevent restructuring and presumably recreation of such basic circuitry. Especially important may be the astrocyte-produced chondroitin sulfate proteoglycans (CSPGs). These substances play an important role in the termination of the developmentally "critical periods" and thereby "consolidate" the originally formed circuitry in its "final" form (e.g., Pizzorusso et al., 2002; Berardi et al., 2004; McGee et al., 2005; Del Rio and Soriano, 2007; Schäfer et al., 2008). And the presence of these CSPGs apparently blocks an adult recreation of traumatically lost networks (e.g., Del Rio and Soriano, 2007; Schäfer et al., 2008).

Maybe one of the cognitive domains, which have been most extensively studied with respect to the mechanisms of post-traumatic functional recovery, is language. Numerous studies have scrutinized (so far without reaching consensus) the neural substrate of post-traumatic functional recovery of patients suffering various types of aphasia due to either stroke or traumatic brain injury (TBI).

Most patients suffering post-traumatic aphasia regain at least a partial proficiency of language, and such a recovery process is accompanied by reorganizational processes within various parts of the brain (e.g., Marsh and Hillis, 2006). These reorganizational processes may be affected by the type of traumatic impact, itself. For instance, gliomas may be removed in either a one- or two-step surgical procedure. And compared to the more massive impact of the single surgical intervention, the two-step procedure leads to a different type of post-traumatic reorganization (e.g., Duffau et al., 2002). Numerous methods have been employed in demonstrating the post-recovery substrate of language processing in patients suffering aphasia – some of the most common methods being imaging techniques (e.g., Perani et al., 2003) and physiological methods such as event-related potentials (e.g., Laganaro et al., 2008).

Aphasias primarily result from injury to the left hemisphere and the contribution of ipsilateral (left hemisphere) mediation of the reacquired linguistic functions have been demonstrated by for instance Szaflarski et al. (2010), Perani et al. (2003), Specht et al. (2009), and Meinzer et al. (2008) – the latter study demonstrating treatment-induced reintegration of various perilesional areas. The most common question asked within this field is, however, whether the contralateral, right, hemisphere plays a significant role in mediating post-traumatic recovery of language. The earliest case in which it was presumed that post-traumatic recovery upon expressive aphasia (also known as non-fluent aphasia – a condition characterized by difficulty producing speech on the background of a relatively preserved comprehension of language) was mediated by mechanisms within the right hemisphere (in this case the mirror-structure of Broca's area), was published by Barlow (1877). Doubts about this case remain, and Finger et al. (2003) concluded that we still know too little to fully interpret cases such as this one. However, without going into the details of whether a case like Barlow's manifest a true case of "vicariation" (the phenomenon that brain areas with different functions can sometimes assume or "take over" the function of an injured brain region – e.g., Finger and Stein, 1982; Slavin et al., 1988) it is clear that numerous studies – utilizing a variety of methods – have found an apparent involvement of structures within the right hemisphere in mediation of post-traumatic recovery in aphasia (e.g., Thomas et al., 1997; Thulborn et al., 1999; Ansaldo et al., 2002; Ansaldo and Arguin, 2003; Perani et al., 2003; Baumgaertner et al., 2005; Specht et al., 2009). In some cases (e.g., Thomas et al., 1997) there are indications that the patterns of shift toward a right hemisphere mediation of linguistic functions differ between types of aphasia. Furthermore, changes in the direction of language being mediated by the right hemisphere are often accompanied by internal reorganizations within the left hemisphere (where linguistic mediation may shift to uninjured regions). These reorganizations may lead to a more bilateral representation of language – due to the concurrent shift of linguistic mediation within the injured left hemisphere and to the contralateral, right hemisphere (e.g., Thompson et al., 2010). In most instances, the recovery-associated shifts toward right hemisphere mediation of language occur without the rehabilitative training specifically aiming at such a shift (but rather working in a more general way toward a recovery of the lost linguistic abilities). There are, however, cases in which (partly successful) training of aphasic patients have aimed at achieving a higher degree of right hemisphere mediation of linguistic tasks (e.g., Crosson et al., 2009 – who used a manipulation task performed with the patient's left hand to initiate naming trials and thereby obtain an independent right hemisphere activation which presumably can ease an interhemispheric shift of linguistic task mediation). It can often correctly be questioned to what extend the changes in neural activity observed via for instance fMRI in a recovering or recovered aphasic patient are specifically related to the reacquisition of the

language. A post-traumatic change in activity within a given structure may be the consequence of any trauma-related process – e.g., "disinhibition" due to lack of input from the injured brain region. Or other processes which are not directly related to the recovery of language. Certain cases are, however, highly informative in this respect. Meinzer et al. (2007), for instance, studied the linguistic recovery of a bilingual aphasic patient. In this case activation of parts of the superior temporal lobe of the right hemisphere was exclusively associated with use of the trained language while no such activation was observed associated with the untrained language.

To summarize, post-traumatic functional recovery upon brain injury – such as in the instances described above – may demonstrate that performance of the recovered task is associated with activity within brain regions, which were not pretraumatically involved in mediation of such tasks. Such results are normally taken to indicate that in one way or another a task mediation originally achieved by the injured part of the brain is now accomplished by neural mechanisms not previously involved in these functional domains.

In philosophy of mind, Block and Fodor (1972) used plastic properties of the brain as evidence for multiple realization and as an argument for functionalism. However, all findings mentioned above would be fully consistent with the proposal that whereas the post-traumatic, recovered function, e.g., the ability to speak, has superficial similarities with the pretraumatic function, the exact language structure has changed according to the change in neural substrate. What appear to be similar surface manifestations of a cognitive process may – when further scrutinized – turn out to be more dissimilar than originally assumed. For instance, a post-traumatically recovered task performance may have allowed the injured individual to regain a full proficiency (to the level of what is achieved with an uninjured brain) but this recovery is nevertheless achieved via cognitive mechanisms which upon further scrutiny turns out to be dissimilar to those of normal individuals (e.g., Mogensen et al., 2004). The need for a more thorough analysis of the surface phenomena observed during functional recovery after brain injury has been both methodologically and theoretically discussed elsewhere (e.g., Mogensen and Malá, 2009; Mogensen, 2011a,b,c; Wilms and Mogensen, in press).

At the same time, however, and as illustrated in the example of aphasia, it seems one may be able to argue for *and* against multiple realization, embracing the view that while the ability to speak, as such, is re-established in a different way in the brain, the exact same way of speaking, when analyzed in all details, is forever lost. That is: while viewed superficially the task performance of the patient demonstrates that general linguistic abilities are reacquired, but a more detailed analysis of the post-traumatic cognitive processes reveals that the post-traumatically established linguistic processes are dissimilar to those seen pretraumatically. This, in and of itself, reveals, at least, two unanswered questions underlying the debate: In which sense can two instances of a mental phenomenon be said to be the same, and with which methods could such an identity be established?

Although such questions appear directly from an analysis of empirical research, they have been debated for decades in philosophy of mind. Notably, Shapiro (2004, 2008) discusses difficulties in considering multiple realization an immediately testable empirical hypothesis. Shapiro (2008) asks the reader to consider three watches: Two analog ones and one digital. The two analog watches differ in timekeeping properties in that one runs fast and the other late. Nevertheless, the difference can be explained by one respect – a spring in one of the watches is longer than in the other. Shapiro argues that timekeeping is realized in the same way in these two watches although they have slightly different timekeeping properties. Timekeeping, nevertheless, is realized differently in the digital watch, he claims. If one of the analog watches had the exact identical timekeeping properties as the digital one, despite the difference in physical properties, it would be evidence for multiple realization. So, the example challenges evidence for multiple realization in the brain: Are the physical differences, relevant to realization, only relevant for that realized property? If so, it is not a case of "true multiple realization."

Bechtel and Mundale (1999) argue that one central issue is that psychological as well as neural properties can be described with different "granularities." They argue that psychological properties often are described at a very "coarse" level of granularity when deciding whether two mental states are identical or not, whereas neural properties are described at a much "finer" level. In order to make a comparison between two mental states, and thus conclude anything about whether some observation is a case of multiple realization, it must first be resolved at which level, mental states are properly described. In **Table 1** we provide a tentative framework for levels of analysis relevant to this issue.

LEVELS OF ANALYSIS IN COGNITIVE PROCESSES

To begin with, mental states may be compared at different levels. While the term "mental states" in itself is difficult to delineate and define, the first part of the paper will look at mental states as discussed in cognitive science: Functions which may be described as states or processes inferred from behavior.

Looking at cognitive states, we will tentatively suggest at least such three levels of analysis where a mental state may be defined and compared to other states.

At the most general level, it is a relatively simple task to determine whether or not an entire domain is intact, lost or fully recovered. Here, the question is simply if there is visual perception, language comprehension, or any such over-all function present. A person may loose her ability to understand language following an injury to the brain, but recover this ability. The ability may be only partially recovered, understanding just some words, but at this general level it is the same function as before. At this level, the presence or

Table 1 | Levels of analysis in the relation between cognitive and neural events. The number of levels and their labeling are tentative yet serve to show the approach.

Level of analysis	Description
Cognitive domain	General, functional manifestation level, e.g., "visual perception" or "language comprehension"
Task- and domain-specific function	More specific function related to a domain, e.g., "visual form discrimination"
Basic functions	Basic-level operations, may be described in mathematical terms

absence of an impairment is typically established without reference to specific test procedures, but rather by a more direct (e.g., "general clinical") observation of the individual. For instance, it may be immediately evident that an aphasic patient is unable or severely impaired when attempting to communicate via language. This can be established without reference to particular tests while more detailed observation and testing is necessary to perform a further characterization of the affected subdomains within language, perception, or other post-traumatically affected domain. Similar arguments can be made regarding the demonstrations of recovery at this level.

With regards to task- and domain-specific functions, the situation becomes more complex. First of all, some may question the conceptual division between this level and the level of cognitive domains, arguing that this is nothing but a more elaborate description of the different domains. However, as a patient after brain injury may have very specific dysfunctions, lacking, e.g., the ability to recognize faces but have no other perceptual and/or memory problems, we believe the two levels can be separated at least as different levels of description or analysis. Generally speaking, it is conceptually and scientifically useful and important to keep the two levels distinct. For instance, two different visual tasks can only be conceived of as "visual" because they can be described at a "cognitive domain" level as well as a more specific task-level.

Basic functions are *prima facie* very different than the task-/domain-specific functions, as they should be conceived of as discretely operating systems without any direct manifestation at an experiential or a behavioral level. We consider these basic functions of the present model to be either identical to or at least at the same level as the Elementary Functions defined within the reorganization of elementary functions (REF) model (Mogensen and Malá, 2009; Mogensen, 2011a,c). Such Elementary Functions are basic information processing units (to be viewed as a kind of programming modules for a modular programming system of cognition). Elementary Functions are not specific to any domain or task, since an individual Elementary Function contributes to multiple higher order programs – each of which has a more task- and domain-specific application. The characterization of an Elementary Function may be more easily achieved in the language of mathematics than in the terminology of psychology.

The three levels of the presently proposed model are not to be seen as hierarchical in the sense that one is reducible to the other. However, the model suggests a conceptual hierarchy in the sense that for instance task-specific functions are categorized based on the cognitive domain (e.g., visual or auditory perception, logical reasoning etc). One may reasonably question whether the "granularity" of the neural processes involved should not be considered as well. For instance, one might question whether reorganization after brain injury will not lead to a structuring of the postlesional areas that is similar to the prelesional ones in such a way that the conception of multiple realization this way will be challenged. However, considering such examples of recovering language abilities after massive injuries to the left hemisphere, such a worry seems to lean on a rather "coarse grained" notion of the involved brain processes. In other words, we argue that such a worry may be applicable to some cases but is less relevant to a more general discussion of the possibility of multiple realization *per se*.

In the analysis of the neural substrate of post-traumatic recovery, it is more often than not impossible to tell at which "level" a neural substrate of postrehabilitation task mediation is a "replacement" of what has been lost. In other words: it is, for instance, unclear to what extent the same information processing ("function") is present in the right hemisphere when recovering aphasic patients show an increased focal activation within that hemisphere. This situation is often a result of the understandable focus on clinical applications and the potentials of a therapeutic method with respect to achieving a successful rehabilitative process. It can be argued that the essential information is that: (1) the patient is clinically improving, and (2) activation and/or establishment of particular (partly unknown) neural processes within a particular region of the brain is associated with this clinical success. Theoretically, however, such a pragmatic approach can be unsuccessful and potentially misleading. Clinically, reduction or elimination at the surface level of the post-traumatic symptoms is the crucial issue. Theoretically, however, it is of essential importance to establish not only whether a general level of functional recovery is seen, but also whether the mechanisms of such a recovery include re-establishment of the exact information processing lost to trauma. If the issue of such a re-establishment is not addressed, a more superficial examination of the post-traumatic rehabilitation of patients can potentially yield a false impression that lost information processing is reacquired.

In quite a number of cases, e.g., visuospatial neglect, the incidence of severe symptoms is dramatically high in the acute phase, yet most seem to recover (Corbetta et al., 2005). Perhaps less dramatically, this picture is found in many different dysfunctions where the pattern of a reduction of symptoms when moving from the acute to the postacute phase is very different in the acute and postacute phase. In the acute phase, patients suffering neglect may show the well-known symptoms of not eating food on the left side of their plate, bumping into objects to the left of them, etc. However, postacutely, patients who recover seem to represent space in the same way as healthy subjects, and no experiment till date has shown that recovered patients have any abnormalities in this respect. In such cases, *prima facie*, it seems rather safe to say that cognitive functions are realizable in different neural substrates at the level of "cognitive domains." So, at this most general level, there seems to be a 1:X relationship between cognitive states and brain states. On this background, it would seem plausible to also suggest that this is true at least at the task- and domain-specific level as well. Some studies, however, can be seen as challenging to this view.

Crosson et al. (2007) developed a method in which aphasic patients were trained in a naming task in a somewhat complex setup designed to achieve activation of right hemisphere mechanisms. Described in a simplified manner, the patients were required to perform a complex task using their left (right hemisphere-associated) hand as the initial step of the naming task. Theoretically, this voluntary activity can achieve a right hemisphere activation, which might facilitate other – more specifically language-associated – activations within that side of the brain. Although not globally effective, the method appeared to produce a significant level of success in many aphasic patients (Crosson et al., 2007). In a subsequent study Crosson et al. (2009) investigated – by the use of fMRI – the neural activation patterns associated with the improved performance of the naming task under such circumstances. The

results showed that the naming activity *per se* (not the performance of the complex motor task) was associated with an increase in focal activation within parts of the right frontal lobe. Apparently, this activation was associated with the linguistic improvements – in one or another respect representing a rightward "shift" of language processing. It is beyond doubts that such achievements are clinically significant and may constitute an important avenue in future developments of therapeutic interventions. What remains open to question, however, is the issue of exactly what is mediated by the frontal regions activated within the right hemisphere.

In most clinical studies, it is extremely hard or even impossible to step beyond the demonstration that a particular recovery process is accompanied by a particular focal activation. In animal models, however, there has been a certain level of success in achieving an analysis of whether or not a post-traumatic ability to mediate a given task performance is in reality the same as the normal mediation of that task. A couple of examples will serve to demonstrate that apparently similar task performances upon further scrutiny reveal themselves to be dissimilar.

Lepore et al. (1985) investigated in cats whether information relevant to the performance of a visual discrimination task can be transferred subcallosally (without the presence of the corpus callosum) between the two hemispheres. The cats of the study were subjected to split-brain operations in which the corpus callosum was severed. Initially, the discrimination task was acquired exclusively via one eye (thereby conveying the information to only one hemisphere – since the optic chiasm had been split). Subsequently, the ability to perform the visual discrimination task was tested via the other (untrained) eye – thereby testing the availability of discrimination-associated information to the untrained hemisphere. This test of subcallosal transfer of visual discrimination–relevant information was performed in two separate experimental setups: a classic two-choice discrimination box (a maze-type setup) in which food was offered as a reinforcement and a jumping-stand in which the motivation for successful task performance is the avoidance of an unpleasant fall. While the cats were able to demonstrate a subcallosal transfer of information in the jumping-stand, no such postlesional ability was demonstrated in the two-choice discrimination box.

In another experiment studying rats with lesions within the prefrontal system, it was also found that animals subjected to a focal lesion were in one experimental setup fully able to perform a task normally associated with the injured structure. But when tested in another version of the same task the animals were clearly impaired (Mogensen et al., 1987). Normally, across species an impaired performance of spatial delayed alternation tasks is considered a key symptom of injury within the prefrontal cortex or associated structures such as the prefrontal part of the neostriatum (e.g., Mogensen, 2003, 2011b; Mogensen et al., 2007, 2008). Mogensen et al. (1987) tested rats subjected to lesions of the prefrontal part of the neostriatum (in which even the prefrontal cortex was likely to have been rendered inoperational since the lesion of the neostriatum also destroyed fibers, which are essential to the normal functioning of the prefrontal cortex) in two variants of spatial delayed alternation tasks. One of these tasks was administered in a T-maze while the other was performed in an operant chamber. In the T-maze, two spatially separate positions (in the form of the two arms) are the goals toward which responses are to be made and the animal is removed from the apparatus between trials. In the operant chamber, choices are made between two spatially separate retractable levers situated in one wall of the chamber and the levers are retracted from the chamber between trials. While the animals demonstrated an impairment when tested in the T-maze, the test procedure administered in the operant chamber did not reveal any lesion-associated impairment – in spite of the fact that the operant chamber-based procedures imposed the same demands as the procedures in the T-maze regarding what are believed to be the crucial cognitive parameters such as intertrial delays and the necessity to constantly alternate between spatially dissimilar positions.

These two examples show that although neural mechanisms spared by the lesion in question (lesions of the corpus callosum or the prefrontal region of the neostriatum) are able to successfully mediate certain variants of a particular task, the neural mechanisms within those spared parts of the brain are (at least initially) unable to allow a successful task performance under circumstances which according to the formal task requirements are the same but are realized in another test-environment – thereby calling for at least partly different cognitive and neural mechanisms.

The third and final example from animal model-based studies addresses more specifically the mechanisms of a post-traumatic recovery process. Rats subjected to bilateral removal of the prefrontal cortex, bilateral hippocampal lesions in the form of transections of the fimbria–fornix or a combination of these two lesions were studied in a water maze-based allocentric place learning task of the mapping type by Mogensen et al. (2004). All three brain injured groups demonstrated an impaired task acquisition. While the group subjected to prefrontal lesions in isolation showed a relatively mild impairment, the groups subjected to hippocampal lesions in isolation or the combined lesions of the hippocampus and the prefrontal cortex, respectively, were equally and severely impaired. All animals did, however, eventually recover to a fully normal proficiency of task performance. Neither the speed of recovery nor the parameters of task performance revealed any difference between the two groups subjected to hippocampal lesions – begging the conclusion that if the hippocampus has been injured, the presence or absence of prefrontal mechanisms makes no difference to the mediation of task performance. Such a conclusion did, however, turn out to be a fallacy. Upon completion of the recovery training, all groups of animals were subjected to a series of "challenges" which aimed at demonstrating potential differences in: (1) the neural substrates mediating the successful task performance in the recovered animals, and (2) the cognitive strategies employed by various groups. These challenges demonstrated that when the hippocampus had been injured in isolation, the mediation of the successful recovery of the task received significant contributions from the prefrontal cortex. Obviously, such contributions were absent in the group subjected to lesions of both the hippocampus and the prefrontal cortex, and in these animals the most likely substrate of the successful post-traumatic functional recovery was the parietal association cortex. Furthermore, the two hippocampally lesioned groups differed significantly regarding the cognitive strategies employed (although these strategies were equally successful). In the presence of the prefrontal cortex, hippocampally lesioned animals were able to demonstrate an apparently normal level of cognitive representa-

tion of the position of the goal position (the platform in the water maze) while such a representation seemed to be absent in animals subjected to the combined lesion.

We would argue that all these three examples, in quite different ways, illustrate that cognitive functions are realizable in different neural substrates at the level of "cognitive domains." However, they all serve to cast doubts on the conclusion that the "second level phenomena" (domain- and task-specific functions) are multiple realized as well. An easy counter-argument to the position that cognition is multiple realized at the most general level only, could be that no matter how many examples one may line up of the lack of such realization, one needs only one certain example of successful multiple realization in order to show that this is possible in principle. Against this, proponents of the position that multiple realization occurs exclusively at the level of "cognitive domains" could argue that in each case where we yet have not found limitations to multiple realizations of cognitive functions at the "second level," the methods used have simply not been sufficiently exhaustive.

In conclusion, we believe it is relatively safe to say that multiple realization occurs at the "cognitive domain" level. At the lowest level of our model, the level of "basic functions," we believe it is too early to say anything specific as research on neuroplasticity is not yet at a sufficiently advanced stage. Currently, the discussion may focus on the "second level functions" where arguments may be formed for either position. If the position is correct that functions at this level are not subject to multiple realization, it would lead to the hypothesis that basic functions are not either.

LEVELS OF ANALYSIS IN CONSCIOUSNESS

One may now reasonably question whether this approach is specific for neural substrates of cognition, or whether it is applicable to other aspects of mental phenomena, e.g., the study of NCC. As mentioned above, cognitive states are defined with regards to their functions, whereas consciousness is defined as subjective experience (Nagel, 1974). Many leading theories, trying to pinpoint an NCC, claim that certain brain areas are necessary for a person to be conscious of particular contents, e.g., Milner and Goodale (Milner and Goodale, 1995; Goodale, 2007) who argue that ventral projection streams from primary visual cortex are necessary for visual consciousness, and diverse, even more influential theories argue that the prefrontal cortex is necessary (Crick and Koch, 1998; Dehaene and Naccache, 2001). Other theories seem more to focus on specific types of processes than specific brain areas (e.g., Tononi, 2004; Lamme, 2010).

Interestingly, in most cases, studies of NCC's are rather explicitly following a definition introduced by Chalmers (2000) that an NCC should be the minimally neural representational system sufficient for representation of that content in consciousness. This definition is in itself open to multiple realization of consciousness, which leaves it as an empirical question only whether one should think of specific conscious experiences as related to brain processes in a 1:1 or 1:X fashion.

Before proceeding, an important initial difference should however be considered. As we consider consciousness to be identical to subjective experience, which at large is the typical position in NCC literature (Seth et al., 2008), we cannot determine whether two instances of a specific conscious experience are identical based on behavioral measures. Instead, we are fully dependent on introspection and the introspecting subjects' reports (Naccache, 2006; Overgaard and Overgaard, 2010). Whereas several recent papers have argued that the use of introspection *per se* is not in conflict with the wish to do "proper science" (Piccinini, 2003; Overgaard, 2006), contrary to the leading notion not many decades ago, a specific problem arises in this context.

Even though it is not within the scope of this paper to fully analyze this problem, it is far from trivial how a person may introspectively compare two "instances" of an experiential state. If you are shown two blue colors in sequence, one lighter than the other, you would probably be able to compare the memories of the two perceptions and conclude that they were not identical. However, let us say you were to report whether you experience the exact same stimulus of a blue color in the exact same way as you did, say, a year ago, or before a brain injury that for a period changed your experience of color, it would certainly be a much more difficult task. One reasonable hypothesis about the identification of an experience of, say, a certain color, is its relation to experiences of other colors. In another example, emotions, most people would be willing to say that the feeling of, say, joy has important similarities with the feeling of joy they experienced as children. Were it not so, this specific emotion would in fact not be recognized as a feeling of joy. Nevertheless, whether joy is experienced in the exact same way is difficult to answer for many methodological reasons – not just, although important, because of memory decay. The feeling of joy may have been recognized and identified as "joy" under different premises when one was a child compared to today, which in important ways may depend on changes in how often one entertain this feeling, under which circumstances the feeling occurs, etc. Of course, the same argument could be pursued with many other examples (different types of emotions, memories, perceptions, etc.).

In recent years, there has been beginning progress in the attempt to develop more sophisticated methods to report subjective experiences adequately (e.g., Hurlburt, 2009; Overgaard and Timmermans, 2010; Sandberg et al., 2010), yet the field is far from giving exact methodological solutions to problems as those sketched above. Nevertheless, we believe the field is sufficiently mature to separate different analytical levels of experience, which may help this difficult discussion underway. In **Table 2** we provide a tentative framework for levels of analysis relevant to this issue.

A first, over-all distinction can be made following Rosenthal's (1997) concept of "creature consciousness" – whether a person or organism is conscious in any sense. This distinction is of special

Table 2 | Levels of analysis in the relation between conscious and neural events. The number of levels and their labeling are tentative.

Level of analysis	Description
"Creature level" consciousness	Whether a person or creature is conscious at any level or of any content
Phenomenal domain	Modality or "category" of contents, e.g., visual or auditory perceptions, emotions, etc.
Phenomenal content	The conscious experience itself with its contents
Character of content	How the content is manifested – with some degree of clarity, intensity, etc.

relevance in the discussion of coma or vegetative state patients where it is still unclear whether we should conceive of them as fully unconscious (yet with some intact cognitive processes) or as conscious to some extent (Owen et al., 2006; Overgaard, 2009). For patients who have lost consciousness for a period and recover it later, one possible interpretation would be that consciousness is now realized in a different neural substrate. This would be a case of NCC plasticity at a most general level.

A further distinction, we believe, can be made between this overall level of analysis and another level – "phenomenal domains," e.g., visual experiences, auditory experiences, emotions, thoughts, etc. "Domains," in this case, refer to experiences for which the difference between them are not just the actual content of the experience, but the "modality" or "kind of presentation."

As already now indicated, this level differs from the actual content, which may or may not include simultaneous differences between modality. In other words, at this level, a difference between, say, the visual experience of two different colors differ just like a visual and auditory experience do.

Finally, each content may itself be different from other instances of the same content. For instance, a conscious perceptual content may be present in different degrees of clarity or intensity. It is a theoretical debate worth having, but not possible in this context, whether two "identical" contents with different intensities as a matter of principle form two different contents, or whether one and the same content may vary in different aspects. In this context, we assume the latter position.

It is even more complex in this area of investigation – compared to the discussions above – to find experiments that can be said to reasonably argue for or against plasticity limited to one level.

With regards to the "creature level," some experiments give reason to believe that the NCC may change after recovery from coma. In a PET study of a single subject, global metabolic function decreased substantially in the vegetative state following CO intoxication – as would be expected. However, after full recovery, the metabolic level stayed low globally (Laureys, 2005). Although some prefrontal areas did increase metabolism, it seems safe to say that whatever areas or processes may be involved in the NCC, they would be different in the pretraumatic and the post-recovery situation.

In one study, Ptito et al. (2008) applied transcranial magnetic stimulation (TMS) to early blind subjects and blindfolded controls. The blindfolded subjects reported experiencing phosphenes following occipital cortex stimulation, as expected from several previous studies (e.g., Amassian et al., 1998). Some of the blind subjects, nevertheless, reported tactile sensations in the fingers. The representation of those fingers appeared to be somatotopically organized within the visual cortex. Importantly, the number of cortical sites at which tactile sensations could be elicited in the blind was related to the number of hours of "Braille reading" per day. The tactile experiences do show that brain areas that normally would be associated with visual processing are now related to tactile experiences, which indicates cross-domain plasticity.

Cross-domain plasticity in case of individuals deprived of input within one modality is not restricted to "alien source" activations of the visual system of the blind. For instance, Lambertz et al. (2005) reported activation of the auditory cortex due to visual stimulation in deaf subjects.

Intra-modal plasticity related to consciousness is, also, rarely studied, yet some experiments give a few indications of such occurrences. For instance, Ramachandran and Hirstein (1998) studied phantom limb phenomena, where the experience of stroking a missing arm was associated with a factual stroking of the face even though tactile inputs from the arm and face map onto different, adjacent, cortical areas. The case is difficult to interpret as clearly showing plasticity in the neural substrate for phenomenal content, as this would demand an argument that phantom arm sensations are fully identical with sensations in a "normal" arm. Unfortunately, very few studies have investigated the "characters of content" empirically (with few exceptions, e.g., Overgaard et al. (2008)), let alone the possibility for plasticity in their neural substrate.

These examples, though clearly not exhaustive, show that an empirical case could be made in favor of NCC plasticity – from a very general level to at least the level of phenomenal content. Counter-arguments could certainly take many different forms. One very general kind of counter-argument would say that in each of the here mentioned examples, the apparent loss of consciousness and its later recovery (in the Laureys case) or the changed correlation between a specific brain area and conscious experiences associated with it (in the other two cases), what has really changed is not the NCC *per se*, but rather other circumstances which are necessary for the NCC – yet not themselves parts hereof. Stopping the heart from beating would certainly make a person lose consciousness, yet this does not in itself show that the heart is a "correlate of consciousness." According to the definition by Chalmers (2000), as earlier mentioned, we are looking for the minimally sufficient correlate, not a necessary correlate. Therefore, one could argue, these examples need to show that what has happened is in fact a change in the NCC.

Whereas this debate in the present context stands as unresolved, this, at least, gives us a framework in which this debate may take place and potentially be resolved by empirical means.

CONCLUSION

Taken together, we believe to have offered a general framework in which a qualified debate about the multiple realization of cognitive and conscious events can take place. As mentioned underway, we wish to underline that the number of levels as well as their exact character in the two "models" are highly tentative and may be subject to revision. Nevertheless, we believe they capture the essence of the kind of structure, we will propose.

With this structure, positions in the debate of multiple realizations are no longer different ways to argue for or against the phenomenon as such, but opens up to more nuanced positions, arguing for plasticity at some levels and against plasticity at other.

We have deliberately said very little about the theoretical consequences of these possible positions, yet this is certainly a question worth pursuing in future work. We do believe that advances in this debate could indicate empirical ways to answer otherwise theoretical questions about the mind–brain relation.

ACKNOWLEDGMENTS

Morten Overgaard was supported by a Starting Grant from the European Research Council. Jesper Mogensen was supported by a grant from the Danish Research Council.

REFERENCES

Amassian, V. E., Cracco, R. Q., Maccabee, P. J., Cracco, J. B., Rudell, A. P., and Eberle, L. (1998). Transcranial magnetic stimulation in the study of the visual pathway. *J. Clin. Neurophysiol.* 15, 288–304.

Ansaldo, A. I., and Arguin, M. (2003). The recovery from aphasia depends on both the left and right hemispheres: three longitudinal case studies on the dynamics of language function after aphasia. *Brain Lang.* 87, 177–178.

Ansaldo, A. I., Arguin, M., and Lecours, A. R. (2002). The contribution of the right cerebral hemisphere to the recovery from aphasia: a single longitudinal case study. *Brain Lang.* 82, 206–222.

Arvidsson, A., Collin, T., Kirik, D., Kokaia, Z., and Lindvall, O. (2002). Neuronal replacement from endogenous precursors in the adult brain after stroke. *Nat. Med.* 8, 963–970.

Bach-y-Rita, P., Collins, C. C., Saunders, F. A., White, B., and Scadden, L. (1969). Vision substitution by tactile image projection. *Nature* 221, 963–964.

Bach-y-Rita, P., Kaczmarek, K. A., Tyler, M. E., and Garcia-Lara, J. (1998). Form perception with a 49-point electrotactile stimulus array on the tongue: a technical note. *J. Rehabil. Res. Dev.* 35, 1–7.

Barlow, T. (1877). On a case of double cerebral hemiplegia, with cerebral symmetrical lesions. *Br. Med. J.* 2, 103–104.

Baumgaertner, A., Schraknepper, V., and Saur, D. (2005). Differential recovery of aphasia and apraxia of speech in an adolescent after infarction of the left frontal lobe: longitudinal behavioral and fMRI data. *Brain Lang.* 95, 211–212.

Bechtel, W., and Mundale, J. (1999). Multiple realizability revisited: linking cognitive and neural states. *Philos. Sci.* 66, 175–207.

Berardi, N., Pizzorusso, T., and Maffei, L. (2004). Extracellular matrix and visual cortical plasticity: freeing the synapse. *Neuron* 44, 905–908.

Berry, M. (1982). Post-injury myelin-breakdown products inhibit axonal growth: an hypothesis to explain the failure of axonal regeneration in the mammalian central nervous system. *Bibl. Anat.* 23, 1–11.

Block, N., and Fodor, J. (1972). What psychological states are not. *Philos. Rev.* 81, 159–181.

Chalmers, D. J. (2000). "What is a neural correlate of consciousness?" in *Neural Correlates of Consciousness: Empirical and Conceptual Questions*, ed. T. Metzinger (Cambridge, MA: MIT Press), 17–40.

Chen, D. F., Jhaveri, S., and Schneider, G. E. (1995). Intrinsic changes in developing retinal neurons result in regenerative failure of their axons. *Proc. Natl. Acad. Sci. U.S.A.* 92, 7287–7291.

Chen, J., Magavi, S. S., and Macklis, J. D. (2004). Neurogenesis of corticospinal motor neurons extending spinal projections in adult mice. *Proc. Natl. Acad. Sci. U.S.A.* 101, 16357–16362.

Corbetta, M., Kincade, M. J., Lewis, C., Snyder, A. Z., and Sapir, A. (2005). Neural basis and recovery of spatial attentional deficits in spatial neglect. *Nat. Neurosci.* 8, 1424–1425.

Crick, F., and Koch, C. (1998). Consciousness and neuroscience. *Cereb. Cortex* 8, 97–107.

Crosson, B., McGregor, K., Gopinath, K. S., Conway, T. W., Benjamin, M., Chang, Y.-L., Moore, A. B., Raymer, A. M., Briggs, R. W., Sherod, M. G., Wierenga, C. E., and White, K. D. (2007). Functional MRI of language in aphasia: a review of the literature and the methodological challenges. *Neuropsychol. Rev.* 17, 157–177.

Crosson, B., Moore, A. B., McGregor, K. M., Chang, Y.-L., Benjamin, M., Gopinath, K., Sherod, M. E., Wierenga, C. E., Peck, K. K., Briggs, R. W., Rothi, L. J. G., and White, K. D. (2009). Regional changes in word-production laterality after a naming treatment designed to produce a rightward shift in frontal activity. *Brain Lang.* 111, 73–85.

Dehaene, S., and Naccache, L. (2001). Towards a cognitive neuroscience of consciousness. *Cognition* 79, 1–37.

Del Rio, J. A., and Soriano, E. (2007). Overcoming chondroitin sulphate proteoglycan inhibition of axon growth in the injured brain: lessons from chondroitinase ABC. *Curr. Pharm. Des.* 13, 2485–2492.

Duffau, H., Denvil, D., and Capelle, L. (2002). Long term reshaping of language, sensory, and motor maps after glioma resection: a new parameter to integrate in the surgical strategy. *J. Neurol. Neurosurg. Psychiatr.* 72, 511–516.

Elbert, T., Pantev, C., Weinbruch, C., Rockstroh, B., and Taub, E. (1995). Increased cortical representation of the fingers of the left hand in string players. *Science* 270, 305–307.

Fawcett, J. W., Housden, E., Smith-Thomas, L., and Meyer, R. L. (1989). The growth of axons in three-dimensional astrocyte cultures. *Dev. Biol.* 135, 449–458.

Finger, S., Buckner, R. L., and Buckingham, H. (2003). Does the right hemisphere take over after damage to Broca's area? The Barlow case of 1877 and its history. *Brain Lang.* 85, 385–395.

Finger, S., and Stein, D. G. (1982). "Vicariation theory and radical reorganization of function," in *Brain Damage and Recovery: Research and Clinical Perspectives,* eds S. Finger and D. G. Stein (New York: Academic Press), 287–302.

Gizewski, E. R., Gasser, T., de Greiff, A., Boehm, A., and Forsting, M. (2003). Cross-modal plasticity for sensory and motor activation patterns in blind subjects. *Neuroimage* 19, 968–975.

Goldberg, J. L., Klassen, M. P., Hua, Y., and Barres, B. A. (2002). Amacrine-signaled loss of intrinsic axon growth ability by retinal ganglion cells. *Science* 296, 1860–1864.

Goodale, M. (2007). "Duplex vision: separate cortical pathways for conscious perception and the control of action," in *The Blackwell companion to consciousness*, eds M. Velmans and S. Schneider (Oxford: Blackwell), 616–627.

Hurlburt, R. T. (2009). "Descriptive experience sampling," in *Oxford Companion to Consciousness*, eds T. Bayne, A. Cleeremanns, and P. Wilken (Oxford: Oxford University Press), 225–227.

Irvine, D. R. F. (2007). Auditory cortical plasticity: does it provide evidence for cognitive processing in the auditory cortex? *Hear. Res.* 229, 158–170.

Kaczmarek, K. A., Webster, J. G., Bach-y-Rita, P., and Tompkins, W. J. (1991). Electrotactile and vibrotactile displays for sensory substitutionsystems. *IEEE Trans. Biomed. Eng.* 38, 1–16.

Karl, A., Birbaumer, N., Lutzenberger, W., Cohen, L. G., and Flor, H. (2001). Reorganization of motor and somatosensory cortex in upper extremity amputees with phantom limb pain. *J. Neurosci.* 15, 3609–3618.

Keller-Peck, C. R., Walsh, M. K., Gan, W.-B., Feng, G., Sanes, J. R., and Lichtman, J. W. (2001). Asynchronous synapse elimination in neonatal motor units: studies using GFP transgenic mice. *Neuron* 31, 381–394.

Laganaro, M., Morand, S., Schwitter, V., Zimmermann, C., and Schnider, A. (2008). Normalisation and increase of abnormal ERP patterns accompany recovery from aphasia in the post-acute stage. *Neuropsychologia* 46, 2265–2273.

Lambertz, N., Gizewski, E. R., de Greiff, A., and Forsting, M. (2005). Cross-modal plasticity in deaf subjects dependent on the extent of hearing loss. *Cogn. Brain Res.* 25, 884–890.

Lamme, V. (2010). How neuroscience will change our view on consciousness. *Cogn. Neurosci.* 1, 204–220.

Laureys, S. (2005). The neural correlate of (un)awareness: lessons from the vegetative state. *Trends Cogn. Sci. (Regul. Ed.)* 9, 556–559.

Leavitt, B. R., Hernit-Grant, C. S., and Macklis, J. D. (1999). Mature astrocytes transform into transitional radial glia within adult mouse neocortex that supports directed migration of transplanted immature neurons. *Exp. Neurol.* 157, 43–57.

Lee, W. C., Huang, H., Feng, G., Sanes, J. R., Brown, E. N., So, P. T., and Nedivi, E. (2006). Dynamic remodeling of dendritic arbors in GABAergic interneurons of adult visual cortex. *PLoS Biol.* 4, e29. doi: 10.1371/journal.pbio.0040029

Lee, W.-C. A., Chen, J. L., Huang, H., Leslie, J. H., Amitai, Y., So, P. T., and Nedivi, E. (2008). A dynamic zone defines interneuron remodeling in the adult neocortex. *Proc. Natl. Acad. Sci. U.S.A.* 105, 19968–19973.

Lepore, F., Ptito, M., Provencal, C., Bedard, S., and Guillemot, J.-P. (1985). Interhemispheric transfer of visual training in the split-brain cat: effects of the experimental set-up. *Can. J. Psychol.* 37, 527.

Magavi, S. S., Leavitt, B. R., and Macklis, J. D. (2000). Induction of neurogenesis in the neocortex of adult mice. *Nature* 405, 951–955.

Marsh, E. B., and Hillis, A. E. (2006). Recovery from aphasia following brain injury: the role of reorganization. *Prog. Brain Res.* 157, 143–156.

McGee, A. W., Yang, Y., Fischer, Q. S., Daw, N. W., and Strittmatter, S. M. (2005). Experience-driven plasticity of visual cortex limited by myelin and Nogo receptor. *Science* 309, 2222–2226.

Meinzer, M., Flaisch, T., Breitenstein, C., Wienbruch, C., Elbert, T., and Rockstroh, B. (2008). Functional re-recruitment of dysfunctional brain areas predicts language recovery in chronic aphasia. *Neuroimage* 39, 2038–2046.

Meinzer, M., Obleser, J., Flaisch, T., Eulitz, C., and Rockstroh, B. (2007). Recovery from aphasia as a function of language therapy in an early bilingual patient demonstrated by fMRI. *Neuropsychologia* 45, 1247–1256.

Merzenich, M. M., and Jenkins, W. M. (1993). Reorganization of cortical representations of the hand following alterations of skin inputs induced by nerve injury, skin island transfers, and experience. *J. Hand Ther.* 6, 89–104.

Milner, A. D., and Goodale, M. A. (1995). *The Visual Brain in Action*. Oxford: Oxford University Press.

Mogensen, J. (2003). "Animal models in neuroscience," in *Handbook of Laboratory Animal Science*, 2nd Edn, Vol. II, *Animal Models*, eds J. Hau and G. L. van Hoosier (Boca Raton, FL: CRC Press LLC), 95–109.

Mogensen, J. (2011a). Almost unlimited potentials of a limited neural plasticity: levels of plasticity in development and reorganization of the injured brain. *J. Conscious. Stud.* (in press).

Mogensen, J. (2011b). "Animal models in neuroscience," in *Handbook of Laboratory Animal Science*, 3rd Edn,

Vol. II, *Animal Models*, eds J. Hau and S. Schapiro (Boca Raton, FL: CRC Press LLC), (in press).

Mogensen, J. (2011c). Reorganization in the injured brain: implications for studies of the neural substrate of cognition. *Front. Psychol.* 2:7. doi: 10.3389/fpsyg.2011.00007

Mogensen, J., Boyd, M. H., Nielsen, M. D., Kristensen, R. S., and Malá, H. (2008). Erythropoietin improves spatial delayed alternation in a T-maze in rats subjected to ablation of the prefrontal cortex. *Brain Res. Bull.* 77, 1–7.

Mogensen, J., Hjortkjær, J., Ibervang, K. L., Stedal, K., and Malá, H. (2007). Prefrontal cortex and hippocampus in posttraumatic functional recovery: spatial delayed alternation by rats subjected to transection of the fimbria–fornix and/or ablation of the prefrontal cortex. *Brain Res. Bull.* 73, 86–95.

Mogensen, J., Iversen, I. H., and Divac, I. (1987). Neostriatal lesions impaired rats' delayed alternation performance in a T-maze but not in a two-key operant chamber. *Acta Neurobiol. Exp. (Wars)* 47, 45–54.

Mogensen, J., Lauritsen, K. T., Elvertorp, S., Hasman, A., Moustgaard, A., and Wörtwein, G. (2004). Place learning and object recognition by rats subjected to transection of the fimbria-fornix and/or ablation of the prefrontal cortex. *Brain Res. Bull.* 63, 217–236.

Mogensen, J., and Malá, H. (2009). Posttraumatic functional recovery and reorganization in animal models. A theoretical and methodological challenge. *Scand. J. Psychol.* 50, 561–573.

Münte, T. F., Altenmüller, E., and Jäncke, L. (2002). The musician's brain as a model of neuroplasticity. *Nat. Rev. Neurosci.* 3, 473–478.

Naccache, L. (2006). Is she conscious? *Science* 313, 1395–1396.

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

Nakatomi, H., Kuriu, T., Okabe, S., Yamamoto, S.-C., Hatano, O., Kawahara, N., Tamura, A., Kirino, T., and Nakafuku, M. (2002). Regeneration of hippocampal pyramidal neurons after ischemic brain injury by recruitment of endogenous neural progenitors. *Cell* 110, 429–441.

Overgaard, M. (2006). Introspection in science. *Conscious. Cogn.* 15, 629–633.

Overgaard, M. (2009). How can we know if patients in coma, vegetative state or minimally conscious state are conscious. *Prog. Brain Res.* 177, 11–19.

Overgaard, M., Fehl, K., Mouridsen, K., and Cleeremans, A. (2008). Seeing without seeing? Degraded conscious vision in a blindsight patient. *PLoS ONE* 3, e3028. doi: 10.1371/journal.pone.0003028

Overgaard, M., and Overgaard, R. (2010). Neural correlates of contents and levels of consciousness. *Front. Psychol.* 1:164. doi: 10.3389/fpsyg.2010.00164

Overgaard, M., and Timmermans, B. (2010). "How unconscious is subliminal perception?" in *Handbook of Phenomenology and Cognitive Science*, eds D. Schmicking and S. Gallagher (Heidelberg: Springer Verlag), 501–519.

Owen, A. M., Coleman, M. R., Boly, M., Davis, M. H., Laureys, S., and Pickard, J. D. (2006). Detecting awareness in the vegetative state. *Science* 313, 1402

Perani, D., Cappa, S. F., Tettamanti, M., Rosa, M., Scifo, P., Miozzo, A., Basso, A., and Fazio, F. (2003). A fMRI study of word retrieval in aphasia. *Brain Lang.* 85, 357–368.

Piccinini, G. (2003). Data from introspective reports: upgrading from commonsense to science. *J. Conscious. Stud.* 10, 141–156.

Pizzorusso, T., Medini, P., Berardi, N., Chierzi, S., Fawcett, J. W., and Maffei, L. (2002). Reactivation of ocular dominance plasticity in the adult visual cortex. *Science* 298, 1248–1251.

Ptito, M., Fumal, A., de Noordhout, A. M., Schoenen, J., Gjedde, A., and Kupers, R. (2008). TMS of the occipital cortex induces tactile sensations in the fingers of blind Braille readers. *Exp. Brain Res.* 184, 193–200.

Ptito, M., Moesgaard, S. M., Gjedde, A., and Kupers, R. (2005). Cross-modal plasticity revealed by electrotactile stimulation of the tongue in the congenitally blind. *Brain* 128, 606–614.

Rakic, P. (1971). Guidance of neurons migrating to the fetal monkey neocortex. *Brain Res.* 33, 471–476.

Rakic, P. (1985). "Mechanisms of neuronal migration in developing cerebellar cortex," in *Molecular Basis of Neural Development*, eds G. M. Edelman, W. M. Cowan, and E. Gull (New York: Wiley), 139–160.

Ramachandran, V. S., and Hirstein, W. (1998). The perception of phantom limbs. *Brain* 121, 1603–1630.

Recanzone, G. H., Schreiner, C. E., and Merzenich, M. M. (1993). Plasticity in the frequency representation of primary auditory cortex following discrimination training in adult owl monkeys. *J. Neurosci.* 13, 87–103.

Robertson, D., and Irvine, D. R. F. (1989). Plasticity of frequency organization in auditory cortex of guinea pigs with partial unilateral deafness. *J. Comp. Neurol.* 282, 456–471.

Rosenthal, D. (1997). "A theory of consciousness," in *The Nature of Consciousness*, eds N. Block, O. Flanagan, and G. Güzeldere (Cambridge, MA: MIT Press), 729–753.

Sandberg, K., Timmermans, B., Overgaard, M., and Cleeremans, A. (2010). Measuring consciousness: is one measure better than the other? *Conscious. Cogn.* 19, 1069–1078.

Schäfer, R., Dehn, D., Burbach, G. J., and Deller, T. (2008). Differential regulation of chondroitin sulfate proteoglycan mRNAs in the denervated rat fascia dentata after unilateral entorhinal cortex lesion. *Neurosci. Lett.* 439, 61–69.

Scharff, C., Kirn, J. R., Grossman, M., Macklis, J. D., and Nottebohm, G. (2000). Targeted neuronal death affects neuronal replacement and vocal behavior in adult songbirds. *Neuron* 25, 481–492.

Scheich, H. (1991). Auditory cortex: comparative aspects of maps and plasticity. *Curr. Opin. Neurobiol.* 1, 236–247.

Schwab, M. E., and Thoenen, H. (1985). Dissociated neurons regenerate into sciatic but not optic nerve explants in culture irrespective of neurotrophic factors. *J. Neurosci.* 5, 2415–2423.

Seth, A, Dienes, Z, Cleeremans, A, Overgaard, M., and Pessoa, L. (2008). Measuring consciousness: relating behavioural and neurophysiological approaches. *Trends Cogn. Sci. (Regul. Ed.)* 12, 314–321.

Shapiro, L. (2004). *The Mind Incarnate*. Cambridge, MA: MIT Press.

Shapiro, L. (2008). How to test for multiple realization. *Philos. Sci.* 75, 514–525.

Slavin, M. D., Laurence, S., and Stein, D. G. (1988). "Another look at vicariation," in *Brain Injury and Recovery: Theoretical and Controversial Issues*, eds S. Finger, T. E. LeVere, C. R. Almli, and D. G. Stein (New York: Plenum Press), 165–178.

Specht, K., Zahn, R., Willmes, K., Weis, S., Holtel, C., Krause, B. J., Herzog, H., and Huber, W. (2009). Joint independent component analysis of structural and functional images reveals complex patterns of functional reorganisation in stroke aphasia. *Neuroimage* 47, 2057–2063.

Szaflarski, J. P., Eaton, K., Ball, A. L., Banks, C., Vannest, J., Allendorfer, J. B., Page, S., and Holland, S. K. (2010). Poststroke aphasia recovery assessed with functional magnetic resonance imaging and a picture identification task. *J. Stroke Cerebrovasc. Dis.* 1–10.

Thai-Van, H., Micheyl, C., Norena, A., Veuillet, E., Gabriel, D., and Collet, L. (2007). Enhanced frequency discrimination in hearing-impaired individuals: a review of perceptual correlates of central neural plasticity induced by cochlear damage. *Hear. Res.* 233, 14–22.

Thomas, C., Altenmüller, E., Marckmann, G., Kahrs, J., and Dichgans, J. (1997). Language processing in aphasia: changes in lateralization patterns during recovery reflect cerebral plasticity in adults. *Electroencephalogr. Clin. Neurophysiol.* 102, 86–97.

Thompson, C. K., den Ouden, D.-B., Bonakdarpour, B., Garibaldi, K., and Parrish, T. B. (2010). Neural plasticity and treatment-induced recovery of sentence processing in agrammatism. *Neuropsychologia* 48, 3211–3227.

Thulborn, K. R., Carpenter, P. A., and Just, M. A. (1999). Plasticity of language-related brain function during recovery from stroke. *Stroke* 30, 749–754.

Tononi, G. (2004). An information integration theory of consciousness. *BMC Neurosci.* 5, 42. doi: 10.1186/1471-2202-5-42

Weiss, T., Miltner, W. H. R., Huonker, R., Friedel, R., Schmidt, I., and Taub, E. (2000). Rapid functional plasticity of the somatosensory cortex after finger amputation. *Exp. Brain Res.* 134, 199–203.

Wilms, I., and Mogensen, J. (in press). Dissimilar outcomes of apparently similar procedures as a challenge to clinical neurorehabilitation and basic research – when the same is not the same. *Neurorehabilitation.*

Xerri, C., Coq, J., Merzenich, M., and Jenkins, W. (1996). Experience-induced plasticity of cutaneous maps in the primary somatosensory cortex of adult monkeys and rats. *J. Physiol.* 90, 277–287.

Yang, T. T., Gallen, C. C., Ramachandran, V. S., Cobb, S., Schwartz, B. J., and Bloom, F. E. (1994). Noninvasive detection of cerebral plasticity in adult human somatosensory cortex. *Neuroreport* 5, 701–704.

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

Received: 04 October 2010; paper pending published: 14 December 2010; accepted: 13 April 2011; published online: 27 April 2011. Citation: Overgaard M and Mogensen J (2011) A framework for the study of multiple realizations: the importance of levels of analysis. Front. Psychology **2**:*79. doi: 10.3389/fpsyg.2011.00079*

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

frontiers in PSYCHOLOGY

HYPOTHESIS AND THEORY ARTICLE
published: 09 May 2011
doi: 10.3389/fpsyg.2011.00086

The radical plasticity thesis: how the brain learns to be conscious

Axel Cleeremans*

Consciousness, Cognition and Computation Group, Université Libre de Bruxelles, Bruxelles, Belgium

Edited by:
Morten Overgaard, Aarhus University, Aarhus University Hospital, Denmark

Reviewed by:
Hakwan C. Lau, Columbia University in the City of New York, USA
Zoltan Dienes, University of Sussex, UK

****Correspondence:***
Axel Cleeremans, Consciousness, Cognition and Computation Group, Université Libre de Bruxelles CP 191, 50 Avenue F.-D. Roosevelt, B1050 Bruxelles, Belgium.
e-mail: axcleer@ulb.ac.be

In this paper, I explore the idea that consciousness is something that the brain learns to do rather than an intrinsic property of certain neural states and not others. Starting from the idea that neural activity is inherently unconscious, the question thus becomes: How does the brain learn to be conscious? I suggest that consciousness arises as a result of the brain's continuous attempts at predicting not only the consequences of its actions on the world and on other agents, but also the consequences of activity in one cerebral region on activity in other regions. By this account, the brain continuously and unconsciously learns to redescribe its own activity to itself, so developing systems of meta-representations that characterize and qualify the target first-order representations. Such learned redescriptions, enriched by the emotional value associated with them, form the basis of conscious experience. Learning and plasticity are thus central to consciousness, to the extent that experiences only occur in experiencers that have learned to *know* they possess certain first-order states and that have learned to *care* more about certain states than about others. This is what I call the "Radical Plasticity Thesis." In a sense thus, this is the enactive perspective, but turned both inwards and (further) outwards. Consciousness involves "signal detection on the mind"; the conscious mind is the brain's (non-conceptual, implicit) theory about itself. I illustrate these ideas through neural network models that simulate the relationships between performance and awareness in different tasks.

Keywords: consciousness, learning, subjective experience, neural networks, emotion

Consider the humble but proverbial thermostat. A thermostat is a simple device that can turn a furnace on or off depending on whether the current temperature exceeds a set threshold. Thus, the thermostat can appropriately be said to be *sensitive* to temperature. But is there some sense in which the thermostat can be characterized as being *aware* of temperature? Contra Chalmers (1996), I will argue that there is no sense in which the thermostat can be characterized as being aware of temperature. There are two important points that I would like to emphasize in developing this argument. The first is that there is no sense in which the thermostat can be characterized as being aware of temperature because it does not *know that* it is sensitive to temperature. The second point is that there is no sense in which the thermostat can be characterized as being aware of temperature because it does not *care* about whether its environment is hot or cold. I will further argue that these two features – knowledge of one's own internal states and the emotional value associated with such knowledge – are constitutive of conscious experience. Finally, I will argue that learning (or, more generally, plasticity) is necessary for both features to emerge in cognitive systems. From this, it follows that consciousness is something that the brain learns to do through continuously operating mechanisms of neural plasticity. This I call the "Radical Plasticity Thesis."

Information processing can undoubtedly take place without consciousness, as abundantly demonstrated not only by empirical evidence (the best example of which is probably blindsight), but also by the very fact that extremely powerful information-processing machines, namely computers, have now become ubiquitous. Only but a few would be willing to grant any quantum of conscious experience to contemporary computers, yet they are undeniably capable of sophisticated information processing – from recognizing faces to analyzing speech, from winning chess tournaments to helping prove theorems. Thus, consciousness is not information processing; experience is an "extra ingredient" (Chalmers, 2007a) that comes over and beyond mere computation.

With this premise in mind – a premise that just restates Chalmers' (1996) *hard problem*, that is, the question of *why* it is the case that information processing is accompanied by experience in humans and other higher animals, there are several ways in which one can think about the problem of consciousness.

One is to simply state, as per Dennett (e.g., Dennett, 1991, 2001) that there is nothing more to explain. Experience is *just* (a specific kind of) information processing in the brain; the contents of experience are *just* whatever representations have come to dominate processing at some point in time ("fame in the brain"); consciousness is *just* a harmless illusion. From this perspective, it is easy to imagine that machines will be conscious when they have accrued sufficient complexity; the reason they are not conscious now is simply because they are not sophisticated enough: They lack the appropriate architecture perhaps, they lack sufficiently broad and diverse information-processing abilities, and so on. Regardless of what is missing, the basic point here is that there is no reason to assume that conscious experience is anything special. Instead, all that is required is one or several yet-to-be-identified functional mechanisms: Recurrence, perhaps (Lamme, 2003), stability of representation (O'Brien and Opie, 1999), global availability (Baars,

1988; Dehaene et al., 1998), integration and differentiation of information (Tononi, 2003, 2007), or the involvement of higher-order representations (Rosenthal, 1997, 2006), to name just a few.

Another perspective is to consider that *experience* will never be amenable to a satisfactory functional explanation. Experience, according to some (e.g., Chalmers, 1996), is precisely what is left over once all functional aspects of consciousness have been explained. Notwithstanding the fact that so defined, experience is simply not something one can approach from a scientific point of view, this position recognizes that consciousness is a unique (a *hard*) problem in the Cognitive Neurosciences. But that is a different thing from saying that a reductive account is not possible. A non-reductive account, however, is exactly what Chalmers' Naturalistic Dualism attempts to offer, by proposing that information, as a matter of ontology, has a dual aspect, – a physical aspect and a phenomenal aspect. "Experience arises by virtue of its status as one aspect of information, when the other aspect is found embodied in physical processing" (Chalmers, 2007b, p. 366). This position leads him to defend the possibility that experience is a fundamental aspect of reality. Thus, even thermostats, for instance, may be endowed with very simple experiences, in virtue of the fact that they can toggle in two different states.

What, however, do we mean when we speak of "subjective experience" or of "quale"? The simplest definition of these concepts (Nagel, 1974) goes right to the heart of the matter: "Experience" is *what it feels like* for a conscious organism to be that organism. There is something it is like for a bat to be a bat; there is nothing it is like for a stone to be a stone. As Chalmers (2007a) puts it: "When we see, for instance, we *experience* visual sensations: The felt quality of redness, the experience of dark and light, the quality of depth in a visual field" (p. 226).

Let us try to engage in some phenomenological analysis at this point to try to capture what it means for each of us to have an experience. Imagine you see a patch of red (Humphrey, 2006). You now have a *red* experience – something that a camera recording the same patch of red will most definitely *not* have. What is the difference between you and the camera? Tononi (2007), from whom I borrow this simple thought experiment, points out that one key difference is that when you see the patch of red, the state you find yourself in is but one among billions, whereas for a simple light-sensitive device, it is perhaps one of only two possible states – thus the state conveys a lot more *differentiated information* for you than for a light-sensitive diode. A further difference is that you are able to *integrate* the information conveyed by many different inputs, whereas the chip on a camera can be thought of as a mere array of independent sensors among which there is no interaction.

Hoping not to sound presumptuous, it strikes me, however, that both Chalmers' (somewhat paradoxically) and Tononi's analyses miss fundamental facts about experience: Both analyze it as a rather abstract dimension or aspect of information, whereas experience – *what it feels like* – is anything but abstract. On the contrary, what we mean when we say that seeing a patch of red elicits an "experience" is that the seeing *does something to us* – in particular, we might feel one or several emotions, and we may associate the redness with memories of red. Perhaps seeing the patch of red makes you remember the color of the dress that your prom night date wore 20 years ago. Perhaps it evokes a vague anxiety, which we now know is also shared by monkeys (Humphrey, 1971). To a synesthete, perhaps seeing the color red will evoke the number 5. The point is that if conscious experience is what it feels like to be in a certain state, then "What it feels like" can only mean the specific set of associations that have been established by experience between the stimulus or the situation you now find yourself in, on the one hand, and your memories, on the other. This is what one means by saying that there is something it is like to be you in this state rather than nobody or somebody else: The set of memories evoked by the stimulus (or by actions you perform, etc.), and, crucially, the set of emotional states associated with each of these memories. This is essentially the perspective that Damasio (2010) defends.

Thus, a first point about the very notion of subjective experience I would like to make here is that it is difficult to see what experience could mean beyond (1) the emotional value associated with a state of affairs, and (2) the vast, complex, richly structured, experience-dependent network of associations that the system has learned to associate with that state of affairs. "What it feels like" for me to see a patch of red at some point seems to be entirely exhausted by these two points. Granted, one could still imagine an agent that accesses specific memories, possibly associated with emotional value, upon seeing a patch of red and who fails to "experience" anything. But I surmise that this would be mere simulation: One *could* design such a zombie agent, but any real agent that is driven by self-developed motivation, and that cannot help but be influenced by his emotional states will undoubtedly have experiences much like ours.

Hence, there is nothing it is like for the camera to see the patch of red simply because it does not care: The stimulus is meaningless; the camera lacks even the most basic machinery that would make it possible to ascribe any interpretation to the patch of red; it is instead just a mere recording device for which nothing matters. There is nothing it is like to be that camera at that point in time simply because (1) the experience of different colors do not do anything to the camera; that is, colors are not associated with different emotional valences; and (2) the camera has no brain with which to register and process its own states. It is easy to imagine how this could be different. To hint at my forthcoming argument, a camera could, for instance, keep a record of the colors it is exposed to, and come to "like" some colors better than others. Over time, your camera would like different colors than mine, and it would also know that in some non-trivial sense. Appropriating one's mental contents for oneself is the beginning of individuation, and hence the beginning of a *self*.

Thus a second point about experience that I perceive as crucially important is that it does not make any sense to speak of experience without an *experiencer* who experiences the experiences. Experience is, almost by definition ("what it feels like"), something that takes place not in *any* physical entity but rather only in special physical entities, namely cognitive agents. Chalmers' (1996) thermostat fails to be conscious because, despite the fact that it can find itself in different internal states, it lacks the ability to remove itself from the causal chain which it instantiates. In other words, it lacks knowledge *that* it can find itself in different states; it is but a mere mechanism that responds to inputs in certain ways. While there is indeed something to be experienced there (the different states the thermostat can find itself in), there is no one home to be the *subject* of these experiences – the thermostat simply lacks

the appropriate machinery to do so. The required machinery, I surmise, minimally involves the ability to *know that* one finds itself in such or such a state.

This point can be illustrated by means of well-known results in the connectionist, or artificial neural network modeling literature. Consider for instance Hinton's (1986) famous demonstration that neural networks trained through associative learning mechanisms can learn about abstract dimensions of the training set. Hinton's (1986) network was a relatively simple back-propagation network trained to process linguistic expressions consisting of an agent, a relationship, and a patient, such as for instance "Maria is the wife of Roberto." The stimulus material consisted of a series of such expressions, which together described some of the relationships that exist in the family trees of an Italian family and of an English family The network was required to produce the patient of each agent–relationship pair it was given as input. For instance, the network should produce "Roberto" when presented with "Maria" and "wife." Crucially, each person and each relationship were presented to the network by activating a single input unit. Hence there was no overlap whatsoever between the input representations of, say, Maria and Victoria. Yet, despite this complete absence of surface similarity between training exemplars, Hinton (1986) showed that after training, the network could, under certain conditions, develop internal representations that capture relevant abstract dimensions of the domain, such as nationality, sex, or age!

Hinton's (1986) point was to demonstrate that such networks were capable of learning richly structured internal representations as a result of merely being required to process exemplars of the domain. Crucially, the structure of the internal representations learned by the network is determined by the manner in which different exemplars interact with each other, that is, by their *functional similarity,* rather than by their mere *physical similarity* expressed, for instance, in terms of how many features (input units) they share. Hinton (1986) thus provided a striking demonstration of this important and often misunderstood aspect of associative learning procedures by showing that under some circumstances, specific hidden units of the network had come to act as detectors for dimensions of the material that had never been presented explicitly to the network. These results truly flesh out the notion that rich, abstract knowledge can simply emerge as a by-product of processing structured domains. It is interesting to note that the existence of such single-unit "detectors" has recently been shown to exist in human neocortex (Kreiman et al., 2002): Single-neuron recording of activity in hippocampus, for instance, has shown that some individual neurons exclusively respond to highly abstract entities, such as the words "Bill Clinton" and images of the American president.

Now, the point I want to make with this example is as follows: One could certainly describe the network as being *sensitive* to nationality, in the sense that it exhibits differential responding (hence, behavioral sensitivity) to inputs that involve Italian agents vs. English agents. But, obviously, the network does not *know* anything about nationality. It does not even know that it has such and such representations of the inputs, nor does it know anything about its own, self-acquired sensitivity to the relevant dimensions. Instead, the rich, abstract, structured representations that the network has acquired over training forever remain embedded in a causal chain that begins with the input and ends with the network's responses. As Clark and Karmiloff-Smith (1993) insightfully pointed out, such representations are "first-order" representations to the extent that they are representations *in the system* rather than representations *for the system* that is, such representations are not accessible to the network *as representations.*

In other words, such a (first-order) network can never know *that* it knows: It simply lacks the appropriate machinery. This points to a fundamental difference between sensitivity and awareness. Sensitivity merely entails the ability to respond in specific ways to certain states of affairs. Sensitivity does not require consciousness in any sense. A thermostat can appropriately be characterized as being sensitive to temperature, just as the carnivorous plant Dionaea Muscipula may appropriately be described as being sensitive to movement on the surface of its leaves. But our intuitions (at least, my intuitions) tell us that such sensitive systems (thermostats, photodiodes, transistors, cameras, carnivorous plants) are not conscious. They do not have "elementary experiences," they simply have no experiences whatsoever. Sensitivity can involve highly sophisticated knowledge, and even learned knowledge, as illustrated by Hinton's (1986) network, but such knowledge is always first-order knowledge, it is always knowledge that is necessarily embedded in the very same causal chain through which first-order processing occurs and that can therefore only be expressed through action as a direct result of perception.

Awareness, on the other hand, always seems to minimally entail the ability of knowing *that* one knows. This ability, after all, forms the basis for the verbal reports we take to be the most direct indication of awareness. And when we observe the absence of such ability to report on the knowledge involved in our decisions, we rightfully conclude that the decision was based on unconscious knowledge. Thus, it is when an agent exhibits *knowledge* of the fact that he is sensitive to some state of affairs that we take this agent to be a conscious agent. This *second-order* knowledge, I argue, critically depends on *learned* systems of meta representations, and forms the basis for conscious experience provided the agent also *cares* about certain states of affairs more than about others.

Consciousness thus not only requires ability to learn about the geography of one's own representations, but it also requires that the resulting knowledge reflects the dispositions and preferences of the agent. This is an important point, for it would be easy to program a thermostat that is capable not only of acting based on the current temperature, but also to report on its own states. Such a talking thermostat would constantly report on the current temperature and on its decisions. Would that make the thermostat conscious? Certainly not, for it is clear that the reporting is but a mere additional process tacked on the thermostat's inherent ability to switch the furnace according to the temperature. What would go some way toward making the thermostat conscious is to set it up so that it *cares* about certain temperatures more than about others, and that these preferences emerge as a result of learning.

What would it take for a network like Hinton's (1986) to be able to access its own representations, and what difference would that make with respect to consciousness? To answer the first question, the required machinery is the machinery of agenthood; in a nutshell, the ability to do something not just with external states of affairs, but rather with one own's representations of such external states. This crucially requires that the agent be able to access,

inspect, and otherwise manipulate its own representations, and this in turn, I surmise, requires mechanisms that make it possible for an agent to redescribe its own representations to itself. The outcome of this continuous "representational redescription" (Karmiloff-Smith, 1992) process is that the agent ends up knowing something about the geography of its own internal states: It has, in effect, *learned* about its own representations. Minimally, this could be achieved rather simply, for instance by having another network take both the input (i.e., the external stimulus as represented proximally) to the first-order network and its internal representations of that stimulus as inputs themselves and do something with them.

One elementary thing the system consisting of the two interconnected networks (the first-order, observed network and the second-order, observing network) would now be able to do is to make decisions, for instance, about the extent to which an external input to the first-order network elicits a familiar pattern of activation over its hidden units or not. This would in turn enable the system to distinguish between hallucination and blindness (see Lau, 2008), or to come up with judgments about the performance of the first-order network (Persaud et al., 2007; Dienes, 2008).

To address the second question (what difference would representational redescription make in terms of consciousness), I appeal to Rosenthal's (1997, 2006) higher-order thought (HOT) theory of consciousness. While I do not feel perfectly happy with all aspects of HOT Theory, I do believe, however, that higher-order representations (I will call them meta-representations in what follows) play a crucial role in consciousness.

An immediate objection to this idea is as follows: If there is nothing intrinsic to the existence of a representation in a cognitive system that makes this representation conscious, why should things be different for meta-representations? After all, meta-representations are representations also. Yes indeed, but with a crucial difference: Meta-representations inform the agent about its own internal states, making it possible for it to develop an understanding of its own workings. And this, I argue, forms the basis for the contents of conscious experience, provided of course – which cannot be the case in an contemporary artificial system – that the system has learned about its representations by itself, over its development, and provided that it cares about what happens to it, that is, provided its behavior is rooted in emotion–laden motivation (to survive, to mate, to find food, etc.).

THE RADICAL PLASTICITY THESIS

I would thus like to defend the following claim: Conscious experience occurs if and only if an information-processing system has *learned* about its own representations of the world in such a way that these representations have acquired value for it. To put this claim even more provocatively: Consciousness is the brain's (emphatically non-conceptual) theory about itself, gained through experience interacting with the world, with other agents, and, crucially, with itself. I call this claim the "*Radical Plasticity Thesis*," for its core is the notion that learning is what makes us conscious.

Before getting to the core of the argument, I should briefly sketch a framework through which to characterize the relationships between learning and consciousness. If the main cognitive function of consciousness is to make adaptive control of behavior possible, as is commonly accepted, then consciousness is necessarily closely related to processes of learning, because one of the central consequences of successful adaptation is that conscious control is no longer required over the corresponding behavior. Indeed, it might seem particularly adaptive for complex organisms to be capable of behavior that does not require conscious control, for instance because behavior that does not require monitoring of any kind can be executed faster or more efficiently than behavior that does require such control. What about conscious experience? Congruently with our intuitions about the role of consciousness in learning, we often say of somebody who failed miserably at some challenging endeavor, such as completing a paper by the deadline, that the failure constitutes "a learning experience." What precisely do we mean by this? We mean that the person can now learn from her mistakes, that the experience of failure was sufficiently imbued with emotional value that it has registered in that person's brain. The experience *hurt*, it made one realize what was at stake, it made us think about it, in other words, it made us consciously aware of what failed and why.

But this minimally requires what Kirsh (1991) has called "explicit representation," namely the presence of representations that directly represent the relevant information. By "direct" here, I mean that the information is represented in such a manner that no further computation is required to gain access to it. For instance, a representation that is explicit in this sense might simply consist of a population of neurons that fire whenever a specific condition holds: A particular stimulus is present on the screen, my body is in a particular state (i.e., pain, or hunger).

By assumption, however, such "explicit" representations are not necessarily conscious. Instead, they are merely good candidates to enter conscious awareness in virtue of features such as their stability, their strength, or their distinctiveness (Cleeremans, 1997; Cleeremans and Jiménez, 2002). What is missing, then? What is missing is that such representations be themselves the target of other representations. And how would this make any difference? It makes a crucial difference, for the relevant first-order *representations* are now part of the agent's known repertoire of mental states; such representations are then, and only then, recognized by the agent as playing the function of representing some other (internal or external) state of affairs.

NECESSARY CONDITIONS FOR AWARENESS

Let us now focus on the set of assumptions that together form the core of a framework that characterizes how learning shapes availability to consciousness (see Cleeremans and Jiménez, 2002; Cleeremans, 2008, for more detailed accounts). It is important to keep it in mind that the framework is based on the connectionist framework (Rumelhart and McClelland, 1986). It is therefore based on many central ideas that characterize the connectionist approach, such as the fact that information processing is graded and continuous, and that it takes place over many interconnected modules consisting of processing units. In such systems, long-term knowledge is embodied in the pattern of connectivity between the processing units of each module and between the modules themselves, while the transient patterns of activation over the units of each module capture the temporary results of information processing.

This being said, a first important assumption is that *representations are graded, dynamic, active, and constantly causally efficacious* (Cleeremans, 1994, 2008). Patterns of activation in neural networks

and in the brain are typically distributed and can therefore vary on a number of dimensions, such as their stability in time, their strength. or their distinctiveness. *Stability* in time refers to how long a representation can be maintained active during processing. There are many indications that different neural systems involve representations that differ along this dimension. For instance, prefrontal cortex, which plays a central role in working memory, is widely assumed to involve circuits specialized in the formation of the enduring representations needed for the active maintenance of task-relevant information. *Strength* of representation simply refers to how many processing units are involved in the representation, and to how strongly activated these units are. As a rule, strong activation patterns will exert more influence on ongoing processing than weak patterns. Finally, *distinctiveness* of representation is inversely related to the extent of overlap that exists between representations of similar instances. Distinctiveness has been hypothesized as the main dimension through which cortical and hippocampal representations differ (McClelland et al., 1995; O'Reilly and Munakata, 2000), with the latter becoming active only when the specific conjunctions of features that they code for are active themselves.

In the following, I will collectively refer to these different dimensions as "quality of representation" (Farah, 1994). The most important notion that underpins these different dimensions is that representations, in contrast to the all-or-none propositional representations typically used in classical theories, instead have a *graded* character that enables any particular representation to convey the extent to which what it refers to is indeed present.

Another important aspect of this characterization of representational systems in the brain is that, far from being static propositions waiting to be accessed by some process, representations instead continuously influence processing regardless of their quality. This assumption takes its roots in McClelland's (1979) analysis of cascaded processing which, by showing how modules interacting with each other need not "wait" for other modules to have completed their processing before starting their own, demonstrated how stage-like performance could emerge out of such continuous, non-linear systems. Thus, even weak, poor-quality traces are capable of influencing processing, for instance through associative priming mechanisms, that is, in *conjunction* with other sources of stimulation. Strong, high-quality traces, in contrast have *generative capacity*, in the sense that they can influence performance independently of the influence of other constraints, that is, whenever their preferred stimulus is present.

A second important assumption is that *learning is a mandatory consequence of information processing*. Indeed, every form of neural information-processing produces adaptive changes in the connectivity of the system, through mechanisms such as long-term potentiation (LTP) or long-term depression (LTD) in neural systems, or Hebbian learning in connectionist systems. An important aspect of these mechanisms is that they are mandatory in the sense that they take place whenever the sending and receiving units or processing modules are co-active. O'Reilly and Munakata (2000) have described Hebbian learning as instantiating what they call *model learning*. The fundamental computational objective of such unsupervised learning mechanisms is to enable the cognitive system to develop useful, informative models of the world by capturing its correlational structure. As such, they stand in contrast with *task learning* mechanisms, which instantiate the different computational objective of mastering specific input–output mappings (i.e., achieving specific goals) in the context of specific tasks through error-correcting learning procedures.

Stability, strength, or distinctiveness can be achieved by different means. Over short time scales, they can result, for instance, from increased stimulus duration, from the simultaneous top-down and bottom-up activation involved in so-called "reentrant processing" (Lamme, 2006), from processes of "adaptive resonance" (Grossberg, 1999), from processes of "integration and differentiation" (Edelman and Tononi, 2000), or from contact with the neural workspace, brought about by "dynamic mobilization" (Dehaene and Naccache, 2001). It is important to realize that the ultimate effect of any of these putative mechanisms is to make the target representations stable, strong, and distinctive. These properties can further be envisioned as involving graded or dichotomous dimensions (see also Maia and Cleeremans, 2005 for an exploration of how connectionist principles are relevant to the study of consciousness).

Over longer time scales, however, high-quality representations arise as a result of learning or cognitive development. Weak, fragile representations become progressively stronger and higher-quality. As a result, they exert more of an influence on behavior. In most cases, this is a good outcome because the stronger a representation is, the less it will require conscious control and monitoring. Thus, in any domain of experience (from being able to stand up to wine-tasting, from recognizing faces to reading) we begin with weak representations, which are characteristic of implicit cognition and do not require control because they only exert weak effects on behavior. Such representations, because of their poor quality, are also only weakly available to form the contents of consciousness. As learning progresses, the relevant representations become stronger, yet not so strong that they can be "trusted" to do their job properly. This is when cognitive control is most necessary. This is also the point where such explicit representations are most likely to form the contents of consciousness. Finally, with further training, the relevant representations become even stronger and eventually fully adapted. As such, these high-quality representations characteristic of automaticity no longer require cognitive control either, but this is so for completely different reasons than the weak representations characteristic of implicit cognition.

Thus, when I respond faster to a target stimulus in virtue of the fact that the target was preceded by a congruent subliminal prime, I can properly say that there exists a state *c* such that its existence made me respond faster, but by assumption I am not sensitive to the fact that this state *c* is different from state *i* where the target stimulus was preceded by an incongruent prime. States *c* and *i* are thus not conscious states – they merely exert their effects on behavior, so reflecting the agent's *sensitivity* to their existence, but crucially not its *awareness* of their existence. The reason such states are not conscious states has to do with the properties of the corresponding first-order states: It is not so much that there is a failure of a higher-order system to target these states, but rather that the first-order states are too weak to be appropriate targets. You cannot know what is not (sufficiently) there.

Likewise, but perhaps more controversially so, habitual, automatic behavior is often described as involving unconscious knowledge: The behavior unfolds whether you intend to or not, it can

unfold with attention engaged elsewhere, and so on. In such cases, behavior is driven by very high-quality representations that have become, through experience, optimally tuned to drive behavior. While such very high-quality representations are appropriate objects for redescription, the redescriptions either no longer play a functional role or are prevented from taking place (for instance because the agent's attention is engaged elsewhere). Automatic behavior is thus not truly unconscious behavior (Tzelgov, 1997). Rather, it is behavior for which awareness has become optional. You can be perfectly aware of behavior that occurs automatically – you just seldom do so for it is neither necessary nor desirable for you to become aware of such behavior. That is precisely *why* the behavior has become automatic: Because it so adapted that it can unfold without the need for conscious monitoring.

Hence a first important computational principle through which to distinguish between conscious and unconscious representations is the following:

> *Availability to consciousness depends on quality of representation, where quality of representation is a graded dimension defined over stability in time, strength, and distinctiveness.*

While being of high-quality thus appears to be a necessary condition for a representation's availability to consciousness, one should ask, however, whether it is a sufficient condition. Cases such as hemineglect or blindsight (Weiskrantz, 1986) clearly suggest that quality of representation alone does not suffice, for even strong stimuli can fail to enter conscious awareness in such conditions. In normal participants, the attentional blink (Shapiro et al., 1997), as well as inattentional (Mack and Rock, 1998) and change blindness (Simons and Levin, 1997), are all suggestive that high-quality stimuli can simply fail to be experienced unless attended to. Likewise, merely achieving stable representations in an artificial neural network, for instance, will not make this network conscious in any sense – this is the problem pointed out by Clark and Karmiloff-Smith (1993) about the limitations of what they called first-order networks: In such networks, even explicit knowledge (e.g., a stable pattern of activation over the hidden units of a standard back-propagation network that has come to function as a "face detector") remains knowledge that is in the network as opposed to knowledge for the network. In other words, such networks might have learned to be informationally sensitive to some relevant information, but they never know that they possess such knowledge. Thus the knowledge can be deployed successfully through action, but only in the context of performing some particular task.

Hence it could be argued that it is a defining feature of consciousness that when one is conscious of something, one is also, at least potentially so, conscious *that* one is conscious of being in that state. This is the gist of so-called HOT theories of consciousness (Rosenthal, 1997), according to which a mental state is conscious when the agent entertains, in an non-inferential manner, thoughts to the effect that it currently is in that mental state. Importantly, for Rosenthal, it is in virtue of occurrent HOTs that the target first-order representations become conscious. Dienes and Perner (1999) have developed this idea by analyzing the implicit-explicit distinction as reflecting a hierarchy of different manners in which the representation can be explicit. Thus, a representation can explicitly indicate a property (e.g., "yellow"), predication to an individual (the flower is yellow), factivity (it is a fact and not just a possibility that the flower is yellow) and attitude (I know that the flower is yellow). Fully conscious knowledge is thus knowledge that is "attitude-explicit."

This analysis suggests that a further important principle that differentiates between conscious and unconscious cognition is the extent to which a given representation endowed with the proper properties (stability, strength, distinctiveness) is itself the target of meta-representations.

Hence a second important computational principle through which to distinguish between conscious and unconscious representations is the following:

> *Availability to consciousness depends on the extent to which a representation is itself an object of representation for further systems of representation.*

It is interesting to consider under which conditions a representation will remain unconscious based on combining these two principles (Cleeremans, 2008). There are at least four possibilities. First, knowledge that is embedded in the connection weights within and between processing modules can never be directly available to conscious awareness and control. This is simply a consequence of the fact that consciousness, by assumption, necessarily involves representations (patterns of activation over processing units). The knowledge embedded in connection weights will, however, shape the representations that depend on it, and its effects will therefore detectable – but only indirectly, and only to the extent that these effects are sufficiently marked in the corresponding representations. This is equivalent to Dehaene and Changeux's (2004) principle of "active firing."

Second, to enter conscious awareness, a representation needs to be of sufficiently high-quality in terms of strength, stability in time, or distinctiveness. Weak representations are therefore poor candidates to enter conscious awareness. This, however, does not necessarily imply that they remain causally inert, for they can influence further processing in other modules, even if only weakly so. This forms the basis for a host of sub-threshold effects, including, in particular, subliminal priming.

Third, a representation can be strong enough to enter conscious awareness, but failed to be associated with relevant meta-representations. There are thus many opportunities for a particular conscious content to remain, in a way, implicit, not because its representational vehicle does not have the appropriate properties, but because it fails to be integrated with other conscious contents.

Finally, a representation can be so strong that its influence can no longer be controlled – automaticity. In theses cases, it is debatable whether the knowledge should be taken as genuinely unconscious, because it can certainly become fully conscious as long as appropriate attention is directed to them (Tzelgov, 1997), but the point is that such very strong representations can trigger and support behavior without conscious intention and without the need for conscious monitoring of the unfolding behavior.

SUFFICIENT CONDITIONS FOR AWARENESS?

Strong, stable, and distinctive representations are thus *explicit* representations, at least in the sense put forward by Koch (2004): They indicate what they stand for in such a manner that their

reference can be retrieved directly through processes involving low computational complexity (see also Kirsh, 1991, 2003). Conscious representations, in this sense, are explicit representations that have come to play, through processes of learning, adaptation, and evolution, the functional role of denoting a particular content for a cognitive system. Importantly, quality of representation should be viewed as a *graded* dimension. This is essential to capture the fact that phenomenal experience, particularly ordinary phenomenal experience, appears graded itself. Gradedness can be achieved in different ways in a complex system such as the brain. One possibility is that representations are inherently graded because their vehicles are patterns of activation distributed over populations of firing neurons. Another is that representations tend to be all-or-none, but always involve multiple levels of a hierarchy (Kouider et al., 2010).

Once a representation has accrued sufficient strength, stability, and distinctiveness, it may be the target of meta-representations: The system may then "realize," if it is so capable, that is, if it is equipped with the mechanisms that are necessary to support self-inspection, that it has learned a novel partition of the input; that it now possesses a new "detector" that only fires when a particular kind of stimulus, or a particular condition, is present. Humphrey (2006) emphasizes the same point when he states that "This self-monitoring by the subject of his own response is the prototype of the "feeling sensation" as we humans know it" (p. 90). Importantly, my claim here is that such meta-representations are learned in just the same way as first-order representations, that is, by virtue of continuously operating learning mechanisms. Because meta-representations are also representations, the same principles of stability, strength, and distinctiveness therefore apply. An important implication of this observation is that activation of meta-representations can become automatic, just as it is the case for first-order representations.

What might be the function of such meta-representations? One possibility is that their function is to indicate the mental attitude through which a first-order representation is held: Is this something I know, hope, fear, or regret? Possessing such metaknowledge about one's knowledge has obvious adaptive advantages, not only with respect to the agent himself, but also because of the important role that communicating such mental attitudes to others plays in both competitive and cooperative social environments.

What is the mechanism through which such redescription is achieved? Minimally, enabling redescription of one's own internal states requires such internal states to be *available* to redescription, where *availability* is contingent, as described above, on such internal states being *patterns of activation* endowed with certain characteristics such as their strength, their stability in time, and their distinctiveness. Note that these assumptions rule out many potential sources of internal knowledge. For instance, the sort of weak, fleeting representations presumably resulting from the presentation of a brief stimulus would be poor candidates to be available to further processing. Likewise, the associative links that exist between representations, if implemented through patterns of connectivity between groups of units (as they are in connectionist networks) would likewise be inaccessible. Finally, and though this is more speculative (but see Brunel et al., 2010), it may also be the case that the highly distributed representations typical of semantic knowledge (i.e., my knowledge of a typical dog) are less available to form the contents of conscious experience then are the highly distinctive representations characteristic of episodic memory.

Second, those representations that meet these minimal requirements for redescription need to be accessed by another, independent part of the system whose function it is to redescribe them. It is important to note here that mere redescription probably does not cut it, for even in a simple feedforward network, each layer can be thought of as being a redescription of the input. The brain is massively hierarchical and thus contains multiple such redescriptions of any input. Instead of being strictly hierarchically organized, however, the redescriptions that count for the mechanism I have in mind should be removed from the causal chain responsible for the first-order processing. Hence, we need some mechanism that can access and redescribe first-order representations in a manner that is independent from the first-order causal chain.

I suggest that the general form of such mechanisms is something similar to what is depicted in **Figure 1**. Two independent networks (the first-order network and the second-order network) are connected to each other in such a way that the entire first-order network is input to the second-order network. Both networks are simple feedforward back-propagation networks. The first-order network consists of thee pools of units: a pool of input units, a pool of hidden units, and a pool of output units. Let us further imagine that this network is trained to perform a simple discrimination task, that is, to produce what is named Type I response in the language of Signal-Detection Theory. My claim is that there is nothing in the computational principles that characterize how this network performs its task that is intrinsically associated with awareness. The network simply performs the task. While it will develop knowledge of the associations between its inputs and outputs over its hidden units, and while this knowledge may be in some cases very sophisticated, it will forever remain knowledge that is "in" the network as opposed to being knowledge "for" the network. In other words, such a (first-order) network can never know *that* it knows: It simply lacks the appropriate machinery to do so. Likewise, in Signal-Detection Theory, while Type 1 responses always reflect sensitivity to some state of affairs, this sensitivity may or may not be conscious sensitivity. That is, a participant may be successful in discriminating one stimulus from another, yet fail to be aware *that* he is able to do so and thus claim, if asked, that he is merely guessing or responding randomly. In its more general form, as depicted in **Figure 1**, such an architecture would also be sufficient for the second-order network to also perform other judgments, such as distinguishing between an hallucination and a veridical perception, or developing knowledge about the overall geography of the internal representations developed by the first-order network (see also Nelson and Narens, 1990).

Can we use such architectures to account for relevant data? That is the question we set out to answer in recent work (e.g., Cleeremans et al., 2007; Pasquali et al., 2010) aimed at exploring the relationships between performance and awareness. We have found that different approaches to instantiating the general principles we have described so far are required to capture empirical findings. In one, as hinted at above, the first-order and the second-order network are part of the same causal chain, but are trained on different tasks, one corresponding to first-order decisions and the second

corresponding to metacognitive decisions. In a second approach, the two networks are truly independent. Note that in either case, our assumptions are oversimplified, for a complete implementation of the theory would require that the second-order network may influence processing as it takes place in the first-order network by means of recurrence.

In the following, I will illustrate the first approach, through which we have focused on architectures in which the first and second-order networks function as part of the same causal chain. Post-decision wagering was recently introduced by Persaud et al. (2007) as a measure of awareness through which participants are required to place a high or a low wager on their decision, such as relative to stimulus identification for example. The intuition behind this measure is that people will place a high wager when they have conscious knowledge about the reasons for their decisions, and a low wager when they are uncertain of their decisions. In this, wagering is thus similar to other subjective measures of awareness (Seth et al., 2008; Sandberg et al., 2010). According to Persaud et al. (2007) wagering provides an incentive for participants not to withhold any conscious information, as well as not to guess, making it a more objective measure of awareness than confidence judgment. Despite recent criticism of Persaud et al.'s claims (Dienes and Seth, 2010; Sandberg, et al., 2010), wagering certainly reflects the extent to which an agent is sensitive to its own internal states. In Cleeremans et al. (2007), we therefore aimed at creating a wagering network, for wagering affords easy quantification and thus appeared more readily amenable to computational simulation than other metacognitive measures such as confidence. In one of our simulations, which I will describe in more detail here, the first-order feedforward back-propagation network (see **Figure 2**) consisted of 7 input units representing digit shapes (as on a digital watch), 100 hidden units, and 10 output units for the 10 digits. The task of the first-order network is a simple one: It consists of identifying the "visual" representations of the digits 0–9. This is achieved by training the first-order network to respond to each input by activating one of its 10 output units. The 100 first-order hidden units connected to a different pool of 100 hidden units of the second-order feedforward network, with 2 output units representing a high and a low wager, as shown in **Figure 2**. The task of the higher-order network consisted of wagering high if it "thought" that the first-order network was providing a correct answer (correct identification of the digit), and to wager low in case the first network gave a wrong answer (misidentification of the digit). Note that as implemented here, there is no substantial difference between wagering and merely expressing confidence judgments.

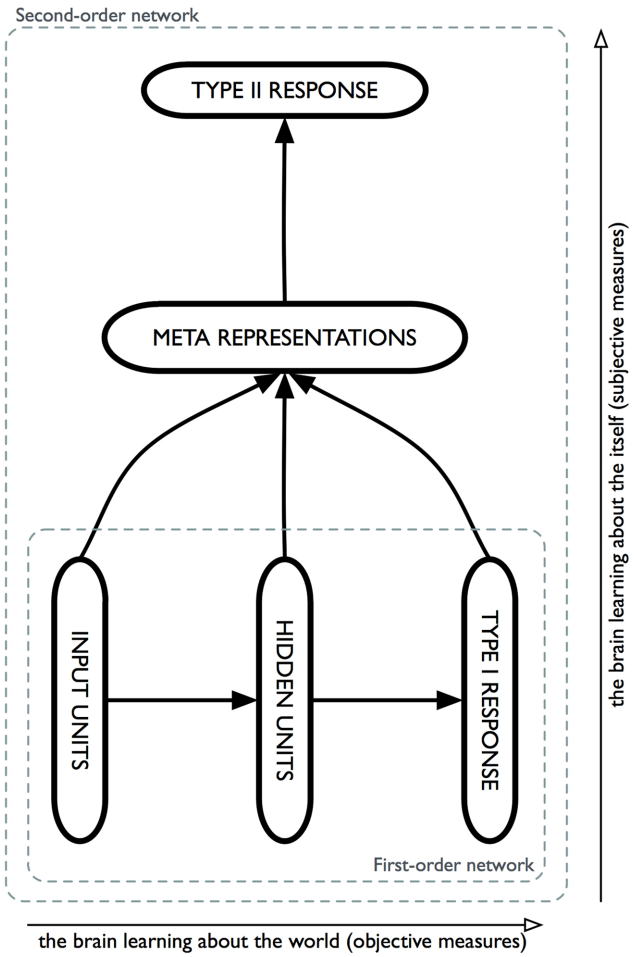


FIGURE 1 | General architecture of a metacognitive network. A first-order network, consisting for instance of a simple three-layers back-propagation network, has been trained to perform a simple classification task and thus contains knowledge that links inputs to outputs in such a way that the network can produce Type I responses. By design, this entire first-order network then constitutes the input to a second-order network, the task of which consists of redescribing the activity of the first-order network in some way. Here, the task that this second-order network is trained to perform is to issue Type II responses, that is, judgments about the extent to which the first-order network has performed *its* task correctly. One can think of the first-order network as instantiating cases where the brain learns about the world, and of the second-order network as instantiating cases where the brain learns about itself.

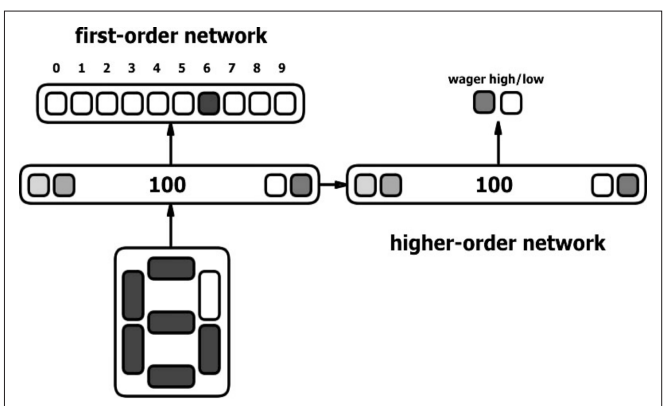


FIGURE 2 | Architecture of a wagering network. A first-order network instantiates a simple pattern classifier trained to classify "visual" input patterns representing the shapes of digits 0–9 in 10 categories. A second-order network is assigned the task of wagering on the first-order network's performance based on the latter's internal representations of the stimulus. The second-order network thus performs judgments about the extent to which the first-order network is correct in its own decisions.

A learning rate of 0.15 and a momentum of 0.5 were used during training of the first-order network. In a first condition of "high awareness," the second network was trained with a learning rate of 0.1, and in a second condition of "low awareness," a learning rate of 10^{-7} was applied. Ten networks were trained to perform their tasks concurrently throughout 200 epochs of training and their performance averaged. The performance of all three networks is depicted in **Figure 3**. Chance level for the first-order network is 10% (there is one chance of out 10 of correctly identifying one digit amongst 10); it is 50% for the second-order network (one chance out of two of placing a correct bet). The figure shows that the first-order network simply gradually learns to improve its classification performance continuously until it achieves 100% correct responses at the end of training. The performance of the "high awareness" second-order network, however, exhibits a completely different pattern. Indeed, one can see that the second-order network initially performs quite well, only to show decreasing performance up until about epoch 40, at which point its performance has sagged to chance level. From epoch 40 onwards, the second-order network's performance increases in parallel with that of the first-order network. This u-shaped performance pattern is replicated, to a lesser degree and with slightly different dynamics, in the "low awareness" second-order network.

One can understand this performance pattern as follows. Initially, the second-order network quickly learns that the first-order network is systematically incorrect in classifying the digits. (which is expected since it has not begun to learn how to perform the task). The safest response (i.e., the response that minimizes error) is thus to always bet low. This, incidentally, is what any rational agent would do. However, as the first-order network quickly begins to exceed chance level performance on its digit classification task, the performance of the second-order network begins to decrease. This corresponds to a stage where the second-order network is beginning to bet "high" on some occasions as it learns to categorize states of the first-order network that are predictive of a correct classification. An interesting pattern of dissociation then occurs, for the second-order network is performing rather poorly just when the first-order network is beginning to truly master its own digit classification task. One can think of that stage as corresponding to a point in training where the system as a whole is essentially acting based on unconscious knowledge: First-order performance on the digit classification task is well above chance level, yet, wagering by the second-order network is close to chance, and is at chance on epoch 40. Later on, after epoch 40, the second-order network has learned enough about when the first-order network will be correct vs. incorrect to begin attempting to maximize its own wagering performance. Thus, epoch 40 corresponds to the second-order network's "most doubtful moment." One could view this as the moment at which the higher-order network abandons a simple "safe" strategy of low wagers and explores the space of first-order hidden unit representations, looking for a criterion that will allow it to separate good from bad identifications.

Thus, as the two networks learn simultaneously to perform their respective tasks, one sees the entire system shifting from a situation where there is no relationship between first- and second-order performance to a situation where the two are correlated. This transition reflects, under our assumptions, a shift between unconscious vs. conscious processing.

In later work (Pasquali, et al., 2010), we have explored similar models based on germane or identical architectures and shown that they are capable of accounting for the data reported by Persaud et al.

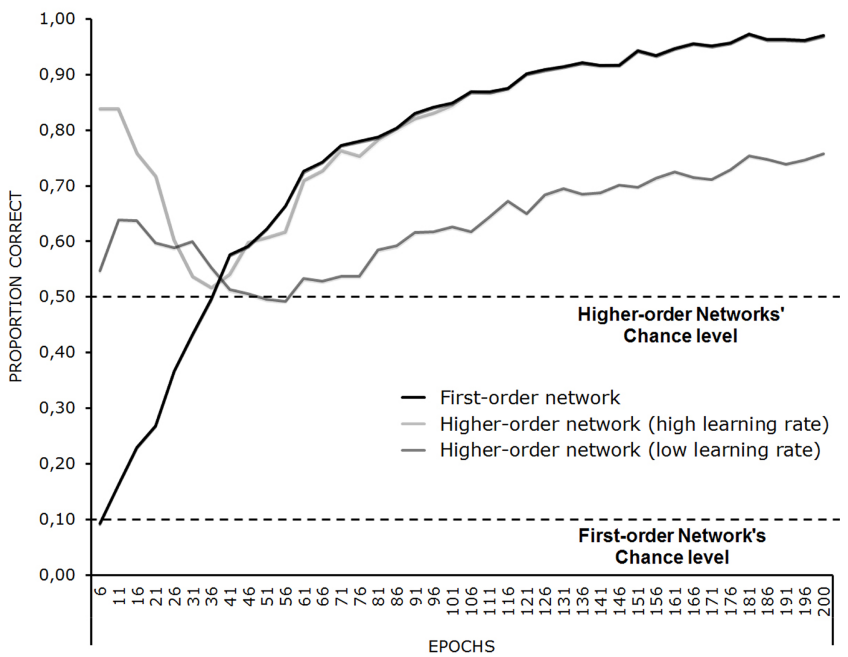


FIGURE 3 | Performance of the first-order and second-order networks, as a function of training expressed as number of epochs.

(2007) in three different domains: Artificial Grammar Learning, Blindsight, and the Iowa Gambling Task. In all three cases, our simulations replicate the patterns of performance observed in human participants with respect to the relationship between task performance and wagering. The blindsight and Artificial Grammar learning simulations instantiate the second approach briefly described above in that they use an architecture in which the processing carried out in second-order network is completely independent from that carried out in the first-order network. In such architectures, the two networks are connected by means of fixed connections that instantiate "comparator units." The Iowa Gambling Task simulation, on the other hand, relies on the same mechanisms as described for the digits task. Interestingly, in this latter case, we were able to additionally capture the fact that asking participants to reflect upon their own performance helps them improve metacognitive awareness (Maia and McClelland, 2004) and hence, the relationship between first-order performance and wagering. The fact that the relationships between first-order and metacognitive performance can vary as a function of task instructions is borne out by a recent study of Fleming et al. (2010) which indicates large individual differences in people's ability to judge their own performance. Strikingly, the authors found that differences in metacognitive ability were subtended not only by differences in the activity of anterior prefrontal cortex, but also by structural differences in the white matter of these regions.

It may seem that the proposed mechanism is identical with signal-detection accounts of metacognition (e.g., Scott and Dienes, 2008). However, there is a crucial difference. Signal-detection accounts typically make the second-order distinction between confidence and guessing (high vs. low wagers) on the very signal that is used for first-order classifications by setting two boundaries on the signal: One boundary that accounts for the first-order classification, and a second boundary (on either side of the first-order boundary) that distinguishes between guessing (cases that fall within the area defined by the second boundaries) and cases that fall outside of these boundaries (on the extremes of the distribution). In such an account, confidence thus depends directly on first-order signal strength (but see Maniscalco and Lau, 2010; Pleskac and Busemeyer, 2010 for further discussion). However, in some of the models we have proposed, the second-order classification does not depend on the same signal as the first-order task. Indeed, instead of wagering high or low based on signal strength, the second-order network re-represents the first-order error as a new pattern of activation. Thus, before it can wager correctly, the second-order network, like the first-order network, has to learn to make a new, single-boundary classification based on this second-order representation (the error representation). Thus, the second-order network actually learns to judge the first-order network's performance independently of the first-order task itself. The difference between our model and Signal-Detection Theory is substantial, for it impinges on whether one considers Type I and Type II performance, that is, first-order and second-order judgments about these decisions entertain hierarchical or parallel relationships with each other. This issue is currently being debated, with some authors defending a dual-route model (Del Cul et al., 2009; Dehaene and Charles, 2010) and others (Lau, 2010; Maniscalco and Lau, 2010) defending hierarchical models. The simulation work described in Pasquali et al. (2010) explored both approaches by means of distinct architectures. Clearly, additional research is necessary to clarify the predictions of each approach and to further delineate their mechanisms.

Beyond giving a cognitive system the ability to learn about its own representations, there is another important function that meta-representations may play: They can also be used to anticipate the future occurrences of first-order representations (see Bar, 2009, on the human brain as a prediction machine). Thus for instance, if my brain learns that SMA is systematically active before M1, then it can use SMA representations to explicitly represent their consequences downstream, that is, M1 activation, and ultimately, action. If neurons in SMA systematically become active before an action is carried out, a metarepresentation can link the two and represent this fact explicitly in a manner that will be experienced as intention. That is: When neurons in the SMA become active, I experience the feeling of intention *because* my brain has learned, unconsciously, that such activity in SMA precedes action. It is this knowledge that gives qualitative character to experience, for, as a result of learning, each stimulus that I see, hear, feel, or smell is now not only represented, but also re-represented through independent meta-representations that enrich and augment the original representation(s) with knowledge about (1) how similar the manner in which the stimulus' representation is with respect to that associated with other stimuli, (2) how similar the stimulus' representation is now with respect to what it was before, (3) how consistent is a stimulus' representation with what it typically is, (4) what other regions of my brain are active at the same time that the stimulus' representation is, etc.

To see how this is different from mere first-order knowledge, consider what happens in the case of hallucination. Imagine a simple three-layers network akin to those described above in which a first layer of units receives perceptual input and is connected to a second layer of internal ("hidden") units that are in turn connected to response units. One can easily train such a simple system to produce specific outputs in response to specific inputs (i.e., activating the "9" unit when presented with the visual pattern corresponding to the digit "9"). After training, each input will cause the emergence of a specific (learned) pattern of action over the network's hidden units, and this will in turn cause a specific response. Crucially, one can now induce a specific response by either presenting a familiar pattern over the network's input units (as it would be in the case of a genuine perception) or by directly activating the network's hidden units with the learned pattern corresponding to that same input (as it could be, for instance, in the case of a memory retrieval whereby the pattern is reinstated by means of other pathways). The point is that the network would respond in exactly the same way in both cases for it simply lacks the ability to identify whether its response was caused by the activation of its input units or by the activation of its hidden units in the absence of any input. In other words, such a network is unable to distinguish between a veridical perception and an hallucination. Doing so would require the existence of another, independent network, whose task it is to learn to associate specific input patterns with specific patterns of activity of the first network's hidden units. That system would then be able to identify cases where the latter exists in the absence of the former, and hence, to learn to distinguish between cases of veridical perception and cases of hallucination. Such internal monitoring is viewed here as

constitutive of conscious experience: A mental state is a conscious mental state when the system that possesses this mental state is (at least non-conceptually) sensitive to its existence. Thus, and unlike what is assumed to be case in HOT Theory, meta-representations can be both subpersonal and non-conceptual.

Overall, this perspective is thus akin to the sensorimotor or enactive perspective (O'Regan and Noë, 2001) and to the general conceptual framework provided by forward modeling (e.g., Wolpert et al., 2004) in the sense that awareness is linked with knowledge of the consequences of our actions, but, crucially, the argument is extended inwards, that is, to the entire domain of neural representations. It can also be extended further outwards, specifically toward social cognition (see also Graziano and Karstner, in press). Our representations of ourselves are shaped by our history of interactions with other agents. Learning about the consequences of the actions that we direct toward other agents uniquely require more sophisticated models of such other agents than when interacting with objects, for agents, unlike objects can react to actions directed toward them in many different ways as a function of their own internal state. A further important point here is that caretakers act as external selves during development, interpreting what happens to developing children for them, and so providing meta-representations where they lack. In this light, theory of mind can thus be understood as rooted in the very same mechanisms of predictive redescriptions as involved when interacting with the world or with one self.

CONCLUSION

Thus we end with the following idea, which is the heart of the "Radical Plasticity Thesis": The brain continuously and unconsciously learns not only about the external world and about other agents, but also about its own representations of both. The result of this unconscious learning is conscious experience, in virtue of the fact that each representational state is now accompanied by (unconscious learnt) meta-representations that convey the mental attitude with which the first-order representations are held. From this perspective thus, there is nothing intrinsic to neural activity, or to information *per se*, that makes it conscious. Conscious experience involves specific mechanisms through which particular (i.e., stable, strong, and distinctive) unconscious neural states become the target of further processing, which I surmise involves some form of representational redescription in the sense described by Karmiloff-Smith (1992). These ideas are congruent both with higher-order theories in general (Rosenthal, 1997; Dienes and Perner, 1999), and with those of Lau (2008), who has characterized consciousness as "signal detection on the mind."

In closing, there is one dimension that I feel is sorely missing from contemporary discussion of consciousness: Emotion (but see, e.g., Damasio, 1999, 2010; LeDoux, 2002; Tsuchiya and Adolphs, 2007). Emotion is crucial to learning, for there is no sense in which an agent would learn about anything if the learning failed to *do something* to it. Conscious experience not only requires an experiencer who has *learned* about the geography of its own representations, but it also requires experiencers who *care* about their experiences.

ACKNOWLEDGMENTS

Axel Cleeremans is a Research Director with the National Fund for Scientific Research (FNRS, Belgium). This work was supported by an institutional grant from the Université Libre de Bruxelles to Axel Cleeremans and by Concerted Research Action 06/11-342 titled "Culturally modified organisms: What it means to be human in the age of culture," financed by the Ministère de la Communauté Française – Direction Générale l'Enseignement non obligatoire et de la Recherche scientifique (Belgium). Portions of this article were adapted from the following publication: Cleeremans (2008), Consciousness: The Radical Plasticity Thesis, In R. Banerjee and B.K. Chakrabati (Eds.), Progress in Brain Science, *168*, 19–33.

REFERENCES

Baars, B. J. (1988). *A Cognitive Theory of Consciousness*. Cambridge: Cambridge University Press.

Bar, M. (2009). Predictions: a universal principle in the operation of the human brain. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 364, 1181–1182.

Brunel, L., Oker, A., Riou, B., and Versace, R. (2010). Memory and consciousness: trace distinctiveness in memory retrievals. *Conscious. Cogn.* 19, 926–937.

Chalmers, D. J. (1996). *The Conscious Mind: In Search of a Fundamental Theory*. New York, NY: Oxford University Press.

Chalmers, D. J. (2007a). "The hard problem of consciousness," in *The Blackwell Companion to Consciousness*, eds M. Velmans and S. Schneider (Oxford: Blackwell Publishing), 225–235.

Chalmers, D. J. (2007b). "Naturalistic dualism," in *The Blackwell Companion to Consciousness*, eds M. Velmans and S. Schneider (Oxford: Blackwell Publishing), 359–368.

Clark, A., and Karmiloff-Smith, A. (1993). The cognizer's innards: a psychological and philosophical perspective on the development of thought. *Mind Lang.* 8, 487–519.

Cleeremans, A. (1994). Awareness and abstraction are graded dimensions. *Behav. Brain Sci.* 17, 402–403.

Cleeremans, A. (1997). "Principles for implicit learning," in *How Implicit is Implicit Learning?* ed. D. C. Berry (Oxford: Oxford University Press), 195–234.

Cleeremans, A. (2008). "Consciousness: the radical plasticity thesis," in *Models of Brain and Mind: Physical, Computational and Psychological Approaches. Progress in Brain Research*, Vol. 168, eds R. Banerjee and B. K. Chakrabarti (Amsterdam: Elsevier), 19–33.

Cleeremans, A., and Jiménez, L. (2002). "Implicit learning and consciousness: a graded, dynamic perspective," in *Implicit Learning and Consciousness: An Empirical, Computational and Philosophical Consensus in the Making?* eds R. M. French and A. Cleeremans (Hove: Psychology Press), 1–40.

Cleeremans, A., Timmermans, B., and Pasquali, A. (2007). Consciousness and metarepresentation: a computational sketch. *Neural Netw.* 20, 1032–1039.

Damasio, A. (1999). *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*. New York, NY: Harcourt Brace and Company.

Damasio, A. (2010). *Self Comes to Mind: Constructing the Conscious Brain*. London: William Heinemann.

Dehaene, S., and Changeux, J.-P. (2004). "Neural mechanisms for access to consciousness," in *The Cognitive Neurosciences*, 3rd Edn, ed. M. Gazzaniga (New York: W.W. Norton), 1145–1157.

Dehaene, S., and Charles, L. (2010). "A dual-route theory of evidence accumulation during conscious access," in *Paper Presented at the 14th Annual Meeting of the Association for the Scientific Study of Consciousness*, Toronto.

Dehaene, S., Kerszberg, M., and Changeux, J.-P. (1998). A neuronal model of a global workspace in effortful cognitive tasks. *Proc. Natl. Acad. Sci. U.S.A.* 95, 14529–14534.

Dehaene, S., and Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition* 79, 1–37.

Del Cul, A., Dehaene, S., Reyes, P., Bravo, E., and Slachevsky, A. (2009). Causal role of prefrontal cortex in the threshold for access to consciousness. *Brain* 132, 2531–2540.

Dennett, D. C. (1991). *Consciousness Explained*. Boston, MA: Little, Brown and Co.

Dennett, D. C. (2001). Are we explaining consciousness yet? *Cognition* 79, 221–237.

Dienes, Z. (2008). "Subjective measures of unconscious knowledge," in

Models of Brain and Mind. Physical, computational and Psychological Approaches. Progress in Brain Research, Vol. 168, eds R. Banerjee and B. K. Chakrabarti (Amsterdam: Elsevier), 49–64.

Dienes, Z., and Perner, J. (1999). A theory of implicit and explicit knowledge. *Behav. Brain Sci.* 22, 735–808.

Dienes, Z., and Seth, A. (2010). Gambling on the unconscious: a comparison of wagering and confidence as measures of awareness in an artificial grammar task. *Conscious. Cogn.* 19, 674–681.

Edelman, G. M., and Tononi, G. (2000). *Consciousness. How Matter Becomes Imagination*. London: Penguin Books.

Farah, M. J. (1994). Neuropsychological inference with an interactive brain: a critique of the "locality" assumption. *Behav. Brain Sci.* 17, 43–104.

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

Graziano, M., and Karstner, S. (in press). Human consciousness and its relationship to social neuroscience: a novel hypothesis. *Cogn. Neurosci.*

Grossberg, S. (1999). The link between brain learning, attention, and consciousness. *Conscious. Cogn.* 8, 1–44.

Hinton, G. E. (1986). "Learning distributed representations of concepts," in *Proceedings of the Eighth Annual Conference of the Cognitive Science Society,* Amherst, MA.

Humphrey, N. (1971). Colour and brightness preferences in monkeys. *Nature* 229, 615–617.

Humphrey, N. (2006). *Seeing Red*. Cambridge, MA: Harvard University Press.

Karmiloff-Smith, A. (1992). *Beyond Modularity: A Developmental Perspective on Cognitive Science*. Cambridge: MIT Press.

Kirsh, D. (1991). "*When is information explicitly represented?*" in *Information, Language, and Cognition*, ed. P. P. Hanson (New York, NY: Oxford University Press).

Kirsh, D. (2003). "Implicit and explicit representation," in *Encyclopedia of Cognitive Science*, Vol. 2, ed. L. Nadel (London: Macmillan), 478–481.

Koch, C. (2004). *The Quest for Consciousness. A Neurobiological Approach*. Englewood, CO: Roberts and Company Publishers.

Kouider, S., de Gardelle, V., Sackur, J., and Dupoux, E. (2010). How rich is consciousness: the partial awareness hypothesis. *Trends Cogn. Sci. (Regul. Ed.)* 14, 301–307.

Kreiman, G., Fried, I., and Koch, C. (2002). Single-neuron correlates of subjective vision in the human medial temporal lobe. *Proc. Natl. Acad. Sci. U.S.A.* 99, 8378–8383.

Lamme, V. A. F. (2003). Why visual attention and awareness are different. *Trends Cogn. Sci. (Regul. Ed.)* 7, 12–18.

Lamme, V. A. F. (2006). Toward a true neural stance on consciousness. *Trends Cogn. Sci. (Regul. Ed.)* 10, 494–501.

Lau, H. (2008). "A higher-order Bayesian decision theory of consciousness," in *Models of Brain and Mind. Physical, Computational and Psychological Approaches. Progress in Brain Research. Progress in Brain Research*, Vol. 168, eds R. Banerjee and B. K. Chakrabarti (Amsterdam: Elsevier), 35–48.

Lau, H. (2010). "Comparing different signal processing architectures that support conscious reports," in *Paper Presented at the 14th Annual Meeting of the Association for the Scientific Study of Consciousness*, Toronto.

LeDoux, J. (2002). *Synaptic Self*. Harmondsworth: Viking Penguin.

Mack, A., and Rock, I. (1998). *Inattentional Blindness*. Cambridge, MA: MIT Press.

Maia, T. V., and Cleeremans, A. (2005). Consciousness: converging insights from connectionist modeling and neuroscience. *Trends Cogn. Sci. (Regul. Ed.)* 9, 397–404.

Maia, T. V., and McClelland, J. L. (2004). A reexamination of the evidence for the somatic marker hypothesis: what participants really know in the Iowa Gambling Task. *Proc. Natl. Acad. Sci. U.S.A.* 101, 16075–16080.

Maniscalco, B., and Lau, H. (2010). Comparing signal detection models of perceptual decision confidence. *J. Vis.* 10, 213.

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

McClelland, J. L., McNaughton, B. L., and O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychol. Rev.* 102, 419–457.

Nagel, T. (1974). What is like to be a bat? *Philos. Rev.* 83, 434–450.

Nelson, T. O., and Narens, L. (1990). Metamemory: a theoretical framework and new findings. *Psychol. Learn. Motiv.* 26, 125–173.

O'Brien, G., and Opie, J. (1999). A connectionist theory of phenomenal experience. *Behav. Brain Sci.* 22, 175–196.

O'Regan, J. K., and Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behav. Brain Sci.* 24, 883–917.

O'Reilly, R. C., and Munakata, Y. (2000). *Computational Explorations in Cognitive Neuroscience: Understanding the Mind by Simulating the Brain*. Cambridge, MA: MIT Press.

Pasquali, A., Timmermans, B., and Cleeremans, A. (2010). Know thyself: metacognitive networks and measures of consciousness. *Cognition* 117, 182–190.

Persaud, N., McLeod, P., and Cowey, A. (2007). Post-decision wagering objectively measures awareness. *Nat. Neurosci.* 10, 257–261.

Pleskac, T. J., and Busemeyer, J. (2010). Two-stage dynamic signal detection: a theory of confidence, choice, and response time. *Psychol. Rev.* 117, 864–901.

Rosenthal, D. (1997). "A theory of consciousness," in *The Nature of Consciousness: Philosophical Debates*, eds N. Block, O. Flanagan, and G. Güzeldere (Cambridge, MA: MIT Press), 729–753.

Rosenthal, D. (2006). *Consciousness and Mind*. Oxford: Oxford University Press.

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

Sandberg, K., Timmermans, B., Overgaard, M., and Cleeremans, A. (2010). Measuring consciousness: is one measure better than the other? *Conscious. Cogn.* 19, 1069–1078.

Scott, R. B., and Dienes, Z. (2008). The conscious, the unconscious, and familiarity. *J. Exp. Psychol. Learn. Mem. Cogn.* 34, 1264–1288.

Seth, A., Dienes, Z., Cleeremans, A., Overgaard, M., and Pessoa, L. (2008). Measuring consciousness: relating behavioural and neuropsychological approaches. *Trends Cogn. Sci. (Regul. Ed.)* 12, 314–321.

Shapiro, K. L., Arnell, K. M., and Raymond, J. E. (1997). The attentional blink. *Trends Cogn. Sci. (Regul. Ed.)* 1, 291–295.

Simons, D. J., and Levin, D. T. (1997). Change blindness. *Trends Cogn. Sci. (Regul. Ed.)* 1, 261–267.

Tononi, G. (2003). "Consciousness differentiated and integrated," in *The Unity of Consciousness: Binding, Integration, and Dissociation*, ed. A. Cleeremans (Oxford: Oxford University Press), 253–265.

Tononi, G. (2007). "The information integration theory," in *The Blackwell Companion to Consciousness*, eds M. Velmans and S. Schneider (Oxford: Blackwell Publishing), 287–299.

Tsuchiya, N., and Adolphs, R. (2007). Emotion and consciousness. *Trends Cogn. Sci. (Regul. Ed.)* 11, 158–167.

Tzelgov, J. (1997). "Automatic but conscious: that is how we act most of the time," in *The Automaticity of Everyday life*, Vol. X, ed. R. S. Wyer (Mahwah, NJ: Lawrence Erlbaum Associates), 217–230.

Weiskrantz, L. (1986). *Blindsight: A Case Study and Implications*. Oxford: Oxford University Press.

Wolpert, D. M., Doya, K., and Kawato, M. (2004). "A unifying computational framework for motor control and social interaction," in *The Neuroscience of Social Interaction*, eds C. D. Frith and D. M. Wolpert (Oxford: Oxford University Press), 305–322.

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

Received: 19 November 2010; paper pending published: 25 February 2011; accepted: 26 April 2011; published online: 09 May 2011.
Citation: Cleeremans A (2011) The radical plasticity thesis: how the brain learns to be conscious. Front. Psychology **2**:*86. doi: 10.3389/fpsyg.2011.00086*
This article was submitted to Frontiers in Consciousness Research, a specialty of Frontiers in Psychology.

frontiers in PSYCHOLOGY

HYPOTHESIS AND THEORY ARTICLE
published: 28 February 2011
doi: 10.3389/fpsyg.2011.00020

Consciousness, plasticity, and connectomics: the role of intersubjectivity in human cognition

Micah Allen*[1] *and* ***Gary Williams*[2]**

[1] *Interacting Minds Project, Centre of Functionally Integrative Neuroscience, Aarhus University, Århus, Denmark*
[2] *Department of Philosophy, Louisiana State University, Baton Rouge, LA, USA*

Edited by:
Morten Overgaard, Aarhus University Hospital, Aarhus University, Denmark

Reviewed by:
David Rosenthal, City University of New York, USA
Giorgio Marchetti, University of Urbino, Italy
Jonathan Cole, Poole Hospital, UK

***Correspondence:**
Micah Allen, Interacting Minds Project, Centre of Functionally Integrative Neuroscience, Aarhus University Hospital, Aarhus University, Århus Sygehus, Norrebrogade 44, Building 10G, 5th floor, 8000 Århus, Denmark.
e-mail: micah@cfin.dk

Consciousness is typically construed as being explainable purely in terms of either private, raw feels or higher-order, reflective representations. In contrast to this false dichotomy, we propose a new view of consciousness as an interactive, plastic phenomenon open to sociocultural influence. We take up our account of consciousness from the observation of radical cortical neuroplasticity in human development. Accordingly, we draw upon recent research on macroscopic neural networks, including the "default mode," to illustrate cases in which an individual's particular "connectome" is shaped by encultured social practices that depend upon and influence phenomenal and reflective consciousness. On our account, the dynamically interacting connectivity of these networks bring about important individual differences in conscious experience and determine what is "present" in consciousness. Further, we argue that the organization of the brain into discrete anti-correlated networks supports the phenomenological distinction of prereflective and reflective consciousness, but we emphasize that this finding must be interpreted in light of the dynamic, category-resistant nature of consciousness. Our account motivates philosophical and empirical hypotheses regarding the appropriate time-scale and function of neuroplastic adaptation, the relation of high and low-frequency neural activity to consciousness and cognitive plasticity, and the role of ritual social practices in neural development and cognitive function.

Keywords: plasticity, consciousness, resting-state networks, development, phenomenology, cognition, culture, intersubjectivity

INTRODUCTION

A constant theme in cognitive science is to define the explanandum of consciousness in terms of qualia or "phenomenal feels," i.e., some ineffable, subjective "what-it-is-like" to experience the world. Moreover, it is often argued that consciousness requires either some kind of higher-order metarepresentation of first-order states (Gennaro, 2004) or that consciousness is itself localized to the pure phenomenal feels or "what-it-is-like" (Dretske, 1993). We contend that the prevailing theoretical spectrum begins from the incorrect assumption that both phenomenal feels and higher-order representations can be collapsed into a single phenomenon. In contrast, we argue that the qualities of phenomenal experience and a subject's higher-order representations of those qualities are separate explananda, while still contending that higher-order representations significantly change the "what-it-is-like" of human experience. This is in accordance with our thesis that reflective consciousness is something that develops in ontogeny and depends upon the plastic individual development of the sensorimotor system in interaction with the default mode network (DMN). Moreover, we contend that both phenomena are highly complex, reciprocally interact, and depend upon the organisms phylogenetic and ontogenetic history of structural coupling between body, brain, and culture.

What drives us to this conclusion? First, the phenomenological tradition, as exemplified by the work of Martin Heidegger and Maurice Merleau-Ponty, emphasizes that our experience of the world is primarily *prereflective* in nature. Accordingly, we would like to construct an account of mental life grounded by the basic insight that cognition is primarily embodied and embedded within an organized environment and social field rather than detached and spectatorial (Stern, 2009). Second, in light of recent evidence of the brain's radical, multisensory plasticity, we will argue that this profound adaptivity at the molecular, network, and systems levels underlies the development and intersubjective function of human consciousness. We will thus argue that both the long-term plasticity underlying skill development and cultural learning and "fast" sensory–motor plasticity underpin our conscious experience of the world and ourselves. Indeed, there are physiological reasons to suspect that both "primary" prereflective processing[1] and "secondary" reflective processing are both dynamic and flexible in nature, grounded in the actual history of the system's encounter with the environment. Whether we are discussing neuron recycling underlying memory consolidation, synaptic reorganization following limb amputation, or alterations in the particular communicative balance between macroscopic neural networks, the old tropes of radical modularism and localization of function are no longer tenable.

Of course, any account of plasticity must face the unique challenge of explaining both "stop" and "go" mechanisms, i.e., the accelerators and "brakes" of cognitive development. While the human brain demonstrates profound plasticity in many domains, it is also the case that two individuals, *ceteris paribus*, will likely

[1] Otherwise known as "online," realtime, adaptive processing.

develop with visual, motor, and auditory cortices in roughly the same place with roughly the same function. Indeed, in many cases, the function of radical neuroplasticity may be a kind of gambler's bargain, in which multiple avenues to reproductive success are made available via profound neural plasticity. Any successful account of consciousness must demonstrate how the brain both adapts to and resists profound environmental, biological, and sociocultural changes.

What is the alternative to the radical representational model or pure sensory–motor account? What kind of system could cope with changing environmental demands with functionally specific responses? "A functional system of this sort would be based on a *dynamic process aimed at achieving an invariant result across changing circumstances*" (Reed, 1996, p. 72). On the picture we are endorsing, we must assume that the organism is dynamically reactive from the very beginning of its ontogenetic history in virtue of autonomous, self-organizing animacy, what Julian Jaynes called *behavioral reactivity*. Interactively learned reflective capacities then further enhance our capacities for action and consciousness while receiving a great deal of inheritance from sensory–motor practice and the phylogeny of the organism.

The role of the neural system then is to *coordinate* or *regulate* the animal's encounter so as to effectively utilize the resources of a changing environment by means of adaptive behavior. On this view, evolution acts upon the *distributed functional system,* leading to adaptive behavior units across a nested set of spatial–temporal scales. Of course, there could not have been cultural evolution if humans had not been able, for example, to see "fire" as something "useful" instead of something that only frightened them. The plasticity of the human mind subserves not only our ability to adapt in the face of shifting challenges, but also to resist them at times, putting the "breaks" on development in order to build individual and collective cognitive "niches." Thus our eventual mastery of fire and other potent environment and mind-shaping tools opened up a world of new biological and cultural stimulants (e.g., nutrients, shelter, bone tools, and hunting rituals) that furthered our particular evolutionary development[2].

This leads to a "transition from models of representation as mirroring or encoding to models of representation as control" (Clark, 1997, p. 47). As Maturana and Varela put it, the patterns of brain activity are not symbolic representations standing in for the stimulus, but rather, state transitions induced by the perturbations or triggering effects of the ecological information embedded within the stimulus. In other words, "Stimuli act upon the organism as control parameters, which upon reaching a certain critical threshold induce a global qualitative discontinuity in the organism (a bifurcation in phase space)" (Thompson, 2007, p. 69). This radical and transitory plasticity of processing is contrasted by the elements of human change that are also partly determined by our own innate biology (e.g., homeostasis, survival instincts), as well as acquired and subjective values (ethical, cultural, religious, etc.). Moreover, a human being does not always change. Sometimes (or even many times) he can resist, decide not to consider stimuli, see them under a different viewpoint, etc. Thus our brain–body systems have evolved to maximize the interaction of reflective and prereflective processes, as well as their interaction and embeddedness within a cultural–intersubjective field. Indeed it is often the case that learned culture and intersubjectivity are themselves the innate control factors restricting our development in particular ways, with examples ranging from "background" factors of social value (race, class, etc.) to the culturally-specific "rules of engagement" for everyday conversation.

Furthermore, we contend that human adults typically engage with the world of objects at a high level of abstraction and linguistic categorization. We are, as Dennett (1993) says, spiders constantly and instinctively spinning our experience into narrative webs that filter and constrain our sensory dynamics. Narrative elaboration appears to be both habitual and open to training at the neural level (Farb et al., 2007). Within eastern traditions, it is common to describe the subject as one who cannot help but assign narrative evaluations to the world of sensory-chaos.

In this paper we develop an account of consciousness that embodies the basic insight that our primary subjective engagement with the world is not constituted by the formal propositional reasoning of Good Old Fashioned Artificial Intelligence (GOFAI), otherwise known as the sense-represent-plan-act model (Wheeler, 2007). However, we do not deny that there are constants in both the primary and secondary processes or that representations play a critical role in cognition and reflective consciousness. Rather, we follow Clark's suggestion that a mature science of mind must invoke both representational information processing and notions of emergent sensorimotor and cultural dynamics (Clark and Toribio, 1994; Clark, 2003, 2008). A variety of phylogenetic factors ensure that through the "cloud" of sensory–motor disorder, certain variables remain more or less constant and transferable within the community by means of joint-attention, abstraction, and categorization.

We will argue that this basic intersubjectivity is critical for human development and underlies our most basic modes of social attunement. Indeed, when discussing the "breaks" on plasticity, it is worth noting that intersubjective practices play a primary role in delimiting an individuals' progress through life. On our account, intersubjectivity is a primary motivating factor for learning through joint action, and depends crucially upon the radically neuroplastic nature of our social-cognitive capacities. Furthermore, we argue that embodied world-directedness and sociocultural cognition are reciprocally related. Research suggests that attending to external stimuli may actively inhibit the kinds of reflective, higher-order thought posited in so-called "theater models" of consciousness (Baars, 1997; Fox et al., 2009).

Accordingly, we contend that recent research on the gross functional connectivity of the human brain sheds light on the problem of understanding both phenomenal feels and meta-consciousness. This new paradigm underlines the importance of ontogenetic plasticity and social–cultural development for determining "what-it-is-like" to be human. We will thus argue that the intrinsic connectivity of the human brain, particularly its organization into discrete anti-correlated networks, supports a view of reflective consciousness as a sociocultural cognitive

[2] Our account thus shares several features with other social–cultural accounts of the ontogeny and functioning of consciousness, see Donald (2001) or Steinberg (2006) for representationally or culturally weighted accounts, respectively.

control mechanism, motivating cortical adaptation and helping to shoehorn individuals into their particular social-cognitive "niche." Delimiting the functional role of the task-negative and positive networks is a step forward in understanding the brain's large-scale organization, and as this special issue suggests, it is not clear why something supervening on a dynamic substrate would not itself be dynamic in nature.

WHAT IS THE EXPLANADA OF CONSCIOUSNESS?

Theorists are divided about the need to rigorously define the concept of consciousness before scientifically explaining it as a natural phenomenon. Some, like Koch (2004), think that "Until the problem is better understood, a more formal definition of consciousness is likely to be either misleading or overly restrictive, or both" (p. 12). Others, like Julian Jaynes, argue "We first have to start from the top, from some conception of what consciousness is, from what our own introspection is. We have to be sure of that, before we can enter the nervous system and talk about its neurology" (p. 18). Because the first approach is bound to improperly delimit the explanandum and thus prove explanatorily evasive, we will follow Jaynes in emphasizing the importance of phenomenologically driven definitions as a mutual constraint on scientific explanation[3]. It is our contention that an empirically sound and phenomenologically driven approach to cognition and consciousness will allow us to begin explaining the enigmatic nature of human subjectivity rather than explaining it away. Moreover, coming to terms with phenomenal experience is at the heart of the solving the mind–body problem and other issues related to the naturalization of subjectivity.

First off, we contend that it is necessary to develop an adequate mental taxonomy of reflective and prereflective experience before attempting naturalistic explanation of such phenomena. Mental taxonomies that distinguish between different levels of consciousness are important because they allow us to trace the phylogenetic and ontogenetic trajectory of reflective consciousness in relation to purely prereflective mentalities such as those driven solely by instinct, perceptual learning, and habit. Mental taxonomies come at a price, however, freezing a shifting phenomena into a static form. Thus we strive to denote specific situations in which alternative functional substrates swap between the "top" and "bottom" position depending upon the context and demands of any given situation (see **Table 1**, below). Even in this formulation it is clear that our innate tendency to explain consciousness in spatial, rather than with temporal or biodynamic metaphors, can be misleading. What sense does it make, for example, to explain conscious control solely in terms of top-down and/or bottom-up processes if both are in reciprocal relation with "bottom" intersubjectivity and "top" executive thought-elements? Before investigating dynamic interrelations of these functional processes, we first distinguish between roughly two different types of mentality in order to better delimit these substrata: those driven purely by prereflective reactivity and those driven by this plus reflective consciousness[4].

Phenomenological reflection indicates that disengaged contemplation of the world's sensory richness is not the foundation of our consciousness. Rather we are typically enacting a phenomenal world that is both shifting and stable in nature. Individuals across cultures share a remarkable degree of perceptual and intersubjective features, yet also display a profound individualism between one another. Thus the challenge for any sociocultural or neurodynamic account of consciousness is to situate our profound adaptivity with our innate phylogenetic inheritance. Thus we observe that across cultures, evolution produces human beings with highly similar gross neuroanatomy, basic perceptual constraints, rearing practices, and

[3] See Varela et al. (1991).

[4] Hobson (2009), Lycan (1997), Schooler (2002), Armstrong (1997), Dewey (1958), and James (1950) all make a distinction along similar lines.

Table 1 | Qualitative differences between prereflective and reflective consciousness, and their interaction.

	General characteristics	Role of plasticity	Time-scale of operation	"What-it-is-like"	Inherits features primarily from
PR	Innate, automatic, embodied, habitual, resistant to perturbations, "online"	Rapid synaptic turn-over, functional adaptation, major development 0–4 years	Short (milliseconds)	Autopilot, flow, seamless, external absorption, effortless, intuitive, extended through tools and the body	Phylogeny
R	Partially learned, socially embedded, can be arrested, sensitive to intersubjectivity, "offline"	Learned Behavior; individual differences, alteration of macroscopic connectome. Major development 4–20 years, throughout life	Long (seconds, minutes, even days and years)	Narrotological, reflective, action-controlling, detached, interiority, folk psychological, calculative, deliberate	Ontogeny
PR R	Intersubjective interaction, individual differences in perception and cognition. Neither "offline" nor "online." Structures PR and R ("pre-noetic")	Influence of local field potentials on global connectivity, rest-stimulus interaction. Develops from birth on	Variable; integration of experience with self-narrative, influence of culture on perception	Smooth expertise without zoning out, integration between online and offline cognition	Both

ontogenetic development. Yet we must also explain the radical plasticity that has enabled humanity to thrive in the face of drastic environmental and cultural–technological shifts[5]. Thus we are able to shift processing from one cortical area to another given minor brain injury or sensory deprivation (Bavelier and Neville, 2002), or retool motion tracking areas for visual word-perception through acquisition of literacy (Dehaene et al., 2010). Moreover, when distinguishing prereflective and reflective consciousness, Merleau-Ponty at times emphasized the dependence of the former on the latter:

> The sensible quality, far from being coextensive with perception, is the peculiar product of an attitude of curiosity or observation. It appears when, instead of yielding the whole gaze to the world, I turn toward the gaze itself, and when I ask myself what precisely it is I see. (Merleau-Ponty, 2006, p. 263)[6]

In some cases, my meta-reflective capacities fundamentally inform my basic experience of the world. We find in other places, also, rich examples where my reflective goals and intentions *pre-noetically* structure what is present in my lived practice, as when Merleau-Ponty describes the footballer in action:

> The field is not given to him, but present as the immanent term of his practical intentions; the player becomes one with it and feels the direction of the "goal," for example, just as immediately as the vertical and the horizontal planes of his own body. (Merleau-Ponty, 2006, p.169)

Thus, although prereflective reactivity in adult humans is often characterized by subpersonal, task-driven "flow" states[7] in which reflective consciousness recedes into the background, we can also find cases in which these states are profoundly impacted by or even co-occur with reflective processes. Julian Jaynes details a complementary case, in which the prereflective engagement actively co-occurs with a detached reflective process:

> My hand, foot, and head behavior...are almost in a different world. In touching something, I am touched; in turning my head, the world turns to me; in seeing, I am related to a world I immediately obey in the sense of driving on the road and not on the sidewalk. And I am not [reflectively] conscious of any of this. And certainly not logical about it. I am caught up, unconsciously enthralled, if you will, in a total interacting reciprocity of stimulation that may be constantly threatening or comforting, appealing or repelling, responding to the changes in traffic and particular aspects of it with trepidation or confidence, trust or distrust, *while my* [narrative] *consciousness is still off on other topics.* (Jaynes, 2000, p. 85, bracketed comments and italics added)

These "zombie" skills suggest that our cognitive system can automatically carry out intentions without the need for meta-conscious oversight, while also demonstrating the subtle dynamics of embedding "top" reflective intentions within active practice. The point is not that automobile drivers are asleep while they drive. Rather, the driver often steers automatically while his or her reflectively conscious mind is ruminating on something else. Thus, their prior and ongoing reflective intentions structure and guide their experience of driving. Similarly, we can see how intersubjective, prereflective elements structure the reflective observation of and interaction with a tool (as in Heidegger), or in Husserl's famous comparison of the first experiences of a Scandic anthropologist in Greenland and those of a naïve tourist. We literally "see intentions" and "experience thoughts" and these explanada are intimately interwoven within one another.

Furthermore, prereflective reactivity is subjective, and through development and interaction comes to be structured by the cultural–linguistic constructs such as the self, the mind, and other folk psychological narratives (Hutto, 2008). We argue that the conceptual categorization afforded by self-reflective folk psychological narratives greatly enhance our capacity for self-reflective action within an internal "mind-space." We thus agree with accounts of language as a form of highly evolved tool use or extended cognition (Tylen et al., 2010). As Andy Clark argues,

> "[T]hinking about thinking" is a good candidate for a distinctively human capacity – one not evidently shared by the non-language using animals that share our planet. Thus, it is natural to wonder whether this might be an entire species of thought in which language plays the generative role – a species of thought that is not just reflected in (or extended by) our use of words but is directly dependent on language for its very existence. (Clark, 1997, p. 209)[8]

In contrast to a continual absorption in realtime temporal dynamics, average human adults with "narratively driven" metacognition are capable of going "offline" to engage in lingual thought-monitoring, deliberative thinking/moral judgment, conscious impulse control, self-consciousness (thoughts about self-image, future, past, etc.), daydreaming, abstract problem solving, reconstructive imagination (visual imagery, internal sketchpad), subvocal rehearsal, rumination, etc. Self-reflexive "ego functions" such as these have been studied by psychologists for decades under headings such as inner speech (Morin, 2005), working memory, thought-monitoring (Frith, 2005), and more recently, mind-wandering (Smallwood and Schooler, 2006; Christoff et al., 2009). From this point forward, we will reserve the term "reflective consciousness" for such offline, decoupled activity. This is done to preserve the phenomenological distinction between low-level sensorimotor cognition and higher-order narrative-driven consciousness. We keep this strong distinction to primarily aid their modeling as dynamic interacting phenomena, not to suggest a fundamental separation or antagonism of their function. Although we review evidence that there exists a fundamental anti-correlation between the neural substrates of online and offline processing, we examine evidence that the strong hypothesis of "pure anti-correlation" is unlikely. Ultimately, we aim to show that while this distinction is useful, it is the dynamic interaction of reflection and action that primarily underpins human consciousness.

[5] See for example, cases in which an innate function such as basic auditory processing may migrate to or co-opt visual processing areas or even vision itself given brain damage (see Bavelier and Neville, 2002 for review) or a consistent alteration in a sensory–motor association (Shimojo and Nakajima, 1981).

[6] On our interpretation, Merleau-Ponty's notion of "sensible quality" is more than just a pure phenomenal quality, but rather, refers to the quality of what-it-is-like to reflect on the sensory stream from a detached perspective.

[7] i.e., those states typical of automatically executed instincts, habits, and skills (Dreyfus, 2002).

[8] Accordingly, we believe that unless you are engaging in second-order dynamics, you are not reflectively conscious at that moment. This is not to suggest that we are somehow asleep when we are not self-reflexive. It just means that there is a qualitative difference between phenomenal feelings and metacognitive reflection on phenomenal feelings.

REFLECTIVE VERSUS PREREFLECTIVE CONSCIOUSNESS

Perhaps the best way to understand the functional role of reflective consciousness in the human cognitive economy is to begin with its counterpart: the fast and efficient attention–salience reactivity system. This vast and intricately connected coordination system has been investigated since the inception of psychology under the rubric of what William James called "automaton theory" and what was later to become taken up and refined as classic behaviorism under Watson and Skinner. In our own time, we have seen the rise of dynamic systems and 4EA theory[9] as an explanatory model for how these automatic, subpersonal processes function so as to regulate our changing response to the environment. A rapidly growing body of research highlights a myriad of "cognitive" functions that are served or even dominated by everyday behavior, the body, intersubjective/interactive processes, and cultural forces. In many cases, these are underwritten by subpersonal body schemas that run automatically and without metarepresentational consciousness of their function[10]. One could say that our mental landscape is pervaded by a prereflective consciousness regardless of whether reflective consciousness is totally absent, co-occurring, or intimately present in the action. Historically, some philosophers (e.g., Huxley) have argued that reflective consciousness as such is epiphenomenal, an "inert spectator" that plays no causal role in the control architecture and can be compared to the steam of a train whistle.

One problem with this view is that it fails to explain the inverse relationship between reflective consciousness and action. As Julian Jaynes observed, "If [reflective] consciousness is the mere impotent shadow of action, why is it more intense when action is most hesitant? And why are we least conscious when doing something most habitual?" (Jaynes, 2000, p. 11, bracketed comment added). Indeed, this question cannot be answered unless we acknowledge the phenomenological reality of effortless prereflective reactivity *and* the effortful deliberations of metacognitive control wherein behavior and attention is modulated by higher-order reflection and narratization. Prereflective flow states, wherein we are "lost" or "absorbed" in the moment of action, are ubiquitous among athletes and other experts who have honed and automatized their skills through continuous training. This training serves a dual purpose: first, to refine automatic processing schemas, second, to enable greater strategic control via adept representational and metarepresentational thought. Thus the human phenomenon of flow is not all or nothing. Even in athletic flow, there are conscious, reflective thoughts and varying degrees of prereflective and reflective awareness of the body, the goal, the action, and so on. Indeed one could give examples in which one is wholly absorbed in a reflective thought, and it is likely that ritualistic mental–social training regimens such as sports and pedagogy are aimed at refining both levels of consciousness as well as their dynamic communication.

Prereflective cognitive processes depend upon ongoing structural transformations in the nervous system which give rise to both simple and complex behaviors across nested sets of temporal scales. According to the 4EA tradition in the philosophy of mind, these subpersonal perception–action cycles are carried out without the use of explicit symbol tokens or "second-order" or "metarepresentational" consciousness. This means that the prereflective, "first order" sensorimotor network is not a Physical Symbol System, but rather, a dynamic, behaviorally reactive and recursive network of sensorimotor connectivity. On this perspective, the system conserves computational energy by using the actual world as a kind of "external memory source" (O'Regan, 1992) to be consulted on-the-fly in response to realtime information needs. As Rodney Brooks famously put it, "It turns out to be better to use the world as its own model" (Brooks, 1991, p. 139).

This has an important upshot for deflating the mind/body problem in regards to how "inner" experience corresponds to an "external world." Strictly speaking, dynamics systems theory claims that for the prereflective system there is no epistemological (or experiential) distinction between "inner" and "outer" because the brain is not modeling the environment in low-level perceptual guidance, but rather, responding to it in terms of its ontogenetic history of structural coupling as a brain–body system. As Maturana and Varela (1987) argued, simple surgical experiments on the ontogenetic development of control assemblies suggest that in regard to automatic, task-oriented behavior such as a frog catching prey "there is no such thing as up and down, front and back, in reference to an outside world, as it exists for the observer doing the study. There is only an *internal correlation* between the place where the retina receives a given perturbation and the muscular contractions that move the tongue, the mouth, the neck, and, in fact, the frog's entire body" (Maturana and Varela, 1987, pp. 125–126) Accordingly, we can reject the Cartesian assumption that first-order behavioral dynamics depend on the sense-represent-plan-act model (Wheeler, 2007). Instead, dynamic systems theory claims that

> intentions are [best] seen as grounded in neural patterns.... Decisions are precisely the brain's falling into one pattern or another, a falling that is modeled as the settling into a basin of attraction that will constrain neural firing in a pattern. There is no linear causal chain of input, processing, and output. Instead there is a continual looping as sensory information feeds into an ongoing dynamic system, altering, or reinforcing pattern formation. (Protevi, 2009, p. 18)[11]

Strictly speaking, the continual operation of first-order cognitive dynamics is subpersonal and not conceptually structured by mentalistic metaphors such as the inside/outside, subject–object, mind–body schema (Lakoff and Johnson, 1980, 1999)[12]. In other words, primary prereflective processes are unaccompanied by metacognitive reflection or the experience of a private "introcosm" or "theater" inside our heads[13]. Based on neuropsychological evidence, we argue that this inside/outside, subject–object schema operationalizes whenever we engage in "introspective gazing" or "reflection"[14].

[9] i.e., the "embodied, embedded, extended, enactive, affective" tradition.

[10] See, for example, Gallagher (2005).

[11] See Freeman (2005).

[12] It is crucial to note that subpersonal cognition is still "subjective" insofar as there is something-it-is-like to experience first-order cognitive dynamics, but this subjectivity is not structured by narrative consciousness and autobiographical memory.

[13] Or hearts, solar plexus, or whatever other folk metaphor of the time might be in play.

[14] Recent lines of research (see Clark (2008), ch. 3 for an overview) suggest the construction of attentional objects such as our own thoughts and perceptions, as well as those of others, directly influences our reflective cognitive system, and vice-versa. Self-reflexive linguistic cognition grounded in lexical metaphors opens up the possibility for a psychological distance between ourselves and the immediate actions of our body that gives rise to novel modes of self-regulation based on the interaction between conscious narratization, imagination, executive control, and working memory (especially the episodic buffer).

Metaphorically speaking, reflective consciousness operationalizes[15] whenever we sharply focus our mind's eye on experiential and sensory qualities rather than being mindlessly "absorbed" into the usability of affordances.

Our approach is thus similar to higher-order perception (HOP) and higher-order thought (HOT) theories. However, we differ significantly insofar as we do not invoke higher-order representations to explain phenomenal feels, but rather, to explain narrative consciousness and self-reflexive cognition. In contrast to higher-order theorists, we do not think higher-order representations are needed to explain phenomenal consciousness (the "what-it-is-like" of an organism). Instead, we think all organisms have a "what-it-is-like" insofar as they are living, embodied beings. However, we do contend that higher-order representations change the what-it-is-like of human cognition to such an extent as to radically change the phenomenal qualities of experience, giving rise to new forms of narratological subjectivity.

Accordingly, we contend that this special psychological interiority or "mind-space" does not correspond to a metaphysical substance (or ghostly process) as assumed by Cartesian dualists, but rather, to a virtual (i.e., temporary and easily dissoluble), analogically constructed "workspace" or "global theater" which acts as a "facility for *accessing, disseminating, and exchanging information*, and for *exercising global coordination and control*" (Baars, 1997, p. 7). Focused introspection upon pure sensation by means of reflective consciousness is itself a metacognitive skill that fundamentally changes our "what-it-is-like," uniting the dynamic sensory–motor processing with the profoundly cultural reflective lens. It is in these unique cases in which the vibrant individual differences in consciousness and metacognition are most manifest. We thus contend that a dynamic systems approach coupled with well-established "theater models" of reflective consciousness gives reasoned answers to philosophical quandaries concerning qualia and subjectivity. Phenomenology suggests that complete absorption into task-oriented, world-directed mental states is often unaccompanied by introspective thought-monitoring or autobiographical memory storage. Instead of being "self-present," we often seem to be "away from ourselves" and "empty minded" when absorbed in the world at large.

Examples like the truck driver or trained meditation practitioner illustrate the oscillatory nature of self-other, reflective–prereflective processing networks. When absorbed in the world, we often find ourselves "coming back" from selfless states unaware of the temporal gaps in consciousness. Indeed, "[narrative] consciousness knits itself over its time gaps and gives the illusion of continuity" (Jaynes, 2000, p. 25, bracketed comment added). Furthermore, research on change blindness suggests that we are often deluded about the level of detail available for report in our episodic memory. In reality, narrative consciousness is able to access only a fraction of what stirs beneath it[16]. In other words, our reflective experience of perceptual detail is a top-down[17], virtual construction based on autobiographical memory. As Alva Noë puts it, "To experience detail virtually, you do not *need* to have all the detail in your head. All you need is quick and easy access to the relevant detail when you need it" (Noë, 2004, p. 50).

Nevertheless, we delude ourselves into thinking that we have a rich picture "inside" our heads when perceiving the world. We suggest that this is a side-effect of language turning experience itself into an object of understanding, amenable to folk psychological metaphors steeped with dualistic presuppositions about the continual presence of consciousness for the control of "rational" thought and action. Reflective consciousness seems to pervade our experience because our mental metaphors are structured by the concept of rational access and control, i.e., the "I." Indeed, our entire autobiographical language is centered around a culture in which the ineffable "self" is both container and director of our experiences. But as Jaynes says, "[Narrative] consciousness is a much smaller part of our mental life than we are conscious of, because we cannot be conscious of what we are not conscious of" (Jaynes, 2000, p. 23, bracketed comment added). This limitation of access is built into the basic structure of conscious introspection, but it is not a burden. Rather, our ability to package prereflective states into increasingly complex and useful representational and metarepresentational "tools" represents a decisive factor in our sudden departure from prelinguistic animals (Tylen et al., 2010).

We contend that appreciation of such facts suggests that the subjective experience of non-human animals is subjective in the same way that the empty mindedness of long distance truck drivers is subjective. Although the truck driver may drive "mindlessly," engaging only the basic sensorimotor subroutines needed to drive safely, it is her automatic conscious rumination that may remind her of an important forgotten task or keep her attention at the road regardless of her fatigue. The real trick then, is to understand how our reflective, linguistic consciousness of the world structures our prereflective engagement with the world. Without the possibility of making mental experience an explicit object of attention by means of a linguistic, self-reflexive "tag" with special experiential associations of interiority, most non-human animals are unable to create the necessary psychological distance from their actions to construct the reflective "introcosm" familiar to humans when they turn inwards upon the "hidden hermitage where we may study out the troubled book of what we have done and yet may do" (Jaynes, 2000, p. 1).

As we have reviewed, an inspection of phenomenology reveals a few basic structural elements of both reflective and prereflective experience. Crucially, consciousness is both hierarchical and dynamic: we are not always reflective or prereflective in nature, but rather, constantly shifting between these poles of reference. In examining the kinds of cases explored by Merleau-Ponty and Heidegger, we are presented with examples in which particular actions or stances are mediated by corporeal states, prior intentions, future thinking, and present-oriented action. For humans moving through the world then, consciousness is not any one static state or achievement, but rather a coordinated movement through various interrelated states of representation and dynamic world-body exchange. As we will show in the next section, the neurobiology of social interaction plays a crucial role in the development of these conscious capacities.

[15] To operate literally means "to perform a function; to exert power or influence." When consciousness operationalizes, it exerts influence over the rest of the brain. This is meant to indicate how consciousness is temporary and easily dissoluble.

[16] See Wegner (2002).

[17] Or better still, top-top (see Roepstorff and Frith, 2004).

THE NEUROPHENOMENOLOGY OF SOCIOCULTURAL CONSCIOUSNESS

Does this phenomenological taxonomy have any grounding in cognitive neuroscience? Recently there has been a great deal of work examining the properties of macroscopic neural networks when subjects are left "task-free" in a brain-scanning environment (Raichle et al., 2001; Fox et al., 2009). These so-called "resting state" functional magnetic resonance imaging experiments (rsfMRI) consistently find statistically coherent relationships within and between gross neuroanatomical networks that are correlated with, and anti-correlated with, a variety of functions relevant for the development and functioning of consciousness. These networks include the DMN, salience network (SAL), and central-executive network (CEN). As we will primarily focus on the function and anatomy of the DMN, it is helpful to briefly review the latter two networks.

Originally described simply as the "task-positive network," attention system, or executive network, traditional, and rsfMRI have lead to the description of two independent but closely related exogenous neural networks. These are deemed "task-positive" due to their consistent tendency to activate during cognitively demanding tasks and deactivate at rest. The CEN refers to the top-down dorsal attention network associated with the online control of behavior, and includes the dorsolateral prefrontal cortex, frontal eye fields, dorsal medial-prefrontal cortex (MPFC), intraparietal sulcus, and superior parietal lobule. The SAL refers to a more ventral network of regions involved in the automatic the detection of error, somatosensory awareness, and the detection of salient non-target stimulus. The SAL network is made up of the dorsal anterior cingulate cortex, frontoinsular cortices, amygdala and ventral midbrain (Buckner and Vincent, 2007[18]; Carhart-Harris and Friston, 2010[19]). Collectively these networks show some degree of integration, both being anti-correlated with the DMN at rest, yet also retaining a significant degree of functional non-overlap.

To better understand the relevance of rsfMRI, let us briefly review its historical development. In the mid-1990s, during the initial development and boom of social-cognitive neuroscience, researchers consistently observed deactivations of the MPFC during "task-positive" conditions (Damoiseaux and Greicius, 2009; Fox et al., 2009). In other words, it appeared that whenever subjects were required to complete tasks requiring focused attention and continual cognitive-executive effort, the MPFC would deactivate. Early controversy revolved around whether this deactivation was merely relative to task-induced positive activations or a "true" deactivation of this area from baseline. In the decade since these early discoveries, a great body of neurophysiological research has gone underway, revealing that these deactivations are not the product of relative activation ratios, but rather, are likely to reflect true task-induced neurophysiological decoupling, interaction, and deactivation depending upon the nature of the task or resting "state" (Greicius and Menon, 2004; Esposito et al., 2006; Buckner and Vincent, 2007; Jerbi et al., 2010). Why are deactivations in the MPFC relevant for the study of both prereflective and reflective consciousness?

First, the MPFC has been implicated in particular types of mental representation (Frith and Frith, 2006), theory-of-mind (Frith, 2007), and narrative processing (Mar, 2004; Mano et al., 2009). The MPFC has long been implicated in social cognition tasks, and can be further subdivided into areas associated with cognitive (posterior region of rostral MFC), affective (anterior rostral MFC), and task outcomes associated with punishment or reward (orbital MFC; Amodio and Frith, 2006). The MPFC is thus uniquely situated to mediate between top-down cognition and bottom-up reward-salience cues, and has been hypothesized as a necessary area for the representation of information in a conscious, socially communicable format (Frith and Frith, 2007). The MPFC remains highly plastic throughout childhood and adolescence, reaching biological maturity during late adolescence (Gogtay et al., 2004; Blakemore, 2008). Further, the development of theory-of-mind is considered a significant landmark in cognitive development that can be accelerated or inhibited by environmental factors (Jenkins and Astington, 1996).

Due to the prominent role of the DMN in these domains as well as within episodic memory, early researchers hypothesized that executive function tasks might require an inhibition of "stimulus-irrelevant thoughts" (SITs) or mind-wandering. Simply put, given the boring nature of sitting prone within a magnet, experimental participants might naturally engage in social-cognitive mind-wandering between the rigorous experimental trials. As task conditions appear to inhibit these networks, it was hypothesized that mind-wandering might constitute a DMN (Raichle et al., 2001). Initially, this formulation was controversial as it is not entirely clear what it might mean if laying in the scanner doing nothing was to be considered a "true baseline" for all cognitive tasks[20].

More methodologically, many researchers complained that due to the slow wave (<0.1 Hz, or about one cycle every 10 s) nature of resting-state connectivity, cardiovascular, and respiratory noise could not be excluded as causes of the apparent resting-state networks (RSNs). Given that initial RSN findings did not adequately control for these confounding variables, and the fact that it remains unclear exactly how the BOLD signal maps onto neural activity, many expressed doubt that these low-frequency fluctuations (LFFs) represented actual neural phenomenon with a functional counterpart. These concerns have since been largely dissuaded due to growing evidence that RSNs are not confounded by physiological noise and do in fact represent cross-culturally replicable[21], robust phenomena of neurophysiological origin (Mantini et al., 2007; Van Dijk et al., 2010). Recent research has demonstrated that the slow-wave oscillations indexed in fMRI that characterize the DMN can be verified as neurophysiological deactivation during task performance via intracerebral EEG (Jerbi et al., 2010). Although it is still unclear exactly what role the RSNs play in cognition, memory, and perception, researchers have begun to converge around some common theories of DMN function.

Since Raichle's early exposition of the DMN, most researchers have moved beyond describing the resting state as a "true baseline." Raichle and Fox, two prominent researchers, have independently stated that

[18]http://www.frontiersin.org/consciousness_research/10.3389/fpsyg.2011.00020/full#B11

[19]http://www.frontiersin.org/consciousness_research/10.3389/fpsyg.2011.00020/full#B12

[20] See Morcom and Fletcher (2007) for a review of rational against the "baseline" hypothesis.

[21] See "the human connectome project" (Biswal et al., 2010) for recent findings collected from approximately 1,800 individuals across 32 international neuroimaging centers.

the spatiotemporal coherence exhibited by spontaneous DMN activity is unlikely to reflect a baseline, but rather, is indicative of deep, almost architectural features of reflective consciousness. This "Mariana's Trench" view of the DMN supports recent theories that relate hyper or hypoactivity connectivity within the network to thought disorders like schizophrenia and OCD. Thus it is not our claim that the DMN is the direct neural correlate of consciousness, as it is likely that the experiences of ruminative reflection and sensory–motor consciousness do not reduce to these low-frequency oscillations. Further, significant evidence of relatively developed (albeit significantly altered) DMN-like connectivity in the brains of pre-term infants (Fransson et al., 2007), and anesthetized adults (Greicius, 2008) lend credence to the pre-noetic view of DMN function. Rather, our claim is that it is the interaction of these networks with the "task-positive" environment and history of the organism out of which social-cognitive consciousness arises. We thus predict that the functions of the DMN are crucial for the conscious integration of experience in self-narrative and interaction, setting the spatiotemporal and intersubjective associations between experiences, both in terms of autobiographical and episodic memory and pre-potent prereflective tendencies. This mapping is further reflected by the distributed set of functions represented by individual nodes of the DMN.

The functions of the individual nodes of the DMN are many, but do appear to share a common, self-referential theme. Comprised primarily of the MPFC, posterior cingulate cortex (PCC), and areas of the inferior parietal lobule (IPL), these regions are implicated in theory-of-mind, belief-understanding, and episodic memory generation. In line with this hypothesis, a recent quantitative meta-analysis of the DMN implicated the network in tasks involving autobiographical narrative, prospective memory, and theory-of-mind (Spreng et al., 2009). A vast array of studies have now linked particular patterns of resting connectivity between these nodes and psychopathologies including ADHD, OCD, depression, and schizophrenia, commonly construed as self-related reflective thought disorders (Greicius, 2008). It has become clear from this research that an individual's particular "connectomics" amongst default mode and other macroscopic neural networks is crucial for the functioning of bottom-up and top-down conscious processes.

Although a precise pattern has yet to emerge from these studies, a common theme throughout is hyper or hypoconnectivity of the DMN at rest, often correlated with measures of ruminative thinking, obsessive urges, or severity of symptoms and abnormalities in DMN deactivation during task-positive experimental conditions. It would thus appear that the "defaultness" of the DMN lies in its comprehensive involvement in tasks involving either the regulation of reflective thinking, the temporal structure of these thoughts, or their somatic references. It is our argument that this reflective, regular, and oftentimes unintentional narrative stream organizes experience into cohesive memories, facilitates action planning, and coordinates the joint action necessary for successful cognition and the ontogenetic development of consciousness.

Today, many researchers have moved beyond discussions centering on the "default" conception, moving instead to the "resting-state network" terminology. This new paradigm highlights the importance of an individual's "functional connectome" for the emergence of a particular consciousness (Buckner, 2010). The connectome approach emphasizes the relative contribution of a given functional localization, denoting that these "task-induced" activities are in a fine-grained dynamic relationship with the ongoing slow-wave activity of the macroscopic brain networks. This is to say that, while certain kinds of reflective and online experience evoke "fast" neural activity, over time the pattern of these excitations leaves its mark upon the "slow" wave, altering how the network processes future information. Indeed task-evoked activity in the medial-frontal gyrus (MFG) during a visual face/place categorization task has been shown to predict subsequent MFG to occipital place areas (PA) low-frequency coupling. In this experiment, the degree of post-task MFG–PA coupling significantly predicted post-scan memory performance (Stevens et al., 2010).

Experiments in animals and humans have revealed similar "rest-stimulus" interactions in which the degree of task-evoked activity can be predicted by prior LFFs or vice versa (See Northoff et al., 2010 for review). Thus, our gradual neuroplastic enculturation predicts reflective processing and our effortful attention to the environment. A great body of research has thus revealed that deactivations of the DMN are intrinsically anti-correlated with the CEN and SAL networks (Sridharan and Levitin, 2008; Carhart-Harris and Friston, 2010). Setting aside considerations of the "defaultness" of the DMN, the finding of these distributed networks and their anti-correlation is immediately interesting for the neurophenomenology of both prereflective and reflective consciousness.

We contend that given the MPFC's crucial role in social-cognitive tasks, including social priming and desirability, the DMN is clearly highly susceptible to stimulation by means of ostensive social cues and social interaction (Schilbach et al., 2006) while also contributing to the reflective stream of conscious thought (Christoff et al., 2009). The MPFC and TPJ sub-nodes of the DMN have been repeatedly implicated in the processing of emotional faces, threat behavior, ToM cartoons, and other task involving the processing of social-cognitive cues (see Frith and Frith, 2007 for review). Given the dual role of the DMN in constraining exogenous neural activity and processing these cues, we hypothesize that the network is involved in both the reflective regulation of social behavior and the bottom-up processes that determine what is "salient" for the prereflective consciousness.

Although the actual detection of and pre-potent response to rapid salient cues is likely mediated and directly processed within the salience and CEN networks, as is the online control of behavior, we argue that it is the MPFC that sets the social "frame" for what is salient and preloads the repertoire of socially appropriate gestures and concepts needed in an online interaction. From a narrative perspective of folk-psychology, this is akin to the interpretation of ones place as an agent within an unfolding interaction. Thus it is the role of DMN-supported self-rumination to evaluate and determine my place within a social hierarchy, to communicate what "set" or context I am to the SAL and CEN networks, and to play an active role in "writing" my own current and future position within a social folk psychological narrative.

In interaction, my ruminative reflection is thus extremely important for the regulation and evaluation of the unfolding scene; yet I must also be detached and authentic, fully automatic in my response. It is the smooth, sometimes anti-correlated interaction of the DMN with the SAL and CEN that enables our seamless interaction with others, underpinning intersubjective behavioral learning (see **Figure 1**, below). Further, we argue that this highly embedded social-cognitive nature (Schilbach et al., 2008) and high neural

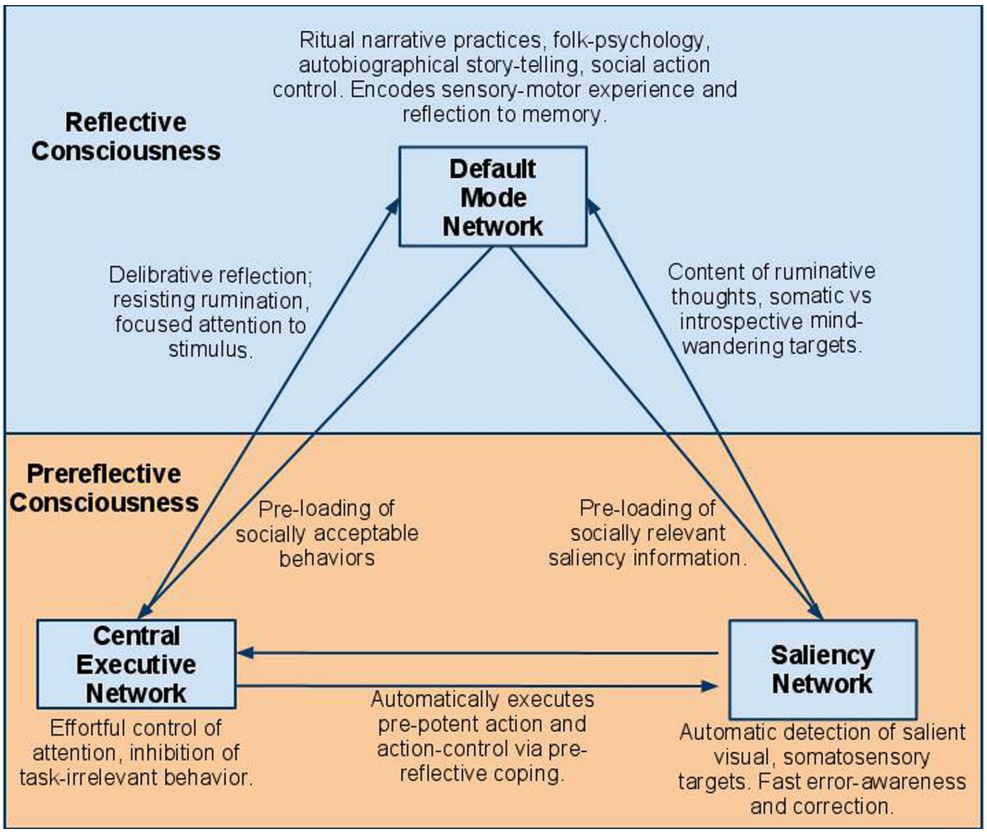


FIGURE 1 | Diagram of resting-state network mappings to prereflective and reflective consciousness. Arrows represent interactions between networks. DMN includes medial-prefrontal cortex, the posterior cingulate cortex, the inferior parietal lobule, the lateral and inferior temporal cortex, and the medial temporal lobes. CEN includes dorsolateral prefrontal cortex, frontal eye fields, dorsal medial-prefrontal cortex, intraparietal sulcus, and superior parietal lobule. The Saliency Network (SAL) includes the dorsal anterior cingulate cortex, frontoinsular cortices, amygdala, and ventral midbrain.

plasticity of the DMN through development (Gogtay et al., 2004; Blakemore, 2008) highlights an important feature of the human cognitive-executive-salience system: social interaction and enculturation are central motivators of plastic brain adaptation.

PLASTICITY: CRUCIAL FOR SOCIOCULTURAL CONSCIOUSNESS

The brain is an evolving, dynamic system. Human brains have often been described as proportionally larger than those of our simian counterparts. Recent research, however, appears to overturn this claim: the human neocortex is not larger than that of chimpanzees when body-size is controlled for. Comparative studies find that prefrontal white matter alone differentiates apes and humans (Schoenemann et al., 2005). White matter is highly plastic, with myelination increases occurring within as few as 11 h of body-awareness training (Tang et al., 2010) and remains highly variable throughout development (Blakemore, 2008) and the lifespan[22]. Finally, local connective tissue appears to possess the greatest level of plasticity, with mappings between synaptic terminals potentially undergoing complete remodeling in as few as 4 weeks (Stettler et al., 2006).

As cortical neuroplasticity in healthy human adults gradually begins to gain widespread acceptance in the scientific community[23], we have begun to realize that a wide variety of psychopathologies can be described in terms of their impact on cortical connectivity. Through the newly developed methods of diffusion-tensor imaging and functional connectivity analysis, diseases that previously resisted neuroimaging classification have begun to reveal the complex way in which the human connectome shapes cognition. To give a simple example, imagine the physiology of focal stress. Given recent connectivity research, the experience of a stressful event can no longer be described as a focal disorder of, for example, stimulus processing, or ruminative thoughts. This is precisely because stress is not localized to any single region or function of the brain. Indeed, the neural stress response supervenes on a distributed system including multiple brain areas, the body, social roles, and the hormone system. The experience of social and mental stress depends upon both the sociocultural history of the actor and the subsequent entrainment of an agent's macroscopic resting networks.

[22] It is worth noting here, though we will not explore the issue further in this paper, that plasticity can be sub-divided into mechanisms corresponding to effortful learning (active) and those associated with recovery from damage or maturation (passive). Here we shall refer to both as "plasticity" and discuss a variety of mechanisms related to both sub-divisions.

[23] See Rubin (2009) for an excellent historical review of neurogenesis and plasticity.

What evidence is there regarding the kind of radical neuroconnective plasticity we are arguing for? By radical neuroconnectivity, we refer to fast adaptations at the cellular, molecular, functional, or anatomical level in response to training and experience. We do not deny that much of neural development is shared across cultures and persons; *ceteris paribus* two individuals will typically develop nearly identical gross neural anatomy[24]. Yet, the connective plasticity of the brain is so radically dynamic, that primary sensory cortices can "take-over" one another given damage (Bavelier and Neville, 2002) or reverse function to incorporate a 180° flip in vision (Shimojo and Nakajima, 1981). Cutting edge neurobiological imaging technologies have now revealed synaptic button turn-over rates of close to 7% per week (Stettler et al., 2006). If this rate holds constant across the axon, the entire synaptic-connective model of a given neural pathway could be remodeled within 3–4 weeks! As little as 4 h of high frequency trans-cranial magnetic stimulation (TMS) of the auditory cortex caused significant thickening of the auditory cortex, both ipsi and contralateral to the sight of the stimulation (May et al., 2007). This thickening, correlated with increased performance on an auditory discrimination task, vanishes 3 days without stimulation.

More recently, research by Tang et al. (2010) reveals that as little as 11 h of body-focused meditation results in significantly increased fractional anisotropy[25] (FA) of anterior cingulate connective pathways, a critical pathway for interaction between CEN and SAL network functions such as behavioral inhibition and action selection. Finally, voxel-based morphometry before and after medical school exams reveals increased neural density in hippocampal learning-associated areas (Draganski et al., 2006), and 3 weeks of Tetris training in young girls revealed functional and structural differences in working memory and spatial processing areas (Haier et al., 2009). Given that working memory and spatial processing are clear predictors of successful language and career learning, and given the implicit impact of executive function networks upon the DMN and vice versa, a mere 3 weeks learning to play Tetris is clearly enough to improve communication between the strategic and sensory–motor elements common to the game. Clearly the brain adapts to its training, but to what degree is this effect realized in the social-cognitive and default mode areas? Do the rigorously ritualistic group-actions we engage in produce systematic alterations in brain structure?

In a landmark study, Gogtay et al. (2004) constructed "movies" of pediatric brain development through the use of MRI repeated every 2 years on 13 healthy children, aged 4–21, over a period of 8–10 years. Analysis of gray matter density changes across the whole brain revealed a striking pattern of neural pruning, with primary association areas being the first to mature, followed by secondary association, frontal, temporal, orbitofrontal, and dorsolateral prefrontal cortex. It is interesting to note that social-cognition associated areas reached maturity prior to classical executive function areas (e.g., DLPFC, rACC) suggesting that these more metacognitive functions may rely on the establishment of social-cognitive mechanisms.

Follow up studies conducted at multiple sites further establish white matter increases throughout the frontal and parietal cortices, coupled with regionally specific increases and/or decreases in gray matter throughout adolescence (Blakemore, 2008). The MPFC undergoes significant changes throughout adolescence, demonstrating altered response to faces and social cues (Blakemore, 2008). Studies of the default mode reveal that although LFF networks can be found even at birth (Fransson et al., 2007) and although they are generally well connected between hemispheres, they remain sparsely interconnected throughout childhood and into early adolescence, gaining maximum connectivity during early adult hood before declining again through late development (Fair et al., 2008). Within elderly populations, differences in DMN structure and function differentiate those with mild Alzheimer's from healthy age-matched controls (Greicius et al., 2004).

SHIFTING NETWORKS AND INDIVIDUAL DIFFERENCES

As argued previously, if we want to maximize our explanatory power, then reflective consciousness must be distinguished from prereflective consciousness. We thus argue for a view of consciousness as the embodied, ontogenetic development of balanced interactions between exteroceptive targeted sensory–motor processes and interoceptive mental representational and cultural forces that are particular to individuals and their sociocognitive history. This distinction is supported by the organization of the brain into endogenously anti-correlated neurological networks for exogenous and endogenous processing, as well as their particular interactions and tendency toward plastic adaptation to the social environ.

What does it mean for neural networks to be endogenously anti-correlated? Although the exact nature of LFF anti-correlations remains unclear, recent experiments have demonstrated both that natural anti-correlations (i.e., systematic complementary deactivations and activations between RSNs) have a neural underpinning (Jerbi et al., 2010) and are consistently predicted by task-difficulty and stimulus demands. However, other fMRI research on experience sampling during sustained attention and introspection tasks (Christoff et al., 2009) reveals that "pure anti-correlation" is less likely than the task and context-dependent distribution of these network's activity, in that these networks are probably not strictly antagonistic, but rather are distributed in their allocation depending on task and context. Thus in some cases the DMN and CEN may exhibit clear anti-correlation (for example during extremely attention-demanding tasks) whereas for various types of social-cognitive or introspective tasks, the DMN may actually coactivate with the CEN or SAL networks depending on task demands.

Given the particularly dynamic and individual nature of these networks, we hypothesize that human consciousness is crucially dependent upon an individual's particular balance of intrinsic and extrinsic brain networks. We do not concede that consciousness is reducible or localizable with these networks, but rather exists only when actively coupled with an individual's sociolinguistic and ontogenetic history. Both prereflective and reflective consciousnesses are, on our account, multiply realizable and constrained by individual differences. The natural plasticity and functional adaptability of both brains and human social networks support this view. We are not born "ready-to-go," but rather, must undergo specialized socially interactive brain training in order to fulfill the massive

[24] Although do note that how "nearly identical" gross neural morphology is between subjects and populations is a primary assumption in many imaging analysis techniques (Friston et al., 1994) and is a topic of heated debate, including evidence that over-prevalent "WEIRD" sampling in neuroimaging has led to a systematic bias in results due to the exclusion of Asian and African-American participants in many studies (Isamah et al., 2010).

[25] A measure of white matter integrity related to axonal myelination.

distribution of niche-specializations that is unique to contemporary society. A growing body of evidence highlights a primary role for social interaction in health and brain development (Dickerson and Kemeny, 2004; Holt-Lunstad et al., 2010). To further elucidate the relationship between prereflective and reflective consciousness in terms of RSNs, consider the following examples.

Any given individual's morphogenetic history will be fundamentally shaped by both interaction with others and genetic inheritance. Interaction serves multiple neurologically formative purposes: the imitation of motor behaviors, the recitation of heuristic habitual behaviors, and the linguistic navigation of social encounters. Human behavior is characterized by a fundamental reliance on highly ritualized skills that are essentially designed to adapt the brain to any given environmental niche. Thus, contemporary society demands that children must be repetitively taught to brush their teeth, pick up after themselves, make their beds, eat well, be nice to others, adhere to appropriate gender norms, and other traits of polite society. In adolescence we require initiation ceremonies, participation in sports and group activity, exercise of the body, courtship behaviors, ritualized practice of mathematical and scientific reasoning, and so on.

It is this rich intersubjective history that allows us to make meaningful choices, deciding to shape our lives through the gradual enculturation of our malleable neural cortex. Yet this plasticity also has a consequence: the socioeconomic factors and small nonconscious habits we ritualistically ascribe to (e.g., "I don't know why, I just don't enjoy X without Y") have lasting consequences on our future possibilities for action. The choices we make may not be "free," yet they are highly consequential for our development and are grounded in our social lives. Thus the most important choices in life do not revolve around the decision to brush ones' teeth or take a meal; these things are entrained such that the truly important (and fundamentally social) decisions, like becoming vegetarian or pursuing a new career far from home come into focus, both in our cultural practices and most cherished memories. Throughout adulthood we require the training of a highly specialized set of ritual behaviors and speech patterns, whether those be preparing sales reports or flipping burgers. The social developmental of functional units of behavior is fundamentally geared toward creating particular types of brains. As John Protevi puts it, "subjectivity [is] an emergent capacity of bodies when they are placed in the appropriate subjectification practices" (Protevi, 2009, p. 31).

A common refusal here might be to say that all of these behaviors require the mere learning of scripts, and do not need to invoke neuroplasticity beyond that required for the development of a functioning language module. This view fails on two counts. Given that activities like studying for exams, learning to juggle, playing a videogame, or learning to meditate all induce structural and functional brain alterations across a wide variety of cognitive–sensory domains[26], it seems unlikely that developmental plasticity is entrained solely within linguistic learning modules. Second, the goal of sociocultural training is not in fact the reflective representation or meta-conscious analysis of a given task-set. Rather, the training to brush ones' teeth, drive to work, or perform courtship rituals is in fact aimed at training the interaction and tempo of interaction between the prereflective sensorimotor and reflective-narrative systems. Successful parenting means that a child will not need to remember to brush her teeth; rather she will automatically do so every night at a given time.

That is to say, regardless of if she is reflectively conscious of needing to brush her teeth, if she goes long-enough without doing it, her prereflective salience network will at some point trigger the default mode related action-controlling ruminative thought, "I should really get on with brushing my teeth"[27]. Although skill learning likely begins with the recitation of ritualized action narratives, embodied practice ensures that in time the individual no longer needs to maintain top-down control. Conversely, over time one may gain sufficient metacognitive experience to listen to the body and determine that one has made some ill mistake, perhaps forgotten to turn the teakettle down. The ability to translate intuitive "gut" feelings into meaningful, reliable decisions requires both a keen introspective practice and a sufficiently well developed self-theory-of-mind. Thus we do not always listen to the cues of our bodies, instead repeating mistakes again and again. This is a kind of delicate balance between our sensory–motor and reflective consciousness, as the two are in a constant reciprocal connection.

If consciousness is related to sensory–motor history, one could ask how they differ from one another in consciousness. How does any one individual differ from, for example a female, an older person, a nun, a juggler, etc.? The answer lies within both the prereflective and reflective consciousness. Prereflectively, these individuals will have different potentials for action; reflectively, they will have subtly different self-metaphors. The juggler may notice his every move, while the nun feels that she is a direct extension of God, with little awareness of her body. A woman may be more or less likely to assert herself depending on the permissiveness of her local culture. Reflectively, the older person will experience herself and her world through the lens of a fundamentally different viewpoint from that of the child, perhaps with the knowledge of long months spent in a war trench, or the exoneration of a successful business venture. It is not the case that we differ merely in memory or the contents of consciousness, as these elements are intricately interwoven into the conscious experience of the individual, preconfiguring our perception of the world.

We contend further that social interaction plays a specific mechanistic role in the development of self-narrative and action-control. At one level, the constant reminders from a mother to her child seek to control his attention and teach new skills. The human prefrontal cortex is extremely sensitive to cues from others. Recent meta-analysis found that "threat to social identity" and "loss of social control" was the greatest elicitors of hypothalamic–pituitary–amygdala cortisol secretion (Dickerson and Kemeny, 2004). Cortisol leads to hyper-activation of the prefrontal cortex and repeated stress-induced cortisol exposure results in thinning of prefrontal density and fronto-amygdalar connectivity. This is, of course, only one among many similar mechanisms and hormones that are sensitive to social interaction and influence development; there are critical periods such as the spurt in theory-of-mind development between ages 3 and 4 and it seems certain that everything from a mother's congratulating smile to video games and school

[26] See Draganski and May (2008) for a review of these and similar findings.

[27] Or perhaps a simple "stare" or gaze from the mother will motivate this behavior.

sports will have an impact on neural development across the brain. Still, we believe there is sufficient evidence of the central importance of intersubjectivity in infant cognitive development. For example, infants are highly responsive to social cues at an early age (Senju and Csibra, 2008), and even new-born twins demonstrate coordinated, coupled social interaction (Castiello et al., 2010). Clearly the brain is equipped from the very beginning to learn about the world from others.

Given the extreme plasticity of the social-cognitive and executive prefrontal networks in the first two decades of life, we contend that interactive and social-cognitive mechanisms play a crucial role in the development of consciousness. It is thus not our claim that social–cortisol response is the only mechanism of plastic adaptation, but rather, one (highly important) mechanism for the kind of sociocultural adaptation under discussion. For the purpose of length, we have not discussed other equally important milestones in the development of (for example) language and motor function, yet these also are likely to depend upon intersubjective plasticity mechanisms to some degree. For this paper, we restrict our review to related mechanisms of systematic anti-correlation between RSNs.

Another piece of evidence for the DMN's susceptibility toward interaction-induced plasticity inducing comes from recent research by Schilbach et al. (2006), demonstrating that a non-cognitive, interactive joint-attention task with virtual avatars actually activates DMN areas while deactivating action-salience systems. Thus what is actually "task-positive" may depend upon the social context within which it occurs; if I am to process an engaging, dynamic, interactive person than my DMN could be quite important for that interaction. On the other hand, if my task is to respond rapidly and accurately to eyeball distractors, the DMN might simply get in the way of this "social" task.

We can now take these analyses and combine them. The DMN exhibits task-free, slow-wave, spontaneous activity that is associated with narrative processing, self-relatedness, reflective consciousness, and ruminative thinking. This association is not strictly conscious; the DMN retains coherence under anesthesia[28] (Peltier et al., 2005; Raichle and Snyder, 2007; Greicius et al., 2008), while connectivity of the DMN remains in locked-in (but not "brain dead") patients (Boly et al., 2009; Vanhaudenhuyse et al., 2010). However, DMN activity does correlate with individual differences and deactivations relate to specific psychopathological traits and personality measures (Sheng et al., 2010).

Furthermore, Lewis et al. (2009) have demonstrated that learning on a visual–motor task, predicts alterations in frontoparietal and visual cortex resting connectivity. We thus propose that the DMN grounds the sometimes-reflective iterative rehearsal of social, self, and action narratives. More globally, the resting-connectome forms a crucial part of the pre-noetic structure[29] of our sensorimotor consciousness, determining what is passed from salience, to reflective rumination, to pre-potent action control. Through specific patterns of neural entrainment, recitation of these ritualized themes brings about alterations in connectivity in these resting networks and alters task-elicited functional specializations (Northoff et al., 2010). In this way the gradual build-up of experience is synthesized in a subjective format and stored for future recollection.

Thus over time the specific sociocultural niche, including socioeconomic factors, access to quality education, parenting style, and even local pollutants (Chen and Schwartz, 2009) contribute to the precise individual balance of these networks. In short, we suggest that sociocultural learning entrains the "what" and "how" of information transfer between the DMN, Salience, and Control networks. Through repetition (reflective and otherwise) certain themes are entrained within the DMN. That is to say, the specifics of ones' cultural context become a constant theme within the overall autobiographical narrative. In this way the reflective consciousness is linked to the particulars of one's culture, and will be shared or different between cultures depending upon the degree of overlap between them. The sensory–motor consciousness (and neural substrate), while highly plastic through early development and capable of recovery in response to injury, ultimately produces an extremely similar outcome regardless of ones' locale.

The greatest area for difference, then, is the particular interaction of the automatic and reflective networks, i.e., those areas where small differences in gesture can have a vast impact on a group, or where careful reflective attention is absolutely necessary for ultimate sensory–motor control. Thus, the information that is available to the salience and CEN networks (e.g., what appears in visual consciousness, and the action-systems repertoire of acceptable pre-potent' responses), will be modulated by whatever social set, context, or role is being primed by the social-cognitive default network. These are in turn set about by the developmental trajectory of the interactive agent.

ARE METACOGNITION AND SOCIAL COGNITION TRAINABLE "SKILLS"?

Here is an obvious truism: some people are better at social cognition than others. Politicians, lawyers, secret service agents, and other trained professionals depend upon highly sharpened belief–desire prediction models. Academics in the humanities must spin long, extremely complicated and obtuse narratives entertaining hundreds if not thousands of years of sociocultural development. And yet, throughout the social-cognitive neurosciences, there exists almost no objective measures of social-cognitive competence. This has made the exploration of plasticity within the medial-prefrontal node of the default mode somewhat more difficult, as there exists no metric by which to evaluate training-related social-cognitive gains. However, in the face of a lack of direct evidence, we can conclude from several sources that these processes are also highly plastic, both intrinsically and in relation to the salience and control networks.

First, developmentally speaking, the medial-prefrontal and temporoparietal regions are among the very last to reach full developmental maturity (Gogtay et al., 2004) with neural development in the human neocortex being marked by massive Hebbian reinforcement (white matter and synaptic connectivity increases)

[28] At first glance this finding may seem contradictory with our hypothesis regarding the involvement of DMN activity in reflective consciousness. That the DMN remains intrinsically coupled during total anesthesia indicates only however that the function of the network is highly automatic. Research in this area indicates the differential dosages of anesthesia induce differential degrees of deactivation and decoupled activity in the DMN. It is our account that DMN *in interaction* with frontoparietal action-salience systems that underlies consciousness.

[29] Literally, "before consciousness," e.g., that which structures but is not present within consciousness (Gallagher and Zahavi, 2008). What is pre-noetic for consciousness? The body, culture, and the neural connectome are all relevant examples.

and neural pruning (gray matter decreases). As neural maturity is achieved via "back to front" development, the neural substrates of theory-of-mind (e.g., MPFC, TPJ, etc) remain open to experiential plasticity and training well into early adulthood (see Blakemore and Choudhury 2006). As we have argued, a primary causal locus for the development of consciousness is the small-group ritualization and enculturation of young brains. Although we are in some sense born with the sensorimotoric equivalent of our early hominid ancestors, we must rapidly entrain ourselves within the highly complex and interwoven social-narrative tapestries that regulate action and prescribe behavior across nested spatiotemporal scales.

Before moving on, we need to briefly mention the notion of temporal receptivity. The given plasticity of a cognitive circuit should at least in part depend upon its window of temporal receptivity. That is to say, it does not make much sense for my social-narrative brain system to be sensitive to the fine tuned sensory–motor dynamics that entrain action-oriented networks. Simply put, blindfolding a participant for an hour of perceptual-motor training should not directly create neuroplastic adaptation in the medial-prefrontal network. Rather, these adaptations are likely to be localized to the high-speed window of the visual–motor system. Indeed, recent research suggests a differential topography of temporal receptive windows throughout the human neocortex: areas associated with reflective cognition tend to have much slower windows (3–36 s and upward) whereas visual cortex responds to information at higher frequency (<1 s; Hasson et al., 2008). We suggest that the greater temporal period of the temporoparietal junction and MPFC correspond more directly to the frequency of slow-wave behavioral–narrative interaction.

Take, for example, a conference dinner. We can here identify multiple temporal scales for relevant information processing. The fast-wave phenomenon like group shifts, behavioral chameleon phenomenon, and embodied mirroring occur far too fast to be tracked exhaustively by the reflective representational system. Rather, it is only the slowly aggregated summation of these events that enter into my narratological, metarepresentational processing, which is itself constrained by the acoustic temporal dynamics and rhythms of spoken interaction. Spiraling this concept outward, we can deduce that narrative processing also includes many slow-wave phenomena, entailing high-speed tasks such as the online identification of misplaced utterances, but also incorporating the gradual updating of self and social narratives. It is not the case that I rapidly and constantly update my self-narrative; rather it is only the aggregate sum of significant interactive events that eventually enters into my narrative. Thus it is the sum of my social interactions that engineer particular types of consciousness through the gross plasticity of our neural system. Our hypothesis is that this gradual integration of experience, led by social engagement with the world, is entrained by the default mode and in turns structures the phenomenal salience of both prereflective and reflective perception.

CONSCIOUSNESS, ANTI-CORRELATION, AND THE TOPOGRAPHY OF MIND

Although we have argued for a bifurcated mental taxonomy between the prereflective and the reflective, we contend that a full understanding of the human mind must move beyond simplistic dualisms of any form, instead embracing the view that the constructs we denote as mental are in shifting, interrelated positions. Thus, although we discuss anti-correlations and their implications for the reflective/prereflective dimension, it is important to note that the "whole story" is likely to be far less clean than this distinction implies.

We are certainly not alone in surmising that the anti-correlation finding minimally suggests a unique form of informational interaction between the action-oriented salience/control and social-cognitive domains (see Carhart-Harris and Friston, 2010). However, basic phenomenology here reveals a few caveats. It is not the case that in prereflective interaction I am no longer able to engage in detached, metarepresentational processing. Nor is it fully the case that in my detached navel-gazing I am shielded from the sensory–motor fluctuation of my body in its environment. Rather, in both cases there exists a fine tuned spatiotemporal distribution of processing and resource allocation between these functional domains.

To illustrate, consider the classical example of "cocktail party coping." To clarify and keep this example quite familiar, let us say the party in question is a post-conference dinner. I, the subject in question, having just entered the room, am immediately presented with the multitude of faces, voices, and explorative eyeball saccades that fixate in wild fluctuation across the room. Further constraining my interaction are the intersubjective power narratives that hang like spectral ether across the room; the "who's whos" and veritas of any social gathering. Entering into conversation, I must not only attend to the complex linguistic content of my new dyad, but also the randomly wavering eye-gazes, body postures, and other embodied semantic content determining the mood of the room. Should I continue speaking, or perhaps take my place at the dinner table? Has my conversational partner become bored, or should I continue our discussion? As we interact, I must continuously update the narrative coming from my mouth and my memory with the information given back to me by my partner. This process will be continually structured by salient target information as well as cultural representational values and is likely to only be minimally "conscious" in the traditional (i.e., intentional, reflective, self-identical) sense of the term.

Consider further the relationship between power narrative and embodied dynamics that unfold in this particular scene. Surely I am not constantly meta-conscious of the continuously unfolding social dynamics. To be so would be almost schizophrenic, and certainly I might suffer social-anxiety should I try to iteratively track all these possible variables. Rather, in line with the reputation costs associated with embodied social behavior, I simply act. I respond automatically to belief states, embodied gestures, and a host of constantly unfolding social-cognitive dynamics. My eyes and face must automatically track my partner's, lest I fail in engaging the chameleon effects that seem so crucial for smooth interpersonal interaction. Clearly we have a situation where my narrative processing is automatically guiding my tracking of salient social cues, and also in the inhibition of action: the social-narrative stream pouring forth from my mouth is consistently inhibited. Simply put, the power-dynamics of my social context are modulating both my behavior (compare ones' posture in a work setting to that of a bar or amongst close friends) and my default speech. If I am to be socially successful, I will inhibit whatever dirty jokes I might tell otherwise. Human social interaction is completely pervaded by these information intensive interchanges

of narrative and embodied coping (Gallagher and Hutto, 2008). Pure anti-correlation cannot obtain. Rather, my executive-control systems and visual salience networks must be in constant reciprocal communication with my social-representational-moral DMN.

CONCLUSION

As is often the case in the cognitive sciences, we have tried to demonstrate that there are multiple areas of critical overlap between a variety of currently disparate research trends. We have reviewed new research findings and theoretical developments in neuroplasticity, cognitive neuroscience, development, joint action and social cognition, phenomenology, and the philosophical investigation of consciousness. As a definitive integration of these areas is obviously beyond the scope of the present paper, we present these findings in hopes of convincing the reader that there is a crucial role for social interaction and plasticity in the ontogeny of human consciousness.

Our account motivates several conclusions. First, an explanation of human consciousness is not primarily a matter of explaining pure phenomenal feels. This is not to deny the importance of phenomenal feels for grounding our experience as living, embodied organisms. Rather, we have argued that to fully explain consciousness is to also give an account of the unique capacity for metacognition and metarepresentation demonstrated through what has been called *offline intelligence*. We have argued that a mature science of consciousness must be careful in avoiding the conflation of online intelligence with offline intelligence. Second, a mature explanation of the human mind must take into account the radical multisensory plasticity underlying both online and offline intelligence.

Further than just integration, we have hoped to advance an account that places the DMN, and the larger paradigm of human connectomics, as a central mechanism for the development of consciousness. We have thus proposed the following:

A. Human consciousness depends upon the ruminative, self-specific narrative stream of reflective thought.
B. This narrative stream is central for the production and control of action, and isspecifically encultured in the DMN by social interaction.
C. Consciousness appears to emerge from the complex interplay of reflective-narrative process and the sensorimotor control/salience dynamic.
D. It is the vast connective plasticity of the human frontoparietal cortex that enables the rapid development and acquisition of 10,000 plus years of evolved tool use, upon which our extended consciousness depends.

There are of course some worries generated by our account. One initial worry might be that we have "over intellectualized" consciousness and lost sight of the original explanandum of "what-it-is-like" to be an organism. However, although philosophical accounts of consciousness as pure phenomenal subjectivity might be initially appealing, we have argued that they are insufficient to account for the rich narratological capacities of contemporary human consciousness and fail to capture what makes us uniquely human. Accordingly, we avoid the theoretical pitfall of denying phenomenal states to non-human organisms while still retaining a robust sense of consciousness that goes beyond embodied sensorimotor experience. Such an account is desirable precisely because it can do justice to the subjectivity of animals while acknowledging the unique cultural–linguistic conditioning underlying the cognitive prowess of adult humans.

ACKNOWLEDGMENTS

The authors wish to thank Shaun Gallagher, Andreas Roepstorff, Chris Frith, Leonhard Schilbach, Jon Cogburn, and John Protevi for comments on early versions of this paper. We also thank the LSU Philosophy Department, the Danish Center for Functionally Integrative Neuroscience, and the Interacting Minds Project for supporting this research.

REFERENCES

Amodio, D., and Frith, C. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nat. Rev. Neurosci.* 7, 268–277.

Armstrong, D. (1997). "What is consciousness?" in *The Nature of Consciousness: Philosophical Debates*, eds N. Block, O. Flanagan, and G. Güzeldere (Cambridge, MA: MIT Press), 721–728. [Original edition, 1978].

Baars, B. J. (1997). *In the Theater of Consciousness*. New York: Oxford University Press.

Bavelier, D., and Neville, H. (2002). Cross-modal plasticity: where and how? *Nat. Rev. Neurosci.* 3, 443–452.

Biswal, B. B., Mennes, M., Zuo, X.-N., Gohel, S., Kelly, C., Smith, S. M., Beckmann, C. F., Adelstein, J. S., Buckner, R. L., Colcombe, S., Dogonowski, A.-M., Ernst, M., Fair, D., Hampson, M., Hoptman, M. J., Hyde, J. S., Kiviniemi, V. J., Kötter, R., Li, S.-J., Lin, C.-P., Lowe, M. J., Mackay, C., Madden, D. J., Madsen, K. H., Margulies, D. S., Mayberg, H. S., McMahon, K., Monk, C. S., Mostofsky, S. H., Nagel, B. J., Pekar, J. J., Peltier, S. J., Petersen, S. E., Riedl, V., Rombouts, Serge, A. R. B., Rypma, B., Schlaggar, B. L., Schmidt, S., Seidler, R. D., Siegle, G. J., Sorghh, C., Teng, G.-J., Veijola, J., Villringer, A., Walter, M., Wang, L., Weng, X.-C., Whitfield-Gabrieli, S., Williamson, P., Windischberger, C., Zang, Y.-F., Zhang, H.-Y., Castellanos, F. X., and Milham, M. P. (2010). Toward discovery science of human brain function. *Proc. Natl. Acad. Sci. U.S.A.* 107, 4734.

Blakemore, S. (2008). The social brain in adolescence. *Nat. Rev. Neurosci.* 9, 267–277.

Blakemore, S., and Choudhury, S. (2006). Development of the adolescent brain: implications for executive function and social cognition. *J. Child Psychol. Psychiatry* 47, 296–312.

Boly, M., Tshibanda, L., Vanhaudenhuyse, A., Noirhomme, Q., Schnakers, C., Ledoux, D., Boveroux, P., Garweg, C., Lambermont, B., Phillips, C., Luxen, A., Moonen, G., Bassetti, C., Maquet, P., and Laureys, S. (2009). Functional connectivity in the default network during resting state is preserved in a vegetative but not in a brain dead patient. *Hum. Brain Mapp.* 30, 2393–2400.

Brooks, R. (1991). Intelligence without representation. *Artif. Intell.* 47, 139–159.

Buckner, R. L. (2010). Human functional connectivity: new tools, unresolved questions. *Proc. Natl. Acad. Sci. U.S.A.* 107, 10769–10770.

Buckner, R. L., and Vincent, J. L. (2007). Unrest at rest: default activity and spontaneous network correlations. *Neuroimage* 37, 1091–1096.

Carhart-Harris, R., and Friston, K. (2010). The default-mode, ego-functions and free-energy: a neurobiological account of Freudian ideas. *Brain* 133, 1265–1283.

Castiello, U., Becchio, C., Zoia, S., Nelini, C., Sartori, L., Blason, L., D'Ottavio, G., Bulgheroni, M., and Gallese, V. (2010). Wired to be social: the ontogeny of human interaction. *PLoS ONE* 5, e13199. doi: 10.1371/journal.pone.0013199

Chen, J.-C., and Schwartz, J. (2009). Neurobehavioral effects of ambient air pollution on cognitive performance in US adults. *Neurotoxicology* 30, 231–239.

Christoff, K., Gordon, A. M., Smallwood, J., Smith, R., and Schooler, J. W. (2009). Experience sampling during fMRI reveals default network and executive system contributions to mind wandering. *Proc. Natl. Acad. Sci. U.S.A.* 106, 8719.

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

Clark, A. (2003). *Natural-Born Cyborgs: Minds, Technologies, and the Future of Human Intelligence.* New York: Oxford University Press.

Clark, A. (2008). *Supersizing the Mind: Embodiment, Action, and Cognitive*

Extension. New York: Oxford University Press.

Clark, A., and Toribio, J. (1994). Doing without representing? *Synthese* 101, 401–431.

Damoiseaux, J., and Greicius, M. (2009). Greater than the sum of its parts: a review of studies combining structural connectivity and resting-state functional connectivity. *Brain Struct. Funct.* 213, 525–533.

Dehaene, S., Pegado, F., Braga, L. W., Ventura, P., Filho, G. N., Jobert, A., Dehaene-Lambertz, G., Kolinsky, R., Morais, J., and Cohen, L. (2010). How learning to read changes the cortical networks for vision and language. *Science* 330, 1359–1364.

Dennett, D. (1993). *Consciousness Explained*. London: Penguin.

Dewey, J. (1958). *Experience and Nature*. New York: Dover Publications. [Original edition, 1925].

Dickerson, S. S., and Kemeny, M. E. (2004). Acute stressors and cortisol responses: a theoretical integration and synthesis of laboratory research. *Psychol. Bull.* 130, 355–391.

Donald, M. (2001). *A Mind So Rare: The Evolution of Human Consciousness*. New York: WW Norton & Company.

Draganski, B., Gaser, C., Kempermann, G., Kuhn, H. G., Winkler, J., Büche, C., and May, A. (2006). Temporal and spatial dynamics of brain structure changes during extensive learning. *J. Neurosci.* 26, 6314.

Draganski, B., and May, A. (2008). Training-induced structural changes in the adult human brain. *Behav. Brain Res.* 192, 137–142.

Dretske, F. (1993). Conscious experience. *Mind* 102, 263–283.

Dreyfus, H. L. (2002). Intelligence without representation – Merleau-Ponty's critique of mental representation the relevance of phenomenology to scientific explanation. *Phenomenol. Cogn. Sci.* 1, 367–383.

Esposito, F., Bertolino, A., Scarabino, T., Latorre, V., Blasi, G., Popolizio, T., Tedeschi, G., Cirillo, S., Goebel, R., and Di Salle, F. (2006). Independent component model of the default-mode brain function: assessing the impact of active thinking. *Brain Res. Bull.* 70, 263–269.

Fair, D A., Cohen, A. L., Dosenbach, N. U. F., Church, J. A., Miezin, F. M., Barch, D. M., Raichle, M. E., Petersen, S. E., and Schlaggar, B. L. (2008). The maturing architecture of the brain's default network. *Proc. Natl. Acad. Sci. U.S.A.* 105, 4028–4032.

Farb, N. A. S., Segal, Z. V. Mayberg, H., Bean, J., McKeon, D., Fatima, Z., and Anderson, A. K. (2007). Attending to the present: mindfulness meditation reveals distinct neural modes of self-reference. *Soc. Cogn. Affect. Neurosci.* 2, 313–322.

Fox, M., Zhang, D., and Snyder, A. (2009). The global signal and observed anticorrelated resting state brain networks. *J. Neurophysiol.* 101, 3270–3283.

Fransson, P., Skiöld, B., Horsch, S., Nordell, A., Blennow, M., Lagercrantz, H., and Åden, U. (2007). Resting-state networks in the infant brain. *Proc. Natl. Acad. Sci. U.S.A.* 104, 15531–15536.

Freeman, W. J. (2005). A field-theoretic approach to understanding scale-free neocortical dynamics. *Biol. Cybern.* 92, 350–359.

Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J.-P., Frith, C. D., and Frackowiak, R. S. J. (1994). Statistical parametric maps in functional imaging: a general linear approach. *Hum. Brain Mapp.* 2, 189–210.

Frith, C. (2005). The self in action: lessons from delusions of control. *Conscious. Cogn.* 14, 752–770.

Frith, C. (2007). The social brain? *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 362, 671.

Frith, C., and Frith, U. (2006). How we predict what other people are going to do. *Brain Res.* 1079, 36–46.

Frith, C., and Frith, U. (2007). Social cognition in humans. *Curr. Biol.* 17, R724–R732.

Gallagher, S. (2005). *How the Body Shapes the Mind*. New York: Oxford University Press.

Gallagher, S., and Hutto, D. (2008). "Understanding others through primary interaction and narrative practice," in *The Shared Mind: Perspectives on Intersubjectivity*, ed. J. Zlatev (Amsterdam: John Benjamins Publishing Co.), 17–38.

Gallagher, S., and Zahavi, D. (2008). *The Phenomenological Mind: An Introduction to Philosophy of Mind and Cognitive Science*. Abingdon, Oxon: Routledge Press.

Gennaro, R. J. (ed.). (2004). *Higher-Order Theories of Consciousness: An Anthology*. Philadelphia: John Benjamins.

Gogtay, N., Giedd, J. N., Lusk, L., Hayashi, K. M., Greenstein, D., Vaituzis, A. C., Nugent, T. F. III, Herman, D. H., Clasen, L. S., Toga, A. W., Rapoport, J. L., and Thompson, P. M. (2004). Dynamic mapping of human cortical development during childhood through early adulthood. *Proc. Natl. Acad. Sci. U.S.A.* 101, 8174.

Greicius, M. (2008). Resting-state functional connectivity in neuropsychiatric disorders. *Curr. Opin. Neurol.* 21, 424–430.

Greicius, M. D., Srivastava, G., Reiss, A. L., and Menon, V. (2004). Default-mode network activity distinguishes Alzheimer's disease from healthy aging: evidence from functional MRI. *Proc. Natl. Acad. Sci. U.S.A.* 101, 4637–4642.

Greicius, M. D., Kiviniemi, V., Tervonen, O., Vainionpää, V., Alahuhta, S., Reiss, A. L., and Menon, V. (2008). Persistent default mode network connectivity during light sedation. *Hum. Brain Mapp.* 29, 839–847.

Greicius, M. D., and Menon, V. (2004). Default-mode activity during a passive sensory task: uncoupled from deactivation but impacting activation. *J. Cogn. Neurosci.* 16, 1484–1492.

Haier, R. J., Karama, S., Leyba, L., and Jung, R. E. (2009). MRI assessment of cortical thickness and functional activity changes in adolescent girls following three months of practice on a visual-spatial task. *BMC Res. Notes* 2, 174. doi: 10.1186/1756-0500-2-174

Hasson, U., Yang, E., Vallines, I., Heeger, D. J., and Rubin, N. (2008). A hierarchy of temporal receptive windows in human cortex. *J. Neurosci.* 28, 2539–2550.

Hobson, J. A. (2009). REM sleep and dreaming: towards a theory of protoconsciousness. *Nat. Rev. Neurosci.* 10, 803–813.

Holt-Lunstad, J., Smith, T. B., and Layton, J. B. (2010). Social relationships and mortality risk: a meta-analytic review. *PLoS Med.* 7, e1000316. doi: 10.1371/journal.pmed.1000316

Hutto, D. D. (2008). *Folk Psychological Narratives*. Cambridge, MA: MIT Press.

Isamah, N., Faison, W., Payne, M. E., MacFall, J., Steffens, D. C., Beyer, J. L., Ranga Krishnan, K., and Taylor, W. D. (2010). Variability in frontotemporal brain structure: the importance of recruitment of African Americans in neuroscience research. *PLoS ONE* 5, e13642. doi: 10.1371/journal.pone.001364

James, W. (1950). *Principles of Psychology*, Vol. I. Mineola: Dover. [Original edition, 1890].

Jaynes, J. (2000). *The Origin of Consciousness in the Breakdown of the Bicameral Mind*. Boston: Houghton Mifflin. [Original edition, 1976].

Jenkins, J., and Astington, J. (1996). Cognitive factors and family structure associated with theory of mind development in young children. *Dev. Psychol.* 32, 70–78.

Jerbi, K., Vidal, J. R., Ossandon, T., Dalal, S. S., Jung, J., Hoffmann, D., Minotti, L., Bertrand, O., Kahane, P., and Lachaux, J.-P. (2010). Exploring the electrophysiological correlates of the default-mode network with intracerebral EEG. *Front. Syst. Neurosci.* 4:27. doi: 10.3389/fnsys.2010.00027

Koch, C. (2004). *The Quest for Consciousness: A Neurobiological Approach*. Englewood, CO: Roberts & Company Publishers.

Lakoff, G., and Johnson, M. (1980). *Metaphors We Live By*. Chicago: University of Chicago Press.

Lakoff, G., and Johnson, M. (1999). *Philosophy in the Flesh: the Embodied Mind and its Challenge to Western Thought*. New York: Basic Books.

Lewis, C. M., Baldassarre, A., Committeri, G., Romani, G. L., and Corbetta, M. (2009). Learning sculpts the spontaneous activity of the resting human brain. *Proc. Natl. Acad. Sci. U.S.A.* 106, 17558–17563.

Lycan, W. G. (1997). "Consciousness as internal monitoring," in *The Nature of Consciousness: Philosophical Debates*, eds N. Block, O. Flanagan, and G. Güzeldere (Cambridge, MA: MIT Press), 755–771.

Mano, Y., Harada, T., Sugiura, M., Saito, D. N., and Sadato, N. (2009). Perspective-taking as part of narrative comprehension: a functional MRI study. *Neuropsychologia* 47, 813–824.

Mantini, D., Perrucci, M., and Del Gratta, C. (2007). Electrophysiological signatures of resting state networks in the human brain. *Proc. Natl. Acad. Sci. U.S.A.* 104, 13170–13175.

Mar, R. (2004). The neuropsychology of narrative: story comprehension, story production and their interrelation. *Neuropsychologia* 42, 1414–1434.

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

May, A., Hajak, G., Gänbauer, S., Steffens, T., Langguth, B., Kleinjung, T., and Eichhammer, P. (2007). Structural brain alterations following 5 days of intervention: dynamic aspects of neuroplasticity. *Cereb. Cortex* 17, 205–210.

Merleau-Ponty, M. (2006). *Phenomenology of Perception*. New York: Routledge and Kegan Paul. [Original edition, 1945].

Morcom, A., and Fletcher, P. (2007). Does the brain have a baseline? Why we should be resisting a rest. *Neuroimage* 37, 1073–1082.

Morin, A. (2005). Possible links between self-awareness and inner speech theoretical background, underlying mechanisms, and empirical evidence. *J. Conscious. Stud.* 12, 115–134.

Noë, A. (2004). *Action in Perception*. Cambridge, MA: MIT Press.

Northoff, G., Qin, P., and Nakao, T. (2010). Rest-stimulus interaction in the brain: a review. *Trends Neurosci.* 33, 277–284.

O'Regan, J. K. (1992). Solving the "real" mysteries of visual perception: the world as an outside memory. *Can. J. Psychol.* 46, 461–488.

Peltier, S. J., Kerssens, C., Hamann, S. B., Sebel, P. S., Byas-Smith, M., and Hu, X. (2005). Functional connectivity

changes with concentration of sevoflurane anesthesia. *Neuroreport* 16, 285.
Protevi, J. (2009). *Political Affect: Connecting the Social and the Somatic*. Minneapolis: University of Minnesota Press.
Raichle, M., MacLeod, A., and Snyder, A. (2001). A default mode of brain function. *Proc. Natl. Acad. Sci. U.S.A.* 98, 676–682.
Raichle, M., and Snyder, A. (2007). A default mode of brain function: a brief history of an evolving idea. *Neuroimage* 37, 1083–1090.
Reed, E. (1996). *Encountering the World: Toward and Ecological Psychology*. New York: Oxford University Press.
Roepstorff, A., and Frith, C. (2004). What ís at the top in the top-down control of action? Script-sharing and top-top control of action in cognitive experiments. *Psychol. Res.* 68, 189–198.
Rubin, B. (2009). Changing brains: the emergence of the field of adult neurogenesis. *Biosocieties* 4, 407–424.
Schilbach, L., Wohlschlaeger, A. M., Kraemer, N. C., Newen, A., Shah, N. J., Fink, G. R., and Vogeley, K. (2006). Being with virtual others: neural correlates of social interaction. *Neuropsychologia* 44, 718–730.
Schilbach, L., Eickhoff, S. B., Rotarska-Jagiela, A., Fink, G. R., and Vogeley, K. (2008). Minds at rest. Social cognition as the default mode of cognizing and its putative relationship to the "default system" of the brain. *Conscious. Cogn.* 17, 457–467.
Schoenemann, P., Sheehan, M., and Glotzer, L. (2005). Prefrontal white matter volume is disproportionately larger in humans than in other primates. *Nat. Neurosci.* 8, 242–252.
Schooler, J. W. (2002). Re-representing consciousness: dissociations between experience and meta-consciousness. *Trends Cogn. Sci.* 6, 339–344.
Senju, A., and Csibra, G. (2008). Gaze following in human infants depends on communicative signals. *Curr. Biol.* 18, 668–671.
Sheng, T., Gheytanchi, A., and Aziz-Zadeh, L. (2010). Default network deactivations are correlated with psychopathic personality traits. *PLoS ONE* 5, e12611. doi: 10.1371/journal.pone.0012611
Shimojo, S., and Nakajima, Y. (1981). Adaptation to the reversal of binocular depth cues: effects of wearing left-right reversing spectacles on stereoscopic depth perception. *Perception* 10, 391–402.
Smallwood, J., and Schooler, J. (2006). The restless mind. *Psychol. Bull.* 132, 946.
Spreng, R., Mar, R., and Kim, A. (2009). The common neural basis of autobiographical memory, prospection, navigation, theory of mind, and the default mode: a quantitative meta-analysis. *J. Cogn. Neurosci.* 21, 489–510.
Sridharan, D., and Levitin, D. (2008). A critical role for the right fronto-insular cortex in switching between central-executive and default-mode networks. *Proc. Natl. Acad. Sci. U.S.A.* 105, 12569–12574.
Steinberg, D. (2006). *Consciousness Reconnected: Missing Links Between Self, Neuroscience, Psychology, and the Arts*. Abingdon: Ratcliffe Press.
Stern, D. (2009). Pre-reflexive experience and its passage to reflexive experience: a developmental view. *J. Conscious. Stud.* 16, 307–331.
Stettler, D. D., Yamahachi, H., Li, W., Denk, W., and Gilbert, C. D. (2006). Axons and synaptic boutons are highly dynamic in adult visual cortex. *Neuron* 49, 877–887.
Stevens, W. D., Buckner, R. L., and Schacter, D. L. (2010). Correlated low-frequency BOLD fluctuations in the resting human brain are modulated by recent experience in category-preferential visual regions. *Cereb. Cortex* 8, 1997–2006.
Tang, Y.-Y., Lu, Q., Geng, X., Stein, E. A., Yang, Y., and Posner, M. I. (2010). Short-term meditation induces white matter changes in the anterior cingulate. *Proc. Natl. Acad. Sci. U.S.A.* 107, 15649.
Thompson, E. (2007). *Mind in Life: Biology, Phenomenology, and the Sciences of Mind*. Cambrige, MA: Harvard University Press.
Tylen, K., Weed, E., Wallentin, M., Roepstorff, A., and Frith, C. D. (2010). Language as a tool for interacting minds. *Mind Lang.* 25, 3–29.
Van Dijk, K. R., Hedden, T., Venkataraman, A., Evans, K. C., Lazar, S. W., and Buckner, R. L. (2010). Intrinsic functional connectivity as a tool for human connectomics: theory, properties, and optimization. *J. Neurophysiol.* 103, 297–321.
Vanhaudenhuyse, A., Noirhomme, Q., Tshibanda, L. J., Bruno, M. A., Boveroux, P., Schnakers, C., Soddu, A., Perlbarg, V., Ledoux, D., Brichant, J. F., Moonen, G., Maquet, P., Greicius, M. D., Laureys, S., and Boly, M. (2010). Default network connectivity reflects the level of consciousness in non-communicative brain-damaged patients. *Brain* 133, 161.
Varela, F., Thompson, E., and Rosch, E. (1991). *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge, MA: MIT Press.
Wegner, D. M. (2002). *The Illusion of Conscious Will*. Cambridge, MA: MIT Press.
Wheeler, M. (2007). *Reconstructing the Cognitive World: The Next Step*. Cambridge, MA: MIT Press.

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

Received: 11 November 2010; accepted: 31 January 2011; published online: 28 February 2011.
Citation: Allen M and Williams G (2011) Consciousness, plasticity, and connectomics: the role of intersubjectivity in human cognition. Front. Psychology ***2****:20. doi: 10.3389/fpsyg.2011.00020*
This article was submitted to Frontiers in Consciousness Research, a specialty of Frontiers in Psychology.

frontiers in PSYCHOLOGY

HYPOTHESIS AND THEORY ARTICLE
published: 26 October 2010
doi: 10.3389/fpsyg.2010.00168

Framework of consciousness from semblance of activity at functionally LINKed postsynaptic membranes

Kunjumon I. Vadakkan*

Division of Neurology, Department of Internal Medicine, University of Manitoba, Winnipeg, MB, Canada

Edited by:
Morten Overgaard, Aarhus University, Denmark

Reviewed by:
Jessica Saenger, University of Düsseldorf, Germany
Giorgio Marchetti, Nokia Siemens Networks – IT, Italy

****Correspondence:***
Kunjumon I. Vadakkan, Division of Neurology, Department of Internal Medicine, University of Manitoba, GF532-820 Sherbrook Street, Winnipeg, MB, Canada R3A1R9.
e-mail: umvadakk@cc.umanitoba.ca

Consciousness is seen as a difficult "binding" problem. Binding, a process where different sensations evoked by an item are associated in the nervous system, can be viewed as a process similar to associative learning. Several reports that consciousness is associated with some form of memory imply that different forms of memories have a common feature contributing to consciousness. Based on a proposed synaptic mechanism capable of explaining different forms of memory, we developed a framework for consciousness. It is based on the formation of semblance of sensory stimulus from (1) synaptic semblances when excitatory postsynaptic potentials arrive at functionally LINKed postsynaptic membranes, and (2) network semblances when these potentials summate to elicit action potential initiating activity in a network of neurons. It is then possible to derive a framework for consciousness as a multi-dimensional semblance. According to this framework, a continuum of semblances formed from background sensory stimuli and oscillating neuronal activities serve to maintain consciousness. Feasibility of this framework to explain various physiological and pathological states of consciousness, its subjective nature and qualia is examined.

Keywords: consciousness, qualia, binding, memory, semblance hypothesis, functional LINKs

INTRODUCTION

Consciousness is described as a biological and therefore, physical feature of certain higher-level nervous systems of humans and a large number of different types of animals (Searle, 1995). Consciousness is viewed as a hard problem (Chalmers, 1995). Biochemical, electrophysiological, and systems-level studies of the neuronal activities have not allowed us to cross the chasm of consciousness (Blackmore, 2002). However, there is general agreement that brain processes are responsible for consciousness (McGinn, 1989; Searle, 1997) and that finding a perfect match between the contents of the neural and mental states will lead to the discovery of what that process is (McGinn, 1989; Chalmers, 1998). Many investigators have suggested that it is time to approach the problem aggressively with new theories (Gray, 1992) and scientific methods (Farah and Murphy, 2009; Koch, 2009).

Many examinations of the relationship between consciousness and neuronal activities have been made in the past, particularly by examining the continuity in neuronal activity as a possible candidate that can produce a conscious state without any breaks. Oscillatory neuronal networks and their significance in regards to the conscious representation of visual sensory inputs have been studied extensively (Lamme and Spekreijse, 2000; Babiloni et al., 2004; Tallon-Baudry, 2009; Wyart and Tallon-Baudry, 2009). Different frameworks for consciousness based on neuronal activity were made and are being examined (Tononi and Edelman, 1998; Crick and Koch, 2003; Tononi, 2004; Orpwood, 2007; Rolls, 2007; Seth, 2008; He and Raichle, 2009). Even though it is thought that activity patterns and causal interactions between specific groups of neurons could be responsible for consciousness, what constitutes consciousness is not yet known.

While deriving the present framework, theoretical explanations for the following were taken as essential criteria for consciousness. (1) A feasible explanation for qualia is regarded as the gold standard of explanations for consciousness. This is from the view that once a framework of consciousness is developed, then the problem of qualia can be understood better (Crick and Koch, 1998). (2) The discovery of a mechanism that can explain subjective qualities (first-person perspective) of conscious experience is an important parameter of success for deriving a framework of consciousness (Searle, 1992, 1997). (3) A feasible mechanism to accommodate the Zombie phenomenon (Koch and Crick, 2001), where meaningful motor activities can be performed without conscious awareness of them. (4) A mechanism explaining consciousness as a graded, not all-or-none, phenomenon, as evidenced by the varying degrees of loss of consciousness in many clinical settings of acute brain stem injury (Alkire et al., 2008).

A broader degree of philosophical reasoning for the present framework of consciousness was made by integrating information from different fields of brain sciences. The problem was approached by broadening the definition for consciousness by incorporating (1) conscious state, often referred as not required for implicit memory but needed for episodic memory, and (2) consciousness (that is lost during comatose state), which is required for implicit memory. An inductive reasoning approach was used to derive the relationship between consciousness and other brain functions that have a proven direct association with some of the plastic changes at the synapses. Following this, the present work examines the framework's suitability to explain various physiological and pathological conditions associated with conscious states. An abstract of this work was presented at the "Toward a Science of Consciousness" conference, at Tucson, Arizona (Vadakkan, 2010).

THE CONTEXT

Studies from various fields of brain sciences have clearly shown that some form of memory is directly associated with consciousness. Short-term memory is thought to be an essential component necessary for consciousness (Ramachandran and Hirstein, 1997). Higher-order consciousness is viewed as dependent on memory for symbols and meanings in the language circuits (Edelman, 1992). With regard to visual consciousness, some form of very short-term transient memory lasting for only a fraction of a second is thought to be essential for consciousness (Crick and Koch, 1998). Using experimental results from a very small number of cases, it has been suggested that semantic memory and general knowledge may contribute to the conscious features of the subjects (Rosenbaum et al., 2007). Conscious machines are thought to require the activation of some memory representations of a device with a certain kind of structure (Minsky, 1991). Since some form of memory is considered to be associated with consciousness, it is reasonable to examine whether a framework for consciousness can be derived from the basic units of memory.

There are two challenging issues that have prevented us from proceeding further. First, the basic mechanism for memory storage is still not known. Second, different forms of memories are thought to have different mechanisms of operation. A feasible approach that can be used at this juncture will be to (1) find a proposal of memory storage that has a common basic unit capable of explaining different types of memories, and (2) find the relationship of the basic unit with the synaptic functions and use the information to develop a framework for consciousness. The framework can then be tested theoretically by examining the relationship between consciousness and memory. Using this method, a framework of consciousness was made from the basic units of the semblance hypothesis of memory (Vadakkan, 2008). Only general principles are explained, leaving aside the specific roles of different sub-regions of the brain such as the reticular activating system, frontal cortex, thalamus, and basal ganglia.

DERIVATION OF THE FRAMEWORK

The derivation is carried out in two steps. Step one is the derivation of the semblance formation. This is followed by the derivation of the framework of consciousness.

STEP 1: BINDING AND SEMBLANCE FORMATION

The nervous system is endowed with the property of making associations among different sensory inputs. This can be argued to occur due to the facts that (1) most of the items ("item" means anything that relates to the real world) can induce more than one sensation in the nervous system of an animal having multiple sensory systems, and (2) most animals have nervous systems that can perceive more than one sense. Therefore, while perceiving an item, an animal's nervous system simultaneously receives different sensations from the item and makes a meaningful association between those different sensory inputs (This is similar to learning). The combinatorial activation pattern of different sensory inputs from an item could possibly induce specific signatures in the nervous system for identifying the item at a later stage, when more than one item with similar sensory inputs are present. This has been described as an explicit association, namely "binding" (Rosenblatt, 1961; von der Malsburg, 1981). Out of all the binding problems, the problem of consciousness is considered as the most mystifying (Roskies, 1999).

Most of the natural events in the environment cause changes such that they evoke simultaneous novel inputs to more than one sensory system. For example, the ripening of a banana changes its color from green to yellow. The ripening also leads to changes in the consistency of the banana and thus in the quality of touch. There is a change in taste and the banana becomes edible. The yellow-soft-taste sensory triad creates binding such that either one of these sensations will be sufficient to provide the semblance of the other two sensations. The ability of the yellow color of the banana to provide us with the semblance that it is edible is the consequence of binding. To understand binding, the representations of different sensations from an item and their interaction within the nervous system during binding need to be explored.

The simultaneous arrival of different sensory inputs from one single item can create changes in the nervous system such that at a later point of time, the arrival of one of the sensory inputs will evoke the semblance of arrival of the matching sensations from the same item. This can be viewed as a property of nervous systems with multiple sensory systems. Using reductionistic approaches, we sought to discover synaptic plasticity changes that could evoke appropriate changes during the initial exposure to the item such that at a later occasion, one of the sensory inputs from the item could evoke sensory semblance for the second item. In other words, we examined the elements that are required at the synapses to synthesize a virtual sensation of the remaining sensations from an item when one sensation is activated.

Normally, the arrival of activity from a stimulus at the presynapse (presynaptic terminal) evokes the activation of its postsynapse (postsynaptic terminal or dendritic spine). Activity reaches the presynapse in the form of an action potential and induces an excitatory postsynaptic potential (EPSP) at the postsynapse after neurotransmission. It must be argued that for a sensory semblance to occur, the synapse that is normally activated by one sensation needs to acquire the ability to get activated by the second sensation after the nervous system is exposed simultaneously to both these sensations coming from a single item.

To narrow down the requirements within the synaptic structure, using a reductionistic approach, the following question was asked: "What is the minimum statement that is required, in order to say that one synapse is activated?" Activation of a postsynapse takes place when neurotransmission occurs from its presynaptic terminal. Therefore, the activation of the postsynapse can be taken as the minimum requirement to state that one synapse is activated. Reframing some of the earlier statements, it can be said that co-activation of the postsynapses from the two different sensory inputs from an item induces specific changes that will later allow one of the stimuli alone to evoke the activation of the postsynapse that was originally activated by the second sensory item. In order to achieve this, the postsynapses of the synapses that are activated by the two different sensory inputs from the item should be located in close proximity. Since different sensory inputs converge at many locations within the brain, possible changes that can take place at these locations were examined in hopes of substantiating various properties of consciousness.

Experimental studies in neuroscience have not yet identified any structural changes between the postsynapses of neighboring synapses at locations where sensory inputs converge. Based on the derivation that was made in the above paragraph, we hypothesize the formation of a functional LINK (capital letters are used in the word LINK to denote its significance) between the postsynapses (**Figure 1**) at locations where synapses from different sensory inputs converge. Features of functional LINKs are discussed in a separate section.

At a later stage, when one sensory stimulus reaches its synapses and postsynaptic membranes, the functional LINKs can be reactivated such that EPSP can spread from these postsynapses to the neighboring functionally LINKed postsynapses that belong to the second sensory stimulus, causing the latter's activation to spark the semblance of the arrival of that sensation (**Figure 2**) (mechanisms other than the spread of EPSP through the functional LINKs were also discussed, Vadakkan, 2008). The LINK is called "functional" since (1) LINK formation is a function of the simultaneous activation of different sensory receptors during initial exposures to the item (learning), and (2) reactivation of the LINK is a function of the arrival of EPSP at either one of the postsynapses.

Activation of a postsynapse without corresponding activation of its presynapse will provoke the virtual sensation of activity arriving at the presynapse from the sensory receptors of the second sensation. This is called synaptic semblance. Semblance means that the postsynapse that receives EPSP through the functional LINK experiences the cellular hallucination that it has received an action potential-induced synaptic transmission from its presynapse.

A large number of different items can induce the activation of the postsynapses B or D and induce synaptic semblance at the synapse on the opposite side of the functional LINK. Thus, even though the functional LINK was formed during binding from a specific item, it can be reactivated by activities of either postsynapses B or D induced by different items. In this context, functional LINKs can be regarded as transferable both during binding and retrieval of bound sensory features representing an item.

Concurrent with the synaptic semblance, the EPSPs reaching the postsynapses through the functional LINKs can travel down to the dendritic tree of the next neuronal order. These EPSPs can undergo spatial and/or temporal summation inducing an action potential at the axon hillock of the neuron. The activity from this neuron propagates in the downstream network of neurons that belong to the second sensory stimulus. Activation of these neurons induces a network semblance (**Figure 3**) that creates the hallucination of receiving a large subset of the sensory inputs from the second sensory stimulus. From the combined effects of the synaptic and network semblances, the nervous system will be able to receive the virtual sensation of a second sensory stimulus.

In order to translate the semblances to the characteristic features of the virtual sensation of the second stimulus, the following inductive reasoning approach is used. What does semblance means to sensory perception? In other words, how can we derive the sensory quality of a semblance? To answer this question, the sensory equivalent of activity at one presynapse need to be figured out as follows. Let us imagine that synaptic semblance occurs at the postsynapse D (**Figure 2**). Let presynapse C belongs to neuron Z

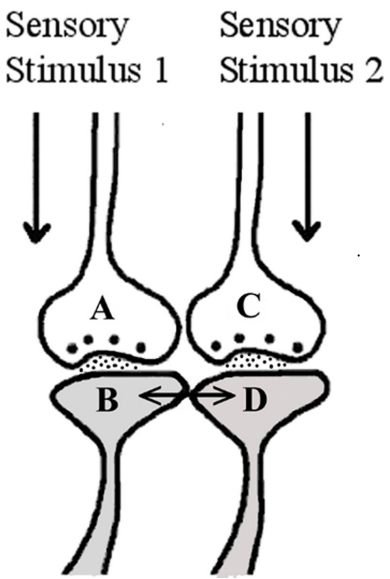


FIGURE 1 | Illustration showing the formation of functional LINK between two postsynaptic membranes B and D during binding. The functional LINK is transient and is a function of the simultaneous arrival of sensory stimuli from an item at the two postsynaptic membranes during binding. This is hypothesized to occur at locations of convergence of sensory inputs. Reactivation of the functional LINK is possible during the arrival of activity at either one of the postsynapses (B or D) at a later time (This figure is used after modification [Vadakkan, 2008] with permission from iUniverse publishers, Bloomington, IN, USA).

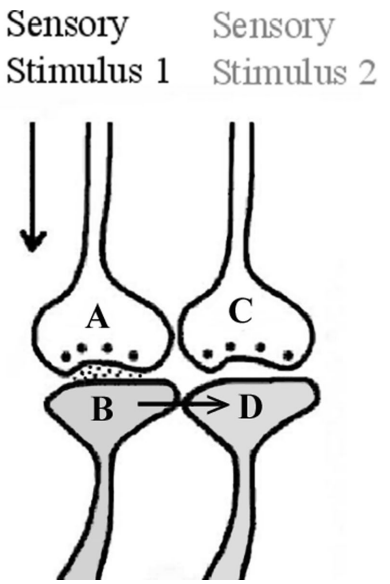


FIGURE 2 | As a result of the binding, when a single sensory stimulus from an item reaches the presynaptic terminal A, it depolarizes its postsynaptic membrane B, and the resulting EPSP spreads through the functional LINK to postsynaptic membrane D. This evokes the cellular hallucination at postsynapse D of an action potential reaching its presynaptic terminal, C. This is called "synaptic semblance". Note that a functional LINK can be reactivated by EPSP reaching it from either postsynapse B or D (This figure is used after modification (Vadakkan, 2008) with permission from iUniverse publishers, Bloomington, IN, USA).

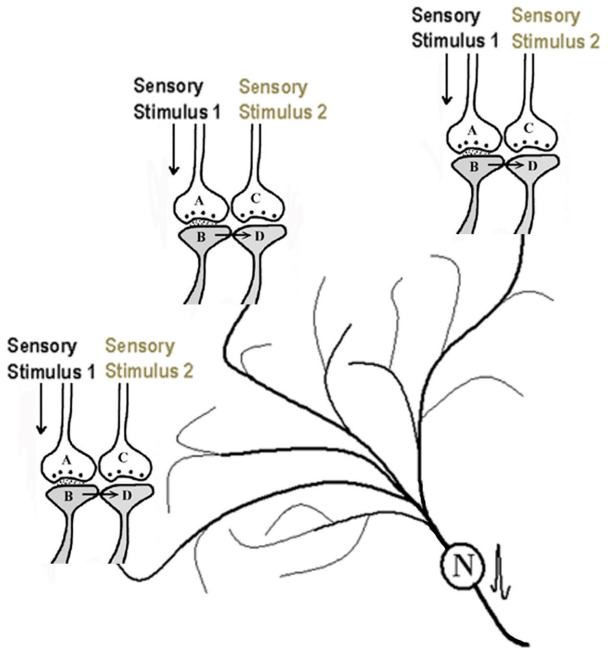


FIGURE 3 | Schematic representation of activation of a partial neuronal network occurring during the arrival of one sensory stimulus from an item after binding (the initial formation of the functional LINK). As explained in **Figure 2**, synaptic semblances are formed at each of the postsynapses marked D. In addition, the EPSPs from all the "D_n" postsynapses can summate (provided spatial and/or temporal summation of sufficient number of EPSPs is possible) to elicit an action potential at the axon hillock of neuron N. Neuron N belongs to the second sensation. Activating neuron N will induce the activation of a partial neuronal network that belongs to the second sensory item inducing network semblance.

(**Figures 4A,B**). Neuron Z, in turn, can only be depolarized by activating a set of axonal terminals belonging to the neurons that synapse to the dendritic spines (postsynapses) on neuron Z's dendritic tree.

The spatial summation of nearly 40 EPSPs or the temporal summation of much less than 40 EPSPs at the soma of neuron Z can trigger an action potential. Therefore, it is possible that activity from a multitude of possible combinations of any 40 EPSPs (from spatial summation) or permutations of less than 40 EPSPs (from temporal summation) arriving to the neuron Z's nearly 4×10^4 dendritic spines has the ability to trigger the same action potential.

Let us imagine that the dendritic tree of neuron Z receives activity from a set of neurons {Y} through existing synaptic connections that they make with the neuron Z. Similar to this, each neuron synapsing on the neurons in the neuronal set {Y} in turn receives synaptic connections from a set of all the possible combinations of neurons upstream from the set of neurons {Y}. Let this set of neurons be {X}. By moving in this way in a retrograde fashion toward the sensory level, sets of neurons at the preceding neuronal orders can be determined. The last step of this process will determine the set of sensory receptors whose activation can theoretically cause the activation of presynapse C. These sensory receptors are superimposed to obtain the sensory receptor map that represents possible sensory locations which can contribute to the activation of presynapse C (**Figure 4D**). This sensory receptor map provides the characteristic features of the virtual sensation that is produced by synaptic semblance occurring at the postsynapses.

Since the number of sensory receptors is finite, a large number of sensory inputs from various items can reactivate one functional LINK that was initially formed by binding of sensory stimuli from one item. Therefore, these functional LINKs will be shared for binding of sensory stimuli from different items. This property can be called transferability of the functional LINKs.

The net effect of synaptic and network semblances occurring during the activation of one sensation from an item that results in the virtual sensation of the second sensation in the latter's absence can be called functional semblance. By computing semblances (sensory receptor maps) from all the activated postsynapses, the quality of the functional semblance can be understood. The combination of semblances from different postsynapses at one neuronal order (that are activated almost simultaneously) (**Figure 4C**) and the permutation of these combination products occurring at the consecutive neuronal orders in a temporal fashion will result in a multi-dimensional net semblance (**Figure 4D**). The configuration of this net semblance is a function of the complexity of the nervous system and will depend on (1) the genetic make-up of the animal that determines the structural LINKs (structural LINKs can be explained as LINKs between the postsynaptic membranes that are genetically determined and are responsible for innate behaviors), number of neurons, orders of neurons, neuronal organization and convergence pattern of sensory inputs, and (2) functional LINKs acquired during life from previous binding events.

STEP 2: DERIVATION OF THE FRAMEWORK FOR CONSCIOUSNESS

Continuous formation of semblances is required for the formation of uninterrupted functional semblance for consciousness. The main inputs that provide a continuum of neuronal activities in the nervous system include:

(1) Background sensory inputs. Since an animal cannot possibly shut down all sensory inputs, its nervous system continues to receive background sensory stimuli from the environment. As a result, many sensory receptors are activated. This will induce many synaptic and network semblances contributing to the formation of C-semblance. C-semblance and consciousness are used interchangeably depending on the context.

(2) Neuronal oscillations. The cortex and hippocampus undergo oscillating neuronal activities both during wake-state and sleep. The sets of neurons that are oscillating are unlikely to get activated in the same order during any specific binding events or retrieval of bound sensations. The oscillating population of neurons will induce semblances by reactivating the functional LINKs at their synapses. Since neuronal oscillation occurs in a population of neurons that are highly non-selective for retrieval of any specific sensory item, the synaptic and network semblances induced can be viewed as contributing to C-semblance.

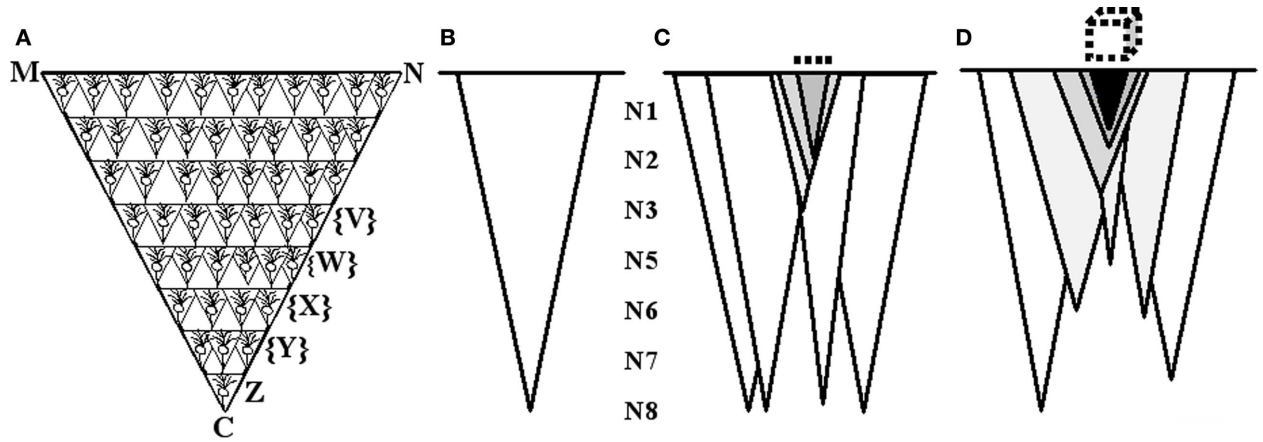


FIGURE 4 | Schematic representation of the formation of net semblance from synaptic and network semblances. (A) Cartoon showing the formation of semblance at one of the presynapses C of the neuron Z when its corresponding postsynapse D (see **Figure 2**) is activated through the functional LINK. The semblance of activity at the presynaptic terminal C indicate that the activity could be arriving from the neuron Z which in turn could be receiving activation from a set of neurons {Y}. In this manner, we can extrapolate the location of neurons of the penultimate orders in a retrograde fashion until we reach the sensory receptors. The final sensory receptor group is represented by the line MN which represents the identity of the sensation. From the sensory stimulation point of view, this reflects the fact that stimulation of the sensory receptors on the line MN can activate the presynaptic terminal C. **(B) A** is summarized and drawn as an isosceles triangle and it represents the semblance from the level of one synapse or neuron to the sensory level. **(C)** Formation of net semblance from a single order of neurons N8. The darkness of the shaded triangles represents overlapping of the semblances that determine the identity of the sensation. The darkest triangular area represents the net semblance which in turn determines the identity of the retrieved sensation. The dotted line on the top represents identity of the sensation occurring from a single order of neurons. **(D)** Schematic representation illustrating the formation of semblances from multiple neuronal orders which are integrated to form multi-dimensional net semblance representing the identity of the sensation to be retrieved. The dotted cube on the top represents the multi-dimensional nature of the semblance. (This figure is used after modification [Vadakkan, 2008] with permission from iUniverse publishers, Bloomington, IN, USA).

Net functional semblance resulting from the background sensory inputs and neuronal oscillations can be viewed as primary semblance, namely C-semblance, intrinsic to the net semblance formation at a given period in an animal's development. This C-semblance is the basis for consciousness.

C-semblance is a function of the reactivated structural LINKs (which are a function of the complexity of the neuronal architecture of a given species) and functional LINKs (which in turn, are a function of the binding events that occurred prior to the time point when consciousness is tested) during background neuronal activity. The optimum possible C-semblance that can be elicited is considered species-specific. Variations in the individual features of the C-semblance can occur depending on (1) the number of functional LINKs formed during life, and (2) the number of functional LINKs reactivated by the background neuronal activities.

Dimensions of C-semblance

The net C-semblance is formed from the integration of all the synaptic and network semblances occurring at different orders of neurons. Due to the normal synaptic delay, semblances occur at different neuronal orders at different time points, requiring time as one of the dimensions for integration. Oscillating nature of the neuronal network is also likely to add new dimensions based on the phase of the oscillation. The final integration product is a multi-dimensional semblance that leads to the virtual sensation of some non-specific sensory stimuli, which is called C-semblance (**Figure 5**).

Dynamics of the formation of C-semblance

The contribution of synaptic and network semblances from different regions of the brain that contribute to the C-semblance may vary. Studying isolated lesions of different parts of the brain can help us understand these regional contributions. Brain lesions occurring at different brain locations lead to different levels of unconscious states (Bernat, 2009). Moderate-sized lesions of the cortex are unlikely to affect consciousness as evidenced by reports of lack of changes in conscious states even after the surgical removal of large portions of cortex (Austin and Grant, 1958). It is possible that semblance formation has uniform distribution in the cortex and that there are an excess of functional LINKs in the cortex than that are required to contribute to the formation of C-semblance.

Abrupt loss of consciousness at a critical concentration of an anesthetic agent suggests that the integrated mechanism underlying consciousness collapse non-linearly (Alkire et al., 2008). From this, it may be interpreted that the opposite is possible during the formation of C-semblance. That is, the conscious state is built from its elements as a non-linear exponential function. This means that most of the time, there are more semblances than what is required for C-semblance. This may explain why consciousness is not affected by the surgical removal of a portion of the cortex (Austin and Grant, 1958). On the other hand, changes that can affect the functional LINKs globally, for example, metabolic, electrolyte and toxic effects, may affect C-semblance quickly.

Synchronization of many neuronal activities required for oscillatory network may be useful in maintaining C-semblance. Firstly, it may contribute to a continuum of neuronal activities, without

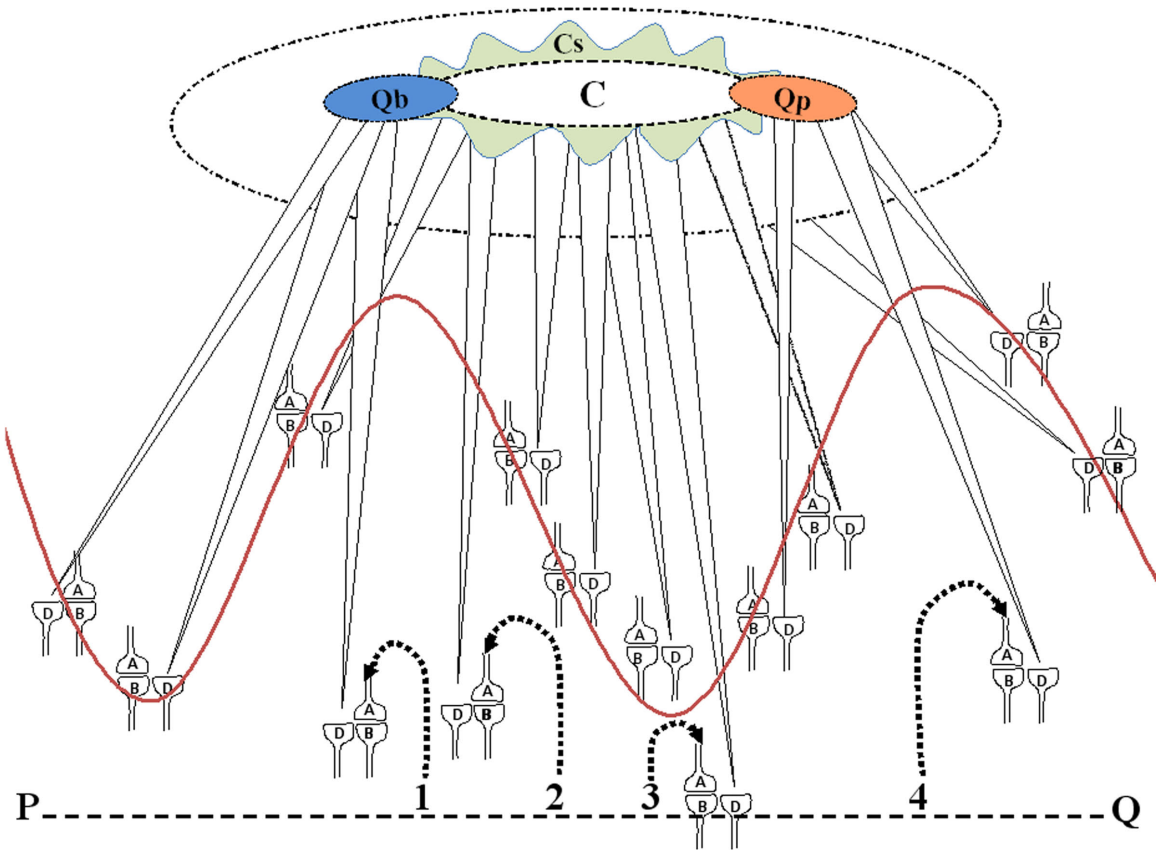


FIGURE 5 | Cartoon showing the formation of primary semblance for consciousness (C-semblance) from the semblances at the LINKed postsynapses as a result of the neuronal oscillations (represented by the wave form) and background sensory inputs (representative inputs are marked as 2 and 3 from the line PQ which represents the environment). Note the elliptical-shaped net functional semblance C representing consciousness. Around the elliptical-shaped C-semblance, is a wave-shaped extension that represents the subjective variations of C-semblance among the members of a species. Note the center-surround configuration of the C-semblance. The center denotes the species-specific features and the surround represents the subjective elements of the C-semblance. Formation of secondary semblances for qualia of blueness of the blue (Qb-semblance) and sensation of pain (Qp-semblance) are shown. For simplicity, synapses are drawn intersecting the wave form instead of the neuronal soma. The synaptic structures and their labeling are from **Figure 2**. Qb and Qp semblances depend on the C-semblance and are shown by their overlap. Semblances from the sensory inputs from the items in the environment (represented by sensory inputs marked 1 and 4 from the line PQ) provide qualia for the evoked sensation. C, C-semblance (consciousness semblance); Cs, subjective consciousness semblance.

any breaks, required for maintaining consciousness. In conditions where the synchrony is broken, it may affect C-semblance. An example of such a condition is absence seizure where conscious states are affected transiently only for a few seconds. Secondly, the synchronized network activity acts as a baseline upon which environmental sensory inputs can induce semblances for memories of associated sensations from different items in the environment. In these respects, synchronization of the neuronal activities may be considered as a binding factor.

Determinants of C-semblance

Species specificity. Every individual within a species is different in the specific number of neurons and fine details of connectivity patterns within its nervous system. However, all members of a species have a general organization that (1) provides similar distribution of structural LINKs at similar locations, and (2) provides opportunities for the formation of almost similar functional LINKs during life. Individuals of a species living in a given environment will associatively bind (learn) the sensory qualities of almost all the items and this is likely to lead to the formation of almost identical functional LINKs within their nervous systems. The evolving similarities in the structural and functional LINKs can lead to similar C-semblances. Depending on the variations in binding that taking place in each individual, the nervous system develops different number and locations of functional LINKs. These differences are discussed under subjective qualities of C-semblance.

Age of the nervous system. Since functional LINKs are likely absent in a newborn animal, C-semblance will depend on the genetically determined structural LINKs between certain postsynapses, called (S → F) LINKs. As the number of binding events an animal undergoes increases, the number of functional LINKs will increase. Since C-semblance in newborn and adult animals varies due to the differences in the number of functional LINKs, the nature of

consciousness may also vary. Since the conscious state is built from semblances in a non-linear exponential function (discussed in the previous section), the net semblance values after the curve becomes an asymptote indicates that the C-semblances of all members of a species with the same age and environmental exposure are nearly the same.

*Interactions of the sensory systems***.** Injuries of the sensory pathways (Guberman and Stuss, 1983; Steinke et al., 1992) can disrupt consciousness. Acute injuries that damage the brain stem may cause sensory deficits that can evoke an unconscious state (Parvizi and Damasio, 2003). Closure of the eyes induces changes in the electroencephalogram (EEG) wave forms due to changes in the oscillatory pattern of the cortical neurons. This is capable of initiating further changes in the configuration of the wake-state C-semblance. As the sleep cycle progresses, the EEG waveforms keep changing indicating concurrent changes in the configuration of the C-semblance eventually leading to various stages of unconscious stages of sleep.

C-SEMBLANCE AND ITS RELATION TO OTHER SEMBLANCES

The primary C-semblance is viewed as resulting from the semblances formed from background neuronal activity. In addition, there are many semblances that are formed in response to specific internal or external sensory inputs. Classifying them may provide a better understanding of different types of semblances and their similarities and differences in operation. Secondary semblances are evoked by specific sensory stimuli that reactivate specific functional LINKs. There are two types of secondary semblances: (1) innate semblances that are modified through developmental stages. Examples include those responsible for sight, sound, smell, touch, taste, and pain, (2) semblances formed through the functional LINKs and were acquired through binding. Examples include memory and path finding.

Tertiary semblances can occur when a novel sensory stimulus is presented to a nervous system that has already completed many binding events in the past. The nervous system uses combinations, permutations, and transferability properties to induce net semblances as a response to this novel cue stimulus. Depending on the number of binding events carried out in the past, the nervous system can produce peaks of net semblances that provide novel response patterns. They can interact with external evidence that allows sensory data to veto or support internal constructs which is regarded as a requirement of a system with cognitive abilities (Abbott, 2008). Both secondary and tertiary semblances are likely to depend on the stability of the conformation of the primary C-semblance. The details of their interaction will need further exploration.

SUBJECTIVE QUALITIES OF C-SEMBLANCE

According to philosophical idealism, things that we perceive around us are the subjective creation of our minds and we have no possibility of knowing what the real world is like (Kant, 1781). Idealism also denies that what we perceive is part of the physical world (Behrendt, 2010). It is viewed that we have no real access to the objective world since we are observing it through our subjective conscious experience (Solomon, 1988). There is an undifferentiated world which exists independently of the observer and differentiation takes place at the level of the subject and only exists for the subject (Schopenhauer, 1844). In this context, there are concerns that attempts to associate consciousness to physical properties of the nervous system will lead to negligence of the subjective qualities of the conscious experience (Searle, 1992, 1997). Therefore, a framework for consciousness is expected to incorporate features that can explain these subjective qualities.

C-semblances are expected to be broadly similar among the age-matched members of a species with normal brain functions living in similar environmental conditions. With regard to its fine details, C-semblance is likely to be specific to individual nervous systems at any specific time point. As discussed in the earlier sections, C-semblance is the net effect of the semblances evoked by background neuronal activities. These background neuronal activities can be either induced by or promote reactivation of the functional LINKs formed as a result of specific sensory associations in the nervous system during life. Therefore, depending on the specific functional LINKs present in an individual nervous system, there will be subjective variations in the C-semblance (**Figure 5**).

For the purpose of simplicity by making pictorial representations, it may be assumed that the C-semblance has a center-surround conformation with center representing the basic C-semblance and the surround representing the subjective variations in consciousness depending on previous binding events (**Figure 5**). Since semblances are formed only through previously formed functional and structural LINKs, sensing the objective world will inevitably include semblances occurring through the functional LINKs formed during previous binding occurred in an individual.

Another major feature of the subjective aspect of consciousness is the awareness of "self," which is explained as the experiential sense of being a vital and self-coinciding subject of experience or first-person perspective on the world (Sass and Parnas, 2003). The present framework provides provisions to explain how this sensation can arise. Associative binding takes place between our physical body and our actions. Thus, associative binding occurs between the owner of actions and the actions themselves. From this it can be seen that actions, from an individual, specific to the individual form semblance for the owner that provides the sensation of self. This feature of "self" depends on the functional LINKs formed from previous binding events.

QUALIA

It is expected that a framework for consciousness should be able to support different functional features of qualia (Dennett, 1988; Ramachandran and Hirstein, 1997; Ramachandran and Hubbard, 2001). The perception of the blueness of blue, the painfulness of pain and a certain sound as noise, requires the operation of a mechanism beyond arrival of sensory inputs from the sensory endings to the higher neuronal orders. This is one of the reasons for the difficulty in formulating a mechanism for consciousness (Chalmers, 1995; Shear, 1997; Crick and Koch, 1998).

Based on the present hypothesis qualia can be viewed as secondary semblances namely, qualia semblances (Q-semblances) which are innate to a species, and modified through developmental stages. For example, the color semblance (Qc-semblance) can be viewed as the net effect of all the semblances evoked at different neuronal orders initiated by the activation of color-specific cones. Since

qualia can only be appreciated in a conscious state, Q-semblances will be dependent on the specific configuration of the primary C-semblance. Similar to the C-semblance, Q-semblances are also multi-dimensional in character and are formed by the combination and permutation of semblances. Q-semblances are dynamic in nature and are a function of reactivation of the LINKs by different sensory stimuli.

It is likely that in the absence of C-semblance, Q-semblances cannot exist. On the contrary, one can imagine conditions having C-semblance, without having certain types of Q-semblances. Even though these Q-semblances are likely to share majority of the semblances from which they were formed with that of the C-semblance, those semblances may not be obligatory for the C-semblance formation. When it comes to the qualia of the consciousness themselves, it is possible that their semblances overlap to a great extent with that of the C-semblance.

LEVELS OF CONSCIOUS AND UNCONSCIOUS STATES

Awareness, attention, and the ability to respond to sensory stimuli are used to estimate the level of consciousness. Based on the defects in these domains, various conscious states can exist. According to the present framework, changes in the conformation of the multi-dimensional C-semblance can explain these different states. For example, C-semblance sufficient to remain conscious, but not being conscious of a specific motor activity leading to a zombie state (Chalmers, 1996; Koch and Crick, 2001) is possible. During occasions of specific conformational states of C-semblance, a cue stimulus can activate a partial neuronal network stimulating a set of motor neurons (**Figure 3**) and lead to the performance of motor actions without the conscious awareness of the action. In a similar way, an inverse zombie-state-producing anesthesia awareness (Mashour and LaRock, 2008) can be explained.

Anoxic and traumatic brain injuries cause different chronic consciousness disorders (Bernat, 2009). In addition, many metabolic disorders and medical conditions produce varying levels of consciousness states. Based on the present framework, these disorders can induce different changes ranging from the blockage of sensory stimuli towards the functional LINKs to inhibition of reactivation of the functional LINKs, resulting in lack of formation of appropriate conformation of the C-semblance.

Conscious states are considered separate from attention (Merikle, 1992; Lamme, 2004; Koch and Tsuchiya, 2007). Similar to this, it was reported that awareness is not a prerequisite for learning (Clark and Squire, 1998). While a conscious state is viewed as a multi-dimensional C-semblance, attention may be viewed as a process required for efficient functional LINK formation for associative learning/binding.

SUPPORTING EVIDENCE

Any account of the neurobiological processes contributing to conscious perception has to be compatible with various known relationships between different brain functions. If semblances from postsynaptic activation through the functional LINKs are responsible for consciousness, then by disturbing the factors that maintain the functional LINKs, consciousness can be disturbed. This can be used as a test to examine the framework for consciousness in the future. Here, the present framework is examined using currently available data from various physiological and pathological conditions of consciousness.

Anesthesia

Anesthetic agents cause the inactivation of a complex brain region in the posterior parietal area (Alkire et al., 2008). At low doses of anesthetics, the sustained firing of neurons during normal aroused brain states changes to a bi-stable burst-pause pattern (Llinas and Steriade, 2006) and oscillates between a depolarized up-state and a hyperpolarized down-state (Steriade et al., 2001). The latter is characterized by complete stoppage of synaptic activity for 0.1 s or more. At higher concentrations of anesthetics, the down-state becomes more prolonged, inducing prolonged stoppage of synaptic activity.

It is possible that in the absence of synaptic activity, functional LINKs stop functioning, leading to a lack of formation of semblances, resulting in unconsciousness. Since time is one of the dimensions that determine the conformation of the C-semblance (by contributing to the permutation of semblances from different neuronal orders that occur at different time points due to synaptic delay) any break in the continuity of activity in the neuronal orders will lead to disturbance in consciousness.

Seizure disorders

It is known that during seizure activity, a large portion of the cortico–thalamic complex is engaged in strong hyper-synchronous stereotypic activity (Hudetz and Imas, 2007; Kroeger and Amzica, 2007). Since a very large number of non-specific neurons fire during a seizure activity, the semblances occurring at the postsynapses will add up to form a very large non-specific net semblance that changes the conformation of C-semblance. This supports the idea that C-semblance requires a specific conformation (or at least a range of feasible conformations) for consciousness.

Sleep

During sleep, there is breakdown of cortical effective connectivity (Massimini et al., 2005). Similar to the bi-stable state seen with anesthetics, the hyperpolarized down-states induced by sleep involve changes in background neuronal oscillations, as evidenced by EEG wave form changes (Sanchez-Vives and McCormick, 2000). As neuronal oscillatory patterns change, the conformation of net semblance changes and results in loss of consciousness.

Hypoxia and hypoglycemia

These can lead to a reduction in the levels of consciousness possibly by interfering with the reactivation of functional LINKs required for sufficient semblances for the formation of C-semblances. Consequent to the defects in consciousness, this can lead to defects in secondary and tertiary semblances.

Chronic consciousness disorders and delirium

The spectrum of clinical conditions of chronic disorders of consciousness range from minimally conscious state to coma (Bernat, 2009). It is possible that with each of these conditions, there is

a corresponding spectrum of multi-dimensional conformations of C-semblance. Similarly, delirium is a disorder with fluctuating levels of consciousness. Reducing the background sensory stimuli can worsen the symptoms of delirium, resulting in reduced conscious interactions with the external world. This highlights the role of background environmental stimuli in the formation of C-semblance.

Thalamic relay of sensory inputs

Activities from almost all sensations pass through different nuclei within the thalamus. Since consciousness depends on background sensory inputs, blocking their transmission can lead to changes in consciousness. The reports that thalamic infarcts cause unconsciousness (Guberman and Stuss, 1983; Steinke et al., 1992) support the present framework and explain the contribution of background sensory inputs to consciousness. In addition, during sleep thalamic neurons undergo a prolonged inhibitory rebound after a single burst (Pedroarena and Llinas, 1997), possibly blocking sensory stimuli from reaching the functional LINKs, interfering with the net semblance formation.

Hippocampus and conscious memory

Lesions of the hippocampus cause selective loss of declarative or explicit or conscious memory (Cohen and Squire, 1980). Damage to other specific areas of the brain impairs unconscious forms of memory (Bechara et al., 1995; Knowlton et al., 1996). It is reported that hippocampal networks may be required for awareness of the memories of learned item (Eichenbaum, 1999). These findings indicate that the convergence of multi-sensory inputs in the hippocampus and the resulting functional LINK formation at the hippocampus and at higher orders of neurons are necessary for specific semblances for explicit or conscious memory.

Unconscious memory and goal pursuits

Unconscious associative memory can occur without conscious awareness (Chaumon et al., 2008). In one study, re-exposure to an odor during sleep that had been presented as context during prior learning (Rasch et al., 2007), was shown to improve the retention of hippocampus-dependent declarative memories. It is possible that the functional LINKs were reactivated during the exposure and brought about changes similar to that during repetitive learning. Similarly, actions can be initiated unconsciously (Custers and Aarts, 2010) and can be explained based on the activation of motor neurons as a result of the activation of a partial neuronal network (**Figure 3**).

Color vision

Qualia of color are expected to occur from a specific multi-dimensional conformation. Therefore, it is reasonable to expect that reducing the number of neuronal orders through which stimuli from color travel, will affect the qualia by reducing semblances occurring at different neuronal orders. This is observed in the following experiment. As the rotating speed of a wheel of colors increases, perceptional abilities of individual colors decreases. The only color that can be perceived at high speed is white. The difficulty in perceiving individual colors at high speeds can be explained as follows.

Color perception is time-dependent and involves more neuronal orders than required for the quale of the color white. Since eyes cannot fixate on a single color on the rotating wheel for sufficient time, the sensory stimuli will not reach a sufficient number of neuronal orders. This will reduce the required number of neuronal orders from which permutation of semblances needs to be computed for the quale of individual colors. This results in the lack of quale for individual colors. At high speeds, overlapping of semblances for multiple colors from the initial orders of neurons will result in net semblance for color white (Qc-w-semblance).

According to the present framework, the net semblance for color vision requires semblances from more neuronal orders than are required for the color white. Patients with cortical achromatopsia are color blind; but they can see the color white. It is likely that in these patients, semblances are not formed at the higher orders of neurons and thus not contributed to the formation of net semblance for individual colors. Based on the present framework, it is possible to find lesions affecting connections to the higher orders or neuronal loss at the higher orders in these patients.

Modulations of conscious experience

A previous work (Bottini et al., 1995) has shown that the activation of the vestibular system aided conscious tactile perception in an individual who had lost all abilities of tactile perception due to damage of the sensory pathways of touch at areas before its convergence with the fibers from the vestibular system. Normally, vestibular and tactile systems share projections to the putamen, insula, somatosensory cortex II, premotor cortex, and supramarginal gyrus. The study reported that the vestibular system introduced bias in the neural system involved in body representation. The physiological mechanism of this phenomenon is not yet known.

Based on the present framework, many functional LINKs would have formed between the vestibular system and the pathways governing touch sensation at the locations of convergence of these sensations (in the patient mentioned in the above study, at the orders of neurons after the region of brain damage) during life. At a later time, vestibular sensations alone can induce semblances at the LINKed postsynapses for the touch sensation. The perception of touch sensation occurs only in a conscious state.

Binding problem

The identification of synaptic mechanisms that can explain binding of association of various sensations has been an unsolved issue for many years. From the example of the ripening of a banana, it can be seen that the yellow-soft-taste sensory triad creates associative functional LINKs between the postsynapses such that either one of these sensations will be sufficient to provide the semblance of the other two sensations.

FUNCTIONAL LINKs BETWEEN THE POSTSYNAPTIC MEMBRANES

Functional LINKs are defined as LINKs that are formed as a function of simultaneous arrival of EPSPs at the postsynapses that are LINKed during binding or reactivated by the arrival of EPSPs at either one of the postsynapses during the retrieval process. These

are named functional LINKs since their presence is a function of the arrival of EPSPs. The transient nature of this process may explain the difficulty in exploring its nature. The nature of the functional LINKs and its effect on consciousness can be tested at various levels as following.

1. At the synaptic level: It is anticipated that some specific changes should be taking place between the postsynapses during binding that will allow reactivation of the functional LINKs during retrieval of the bound sensations. The following potential mechanisms that can support the time-scales of retrieval of the bound sensations or memories require further examination.
   (a) Different locations of the nervous system show increased blood oxygenation level dependent (BOLD) signal-intensities in functional magnetic resonance imaging (fMRI) studies both during learning and memory retrieval (Cohen et al., 1999; LaBar and Cabeza, 2006). It is reasonable to assume that similar oxygen requiring processes also take place during binding. Oxygen is possibly involved not only to provide energy for the biological process, but also directly involved in the formation of the functional LINKs. Since hippocampus and amygdala are brain regions where different sensory inputs converge (Amaral and Witter, 1995; Doron and Ledoux, 1999), it is most likely that these areas develop large number of functional LINKs. A preliminary report suggestive of the presence of an oxygenation-state dependent functional LINKs was previously made (Vadakkan, 2008). Further exploration of the findings may provide more information about the nature of the functional LINKs. Role of oxygen in the formation and reactivation of the functional LINKs can be tested by using uncouplers of oxidative phosphorylation to examine the selective role of oxygen in the functional LINK formation. Artificially introducing changes at the locations of convergence of sensory inputs that can induce formation of functional LINKs in experimental conditions may be used to study their nature.
   (b) The potential spaces between the synapses namely the extracellular matrix (ECM) play an important role in keeping the ionic balance required for maintaining the resting membrane potential. Ephaptic changes that can be brought about between the postsynaptic membranes during binding and its subsequent effect on reactivation of either one of the membranes by activation of the other can be carried out by biophysical studies.
   (c) Membrane proteins that can sense the arrival of EPSPs in the neighboring postsynaptic membranes during retrieval of the identities of the bound sensations can induce synaptic semblance reminiscent of the arrival of action potentials from the latter's presynaptic terminals. Proteins that can sense the sequence of ionic changes in the common ECM space during the arrival of action potential may provide valuable information.
   (d) Inductive depolarization of one postsynapse by another is a feasible mechanism that can be tested. Electroconvulsive stimulation (ECS) that alters the level of consciousness during the post-procedure period may act through disturbing the membrane electrical properties.
2. At the neuronal level: Electrophysiological studies can be carried out to examine the effect of functional LINKs at the neuronal level. A study examining cue-associated learning of reward at the thalamo–amygdalar synapses in the lateral amygdala (Tye et al., 2008) had shown an increase in the miniature EPSP (mEPSP) amplitude after the learning event. This can be explained as a function of additional AMPA channel currents through the functionally LINKed postsynapses. The same work also showed learning-induced firing of additional neurons in the lateral amygdale. This can be explained in terms of activation of a partial neuronal network (**Figure 3**) through the functional LINKs. Since binding can be viewed as an associative learning, the above explanations can be further explored to understand the nature of the functional LINKs used in building the framework.
3. At the systems level: Examination of the effect of functional LINKs on neuronal oscillation can be studied by testing whether blocking the functional LINKs can affect the retrieval of the bound sensations. fMRI experiments to study the role of functional LINKs in achieving binding and its subsequent effect in the retrieval of the identities of the bound sensations from an item can be carried out.
4. At the behavioral level: If different behavioral effects can be explained based on the role of functional LINKs, then their relationship with consciousness can be examined. For example if functional LINKs can be used to explain some of the features of schizophrenia, then the relationship between consciousness and schizophrenia can be studied.
5. At the epistemological level: It is possible to examine the feasibility of the functional LINKs as a mechanism to explain different brain functions. This will allow us to examine the presence of any cross-correlation between these brain functions and consciousness. From similar approaches carried out to study the role of the functional LINKs in memory (Vadakkan, 2008), it was observed that oxygen, glucose, dopamine, and amphetamine likely augment the formation of the functional LINKs; whereas activation of glycine receptors and gamma amino-butyric acid (GABA)-activated chloride channels, electroconvulsive stimulation and anesthetic agents likely block the functional LINKs. The similarities of the effect of these agents on memory and consciousness can be explored further to understand the nature of the functional LINKs.

SIGNIFICANCE OF C-SEMBLANCE

C-semblance is viewed as the multi-dimensional virtual sensation of a sensory stimulus resulting from background sensory stimuli and oscillatory neuronal activities taking place at various levels during wake-state. Concurrent with the induction of semblances, the EPSPs pass through the functional LINKs to the postsynapses (dendritic spines) of the dendritic tree of the next order of neurons and induce action potentials, provided sufficient spatial and/or

temporal summation of EPSPs occurs. During these events, the dendritic trees of many neurons will be short of sufficient EPSPs to induce action potentials at their axonal hillocks. In these neurons that are activated by the background sensory stimuli and neuronal oscillations, the summated EPSPs reaching at sub-threshold levels provide important physiological functions.

At this stage where background neuronal activities that continue to evoke C-semblance keeping the animal conscious, a new cue stimulus reaching the nervous system induces additional functional LINKs that add minor increments in EPSPs and can trigger an action potential in the above-mentioned neurons (that belong to the second sensation) that are otherwise activated at sub-threshold levels. Activation of a set of neurons that belong to one sensation from an item will provide network semblance for retrieval of sensory features from other sensations. This enables the nervous system to utilize the property of binding for its functions. Thus, maintenance of C-semblance aids in retrieval of additional sensory features of an item when only one sensation from the item reaches the nervous system. The same mechanism can explain memory retrieval after learning.

DISCUSSION

The present work has synthesized a feasible framework using information from various fields of brain sciences. This framework is then examined for the feasibility to support most of the functional requirements for maintaining a conscious state during which other brain functions can be carried out. In summary, the net semblance of activity from the sensory receptors occurring through the transient functional LINKs between the postsynapses by the background sensory receptor activations and neuronal oscillations is explained as the basis of consciousness. The factors that affect the functional LINKs should change the conformation of C-semblance and lead to different levels of consciousness modifying other brain functions that are dependent on C-semblance.

There are difficult questions that need to be answered. For example, can the proposed framework explain basic language comprehension? How does one consciously experience the meaning of the word "and"? Even though, a feasible explanation may be given as follows, is it sufficient? Based on the previous associative learning experience, different semblances for the word "and" are evoked. From them, the right one can be chosen, once we provide the nervous system with the context. Now, the items that need to be added and the word "and" elicit another semblance that contributes to the conscious experience of the process. For example, the word "and" in a mathematical context can bring conscious experience of a plus (+) sign. In the context of mixing two different colors, one would consciously experience the word "and" as mixing them together.

Unique conscious experiences in very complex circumstances are often difficult to explain. However, the present framework provides innumerable combinatorial and permutation possibilities that can create a large number of complex multi-dimensional semblances from nearly 10^{15} synapses that are located in multiple neuronal orders. It is possible that retrieval of semblances by certain sensations can cause collapse of the multi-dimensional conformation of C-semblance, similar to discontinuity in the graphs of certain calculus functions leading to loss of consciousness.

The present framework has applied new assumptions that binding introduces reversible changes at the level of synapses and during retrieval a virtual sense of the bound sensations is created. The present work explains consciousness using a framework derived from synaptic plasticity changes. Even though the present work has shown several supportive evidences, empirical verification of the functional LINKs needs to be undertaken. It may be possible to carry out experimental approaches to manipulate the functional LINKs in order to study their effects on consciousness. In addition, it may also become possible to translate the framework for experiments in physical systems.

The present framework raises many challenging questions. How many orders of neurons are required for the conscious state? Can machines have consciousness? Is it wrong to assume that a machine that suitably simulates human brain has consciousness (Minsky, 1986)? Computational work may help to estimate the threshold numbers of neuronal orders that can produce the final effective semblance for consciousness. Since consciousness is a property of the internal state of the nervous system, devising methods to express it using proper read-outs may allow us to understand its nature. The present framework should be treated as unproven until it is verified against further supporting evidence.

ACKNOWLEDGMENTS

I acknowledge the support from the Neurosearch Center in Toronto, Canada. I thank the reviewers for their questioning minds and insight during various stages of the review process.

REFERENCES

Abbott, L. F. (2008). Theoretical neuroscience rising. *Neuron* 60, 489–495.

Alkire, M. T., Hudetz, A. G., and Tononi, G. (2008). Consciousness and anesthesia. *Science* 322, 876–880.

Amaral, D. G., and Witter, M. P. (1995). "Hippocampal formation," in *The Rat Nervous System*, 2nd Edn, ed. G. Pacinos (San Diego, CA: Academic Press), 443–493.

Austin, G. M., and Grant, F. C. (1958). Physiologic observations following total hemispherectomy in man. *Surgery* 38, 239–258.

Babiloni, C., Babiloni, F., Carducci, F., Cappa, S. F., Cincotti, F., Del Percio, C., Miniussi, C., Vito Moretti, D., Rossi, S., Sosta, K., and Rossini, P. M. (2004). Human cortical rhythms during visual delayed choice reaction time tasks. A high-resolution EEG study on normal aging. *Behav. Brain Res.* 153, 261–271.

Bechara, A., Tranel, D., Damasio, H., Adolphs, R., Rockland, C., and Damasio, A. R. (1995). Double dissociation of conditioning and declarative knowledge relative to the amygdala and hippocampus in humans. *Science* 269, 1115–1118.

Behrendt, R. P. (2010). Contribution of hippocampal region CA3 to consciousness and schizophrenic hallucinations. *Neurosci. Biobehav. Rev.* 34, 1121–1136.

Bernat, J. L. (2009). Chronic consciousness disorders. *Annu. Rev. Med.* 60, 381–392.

Blackmore, S. (2002). Crossing the chasm of consciousness. *Trends Cogn. Sci.* 6, 276–277.

Bottini, G., Paulesu, E., Sterzi, R., Warburton, E., Wise, R. J., Vallar, G., Frackowiak, R. S., and Frith, C. D. (1995). Modulation of conscious experience by peripheral sensory stimuli. *Nature* 376, 778–781.

Chalmers, D. (1995). Facing up to the problem of consciousness. *J. Conscious. Stud.* 2, 200–219.

Chalmers, D. (1996). *The Conscious Mind: In Search of a Fundamental Theory.* Oxford: Oxford University Press.

Chalmers, D. (1998). "On the search for the neural correlate of consciousness," in *Toward a Science of Consciousness II: The Second Tucson Discussions and Debates*, eds S. Hameroff, A. Kaszniak, and A. Scott (Cambridge: MIT Press), 219–230.

Chaumon, M., Drouet, V., and Tallon-Baudry, C. (2008). Unconscious associative memory affects visual processing before 100 ms. *J. Vis.* 8, 10 1–10.

Clark, R. E., and Squire, L. R. (1998). Classical conditioning and brain systems: the role of awareness. *Science* 280, 77–81.

Cohen, N. J., Ryan, J., Hunt, C., Romine, L., Wszalek, T., and Nash, C. (1999). Hippocampal system and declarative (relational) memory: summarizing the data from functional neuroimaging studies. *Hippocampus* 9, 83–98.

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

Crick, F., and Koch, C. (1998). Consciousness and neuroscience. *Cereb. Cortex* 8, 97–107.

Crick, F., and Koch, C. (2003). A framework for consciousness. *Nat. Neurosci.* 6, 119–126.

Custers, R., and Aarts, H. (2010). The unconscious will: how the pursuit of goals operates outside of conscious awareness. *Science* 329, 47–50.

Dennett, D. C. (1988). "Quining Qualia," in *Consciousness in Modern Science*, eds A. Marcel and E. Bisiach (Oxford: Oxford University Press), 121–158.

Doron, N. N., and Ledoux, J. E. (1999). Organization of projections to the lateral amygdala from auditory and visual areas of the thalamus in the rat. *J. Comp. Neurol.* 412, 383–409.

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

Eichenbaum, H. (1999). Conscious awareness, memory and the hippocampus. *Nat. Neurosci.* 2, 775–776.

Farah, M. J., and Murphy, N. (2009). Neuroscience and the soul. *Science* 323, 1168.

Gray, J. (1992). Consciousness on the scientific agenda. *Nature* 358, 277.

Guberman, A., and Stuss, D. (1983). The syndrome of bilateral paramedian thalamic infarction. *Neurology* 33, 540–546.

He, B. J., and Raichle, M. E. (2009). The fMRI signal, slow cortical potential and consciousness. *Trends Cogn. Sci.* 13, 302–309.

Hudetz, A. G., and Imas, O. A. (2007). Burst activation of the cerebral cortex by flash stimuli during isoflurane anesthesia in rats. *Anesthesiology* 107, 983–991.

Kant, I. (1781). *Critique of Pure Reason*, translation by J. M. D. Meiklejohn (2003). Mineola, NY: Dover.

Knowlton, B. J., Mangels, J. A., and Squire, L. R. (1996). A neostriatal habit learning system in humans. *Science* 273, 1399–1402.

Koch, C. (2009). Do not underestimate science. *Science* 325, 392.

Koch, C., and Crick, F. (2001). The zombie within. *Nature* 411, 893.

Koch, C., and Tsuchiya, N. (2007). Attention and consciousness: two distinct brain processes. *Trends Cogn. Sci.* 11, 16–22.

Kroeger, D., and Amzica, F. (2007). Hypersensitivity of the anesthesia-induced comatose brain. *J. Neurosci.* 27, 10597–10607.

LaBar, K. S., and Cabeza, R. (2006). Cognitive neuroscience of emotional memory. *Nat. Rev. Neurosci.* 7, 54–64.

Lamme, V. A. (2004). Separate neural definitions of visual consciousness and visual attention; a case for phenomenal awareness. *Neural. Netw.* 17, 861–872.

Lamme, V. A., and Spekreijse, H. (2000). Modulations of primary visual cortex activity representing attentive and conscious scene perception. *Front. Biosci.* 5, D232–D243.

Llinas, R. R., and Steriade, M. (2006). Bursting of thalamic neurons and states of vigilance. *J. Neurophysiol.* 95, 3297–3308.

Mashour, G. A., and LaRock, E. (2008). Inverse zombies, anesthesia awareness, and the hard problem of unconsciousness. *Conscious. Cogn.* 17, 1163–1168.

Massimini, M., Ferrarelli, F., Huber, R., Esser, S. K., Singh, H., and Tononi, G. (2005). Breakdown of cortical effective connectivity during sleep. *Science* 309, 2228–2232.

McGinn, C. (1989). Can we solve the mind-body problem? *Mind* 98, 349–366.

Merikle, P. M. (1992). Perception without awareness. Critical issues. *Am. Psychol.* 47, 792–795.

Minsky, M. (1986). *The Society of Mind.* New York: Simon and Schuster Inc. Publishers.

Minsky, M. (1991). "*Machinery of Consciousness,*" Proceedings National Research Council of Canada, 75th Anniversary Symposium on Science in Society.

Orpwood, R. (2007). Neurobiological mechanisms underlying qualia. *J. Integr. Neurosci.* 6, 523–540.

Parvizi, J., and Damasio, A. R. (2003). Neuroanatomical correlates of brainstem coma. *Brain* 126, 1524–1536.

Pedroarena, C., and Llinas, R. (1997). Dendritic calcium conductances generate high-frequency oscillation in thalamocortical neurons. *Proc. Natl. Acad. Sci. U.S.A.* 94, 724–728.

Ramachandran, V. S., and Hirstein, W. (1997). Three laws of qualia; What neurology tells us about the biological functions of consciousness. *J. Conscious. Stud.* 4, 429–457.

Ramachandran, V. S., and Hubbard, E. M. (2001). Synaesthesia – a window into perception, thought and language. *J. Conscious. Stud.* 8, 3–34.

Rasch, B., Buchel, C., Gais, S., and Born, J. (2007). Odor cues during slow-wave sleep prompt declarative memory consolidation. *Science* 315, 1426–1429.

Rolls, E. T. (2007). A computational neuroscience approach to consciousness. *Neural. Netw.* 20, 962–982.

Rosenbaum, R. S., Stuss, D. T., Levine, B., and Tulving, E. (2007). Theory of mind is independent of episodic memory. *Science* 318, 1257.

Rosenblatt, F. (1961). *Principles of Neurodynamics: Perceptions and the Theory of Brain Mechanisms.* Washington, DC: Spartan Books.

Roskies, A. L. (1999). The binding problem. *Neuron* 24, 111–125.

Sanchez-Vives, M. V., and McCormick, D. A. (2000). Cellular and network mechanisms of rhythmic recurrent activity in neocortex. *Nat. Neurosci.* 3, 1027–1034.

Sass, L. A., and Parnas, J. (2003). Schizophrenia, consciousness, and the self. *Schizophr. Bull.* 29, 427–444.

Schopenhauer, A. (1844). *The World as Will and Representation*, translated by E. F. J. Payne, 2 vols. New York: Dover.

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

Searle, J. R. (1995). *Construction of Social Reality*. New York: The Free Press.

Searle, J. R. (1997). *The Mystery of Consciousness*. London: Granta Books.

Seth, A. K. (2008). Causal networks in simulated neural systems. *Cogn. Neurodyn.* 2, 49–64.

Shear, J. (1997). *Explaining Consciousness: The Hard Problem*. Cambridge, MA: MIT Press.

Solomon, R. (1988). *Continental Philosophy since 1750: The Rise and Fall of the Self*. Oxford: Oxford University Press.

Steinke, W., Sacco, R. L., Mohr, J. P., Foulkes, M. A., Tatemichi, T. K., Wolf, P. A., Price, T. R., and Hier, D. B. (1992). Thalamic stroke. Presentation and prognosis of infarcts and hemorrhages. *Arch Neurol* 49, 703–710.

Steriade, M., Timofeev, I., and Grenier, F. (2001). Natural waking and sleep states: a view from inside neocortical neurons. *J. Neurophysiol.* 85, 1969–1985.

Tallon-Baudry, C. (2009). The roles of gamma-band oscillatory synchrony in human visual cognition. *Front. Biosci.* 14, 321–332.

Tononi, G. (2004). An information integration theory of consciousness. *BMC Neurosci.* 5, 42.

Tononi, G., and Edelman, G. M. (1998). Consciousness and complexity. *Science* 282, 1846–1851.

Tye, K. M., Stuber, G. D., de Ridder, B., Bonci, A., and Janak, P. H. (2008). Rapid strengthening of thalamo-amygdala synapses mediates cue-reward learning. *Nature* 453, 1253–1257.

Vadakkan, K. I. (2008). *Semblance Hypothesis of Memory*. Bloomington, USA: iUniverse Publishers.

Vadakkan, K. I. (2010). *Framework of Consciousness from the Semblance Hypothesis of Memory. Toward A Science of Consciousness*. Tucson: University of Arizona, Abstract P-8.

von der Malsburg, C. (1981). "The correlation theory of brain function. MPI Biophysical Chemistry, Internal Report 81-2. Reprinted in Models of Neural Networks II (1994)," in *Models of Neural Networks II,* eds E. Domany, J. L. van Hemmen, and K. Schulten (Berlin: Springer), 95–119.

Wyart, V., and Tallon-Baudry, C. (2009). How ongoing fluctuations in human visual cortex predict perceptual awareness: baseline shift versus decision bias. *J Neurosci* 29, 8715–8725.

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

Received: 26 July 2010; accepted: 23 September 2010; published online: 26 October 2010.

Citation: Vadakkan KI (2010) Framework of consciousness from semblance of activity at functionally LINKed postsynaptic membranes. Front. Psychology ***1****:168. doi: 10.3389/fpsyg.2010.00168*

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

frontiers in
PSYCHOLOGY

OPINION ARTICLE
published: 11 May 2011
doi: 10.3389/fpsyg.2011.00087

What brain plasticity reveals about the nature of consciousness: commentary

Chris D. Frith*[1,2]

[1] Center for Functional Integrative Neuroscience, Aarhus University, Aarhus, Denmark
[2] Wellcome Trust Centre for Neuroimaging, University College London, London, UK
*Correspondence: cfrith@fil.ion.ucl.ac.uk

WHAT IS CONSCIOUSNESS?

Consciousness continues to be an ill-defined concept, so I shall start by discussing how the term is used in this special issue. As discussed in Overgaard and Overgaard (2010), there is an important distinction between level of consciousness and content of consciousness. Level of consciousness refers to a dimension that varies from coma at one extreme, through sleep and, at the other extreme, alert wakefulness. Philosophers call this *creature consciousness* since it applies to the whole creature (Rosenthal, 2009). Level of consciousness is of particular relevance to the studies of patients in persistent vegetative state discussed by Laureys and colleagues (Demertzi et al., 2011). However, most of the contributions to this special issue are concerned with the content of consciousness.

Alert wakefulness is characterized by consciousness of specific mental states. The states that we are aware of are the contents of consciousness. Philosophers call this *state consciousness*. This is somewhat confusing, given that, when people talk about altered states of consciousness, they are usually referring to different levels of consciousness rather than different contents of consciousness. So I will continue to use the terms level and contents of consciousness.

In studies of the neural correlates of consciousness there is great interest in contrasting the neural activity associated with stimuli that influence the contents of consciousness with the neural activity associated with the same stimuli when they affect behavior in the absence of any change in the content of consciousness (Frith et al., 1999). When a stimulus elicits neural activity and affects behavior it does not necessarily follow that we are aware of that stimulus.

A certain minimal level of consciousness is necessary for there to be any contents of consciousness, but the level of consciousness does not determine what the contents of consciousness will be. As is demonstrated in this special issue, brain plasticity has an important role in determining the contents of consciousness.

There is also a reflective aspect of consciousness which is modeled by Cleeremans (2011) and which Allen and Williams (2011) suggest may be uniquely human. Are conscious mental states thoughts about thoughts? Is consciousness by its very nature reflective? There is clearly a relationship between this aspect of consciousness and metacognition. I shall return to consideration of this relationship at the end of this introduction.

CERTAIN NEURAL STRUCTURES ARE NECESSARY, BUT NOT SUFFICIENT FOR CONSCIOUSNESS

It is well established that brain lesions can alter the contents of consciousness. To take just one example, lesions to the extra-striate cortex can eliminate awareness of color (achromatopsia, Zeki, 1990). Brain stimulation can also generate the contents of consciousness. For example, electrical stimulation of extra-striate cortex can generate hallucinations of various visual features including color (Lee et al., 2000). This has lead to the idea of essential nodes for the specific conscious contents (e.g., V4 for color, Zeki and Bartels, 1999). In accord with this idea, Silvanto and Rees (2011) conclude that, in the absence of primary visual cortex, humans seem to have extremely limited capacity for visual awareness (but see Ffytche and Zeki, 2011).

However, these essential nodes are clearly not sufficient for consciousness. If the level of consciousness is too low, as in coma, then experience does not occur even though the essential nodes are intact. Laureys and colleagues (Demertzi et al., 2011) suggest that the lack of awareness in such cases is due to the loss of certain kinds of long-range connectivity in the brain. Similar effects may be produced by anesthesia in which long-ranged connectivity is lost temporarily (Alkire et al., 2008).

BRAIN PLASTICITY PROVIDES IMPORTANT CLUES FOR UNDERSTANDING THE NATURE OF CONSCIOUSNESS AND ITS RELATION TO THE BRAIN

The theme of this special issue is the observation that the loss of awareness associated with brain damage is frequently reversed. Unilateral spatial neglect, for example, is a disorder of consciousness associated with stroke from which recovery can be quite rapid (Cappa and Perani, 2010). In some cases it may be that brain tissue has been temporarily been deactivated, and subsequently recovers. However, in the many cases where brain tissue has been permanently damaged, we have to ask how such recovery is possible. The doctrine of essential nodes would suggest that, if the node has been destroyed, recovery should not be possible.

Mogensen (2011) presents an excellent discussion of this problem. Does recovery depend upon the growth of new connections? Does the patient develop new cognitive strategies? One important conclusion is that the brain activity supporting recovery need not be in the same location as that originally supporting the experience. This observation supports two conclusions: (1) Conscious experience (qualia) can be re-acquired through some sort of learning process. (2) The nature of the experience (qualia) is not solely determined by the nature/location of the brain activity supporting it. This second conclusion is dramatically illustrated in the studies from Kupers et al. (2011) in which congenitally blind subjects were trained to "see" using a tactile stimulator. This technique involves turning the 2D images picked up by a video camera into a corresponding pattern of 2D tactile stimulation applied to the tongue. After being trained to recognize simple patterns with this stimulator, brain imaging revealed that performance of the task elicited activity in visual cortex. In addition

transcranial magnetic stimulation applied to visual cortex lead to the experience tactile qualia. This is evidence against the idea, know as cortical dominance (Hurley and Noë, 2003), that qualia are determined by the cortical location of the associated brain activity. But what then is the property of nervous activity that determines the difference in the experience of the different senses?

NEW QUALIA CAN BE LEARNED

However, it is not only after brain damage that qualia can be relearned and even learned for the first time. There are many examples of learning in the normal case. For example, between 6 and 12 months infants lose awareness of speech sound distinctions not present in their native language. With sufficiently early intervention this loss of awareness can be reversed, but interestingly only through direct interaction with a speaker, rather than passive exposure to audio or video-tapes (Kuhl et al., 2003).

Normal subjects can also learn to become conscious of stimuli previously outside awareness. Schwiedrzik et al. (2009) used meta-contrast masking to achieve chance performance in the detection of stimuli. After 5 days of training sensitivity was significantly increased and subjects reported awareness of the stimuli. Gottfried and his colleagues (Li et al., 2008) exposed volunteers to odor molecules (rose oxide and 2-butanol) that exist in two mirror image forms (enantiomers). At the beginning of the experiment the participants were entirely unable to smell any differences between the two mirror image forms, as is the case for most people. After only seven trials (for each odor) of standard Pavlovian conditioning, participants exhibited fear responses to the odor associated with shock and not to the other form, indicating that they now could distinguish between the mirror image odors. A further perceptual experiment showed that participants could now consciously detect the difference in smell.

In these examples, it seems likely that the potential to make perceptual distinctions was already present in the brain and that training revealed and enhanced this ability. For example, given the nature of the human eye we would not expect training to lead to awareness of infrared or ultraviolet light.

But the potential for awareness can also be artificially modified. Genetic manipulation in both mice and monkeys can alter the perception of color. Male squirrel monkeys are normally dichromats, but, even as adults, can be turned into trichromats through insertion of the missing opsin gene (Mancuso et al., 2009).

BUT HOW DO WE COMPARE QUALIA?

The observation that new qualia can be acquired, whether through learning or gene therapy, reminds us of a fundamental problem in consciousness research. How can we compare qualia from one person to another? Or within the same person at different times? As Overgaard and Mogensen (2011) ask, when a brain damaged patient recovers an awareness that had been lost, how can we know whether it is the same as the awareness that was present before the brain damage? If this recovered awareness is instantiated by activity in a different brain region and depends upon a different cognitive strategy it might well be different. Are there methods for determining whether two seemingly identical conscious states are actually different?

We have long known that people do have different sensory experiences. An obvious example is color blindness. The presence of the receptors necessary for color vision is under genetic control and some people have only two receptors instead of three, leading to different forms of color blindness (dichromacy), depending on which particular pigment is missing. The visual qualia of the color blind is clearly different, but trichromats still have some idea of what color blindness is like. It has now been found that some women have more than three retinal photopigment genes. These women also perceive significantly more color appearances than men or women with the usual three photopigment genes (Jameson et al., 2001). In this case the discovery of the biological difference led to the identification of the difference in the experience of color that can be explored empirically by asking subjects to make fine color discriminations.

Another example concerns individual differences in the spatial extent of primary visual cortex (V1). People with larger V1 are more susceptible to size illusions (Schwarzkopf et al., 2011). Having identified this biological difference we now explore the idea that these people have a subtly different experience of space.

Probably the most striking success in comparing qualia across people is Bartoshuk's et al. (2004) demonstration of the existence of supertasters. These are people who experience the sense of taste with a far greater intensity than average. This discovery depended upon the development of scales for subjective experience that do not eliminate individual differences. There is still much work to be done in developing scales for quantifying subjective experience (see, for example, Sandberg et al., 2010), but it is clearly possible for such comparisons to be made.

HOW CAN WE LEARN NEW QUALIA?

Outside the laboratory human beings spend a lot of time in discussing their experiences. We enjoy telling each other what something was like. When we share experiences with others in this way, we can learn about two kinds of things. We can learn that other people have different experiences from ourselves. However, by pooling our experiences we can also get a better estimate of that the world is like, since, most of the time, two heads are better than one (Bahrami et al., 2010).

In order to pool our experiences we need to down play our differences and take the best features from each experience. Since successful joint action (as well as joint perception) depends upon such pooling, this may be why we are so often unaware of subtle, but consistent differences in experience. The implication is that, as a result of sharing experiences, our qualia may shift toward that of the person we are sharing with. I predict that the greatest shift will occur in the least expert member of the group. So I find most plausible the suggestion from Allen and Williams (2011), that we learn new qualia by interacting with others. This seems to be the case, for example, with activities like wine tasting (Smith, 2007). But for sharing our experiences we have to introspect upon and communicate our experience. This requirement emphasizes the reflective aspect of consciousness that is probably uniquely human. Reflecting upon our own experience is an example of metacognition, that is thinking about our thoughts.

There are considerable advantages for concentrating on this aspect of consciousness since metacognition is more precisely defined. Furthermore powerful techniques

are now available for the quantification of metacognition (e.g., Galvin et al., 2003) and such measures have been applied to show that disruption of activity in dorsolateral prefrontal can change meta-cognitive sensitivity without altering discrimination performance (Rounis et al., 2010). Cleeremans (2011) uses the concept of metacognition to develop a computational model of how a brain can learn to be conscious by constructing a theory of its own behavior.

For me, a particularly interesting idea for further exploration is that this process of learning to be conscious of new things (i.e., to acquire new qualia) critically depends upon social interactions. In the various examples I mentioned above the learning of new qualia depended upon feedback from a teacher. To learn to experience the difference between the mirror image smell molecules required the experimenter to signal the distinction. More particularly, the American babies only learned to make the distinctions involved in Mandarin Chinese phonology through direct interaction with a speaker (Kuhl et al., 2003).

CONCLUSION

This special issue on the relevance of brain plasticity to the understanding of consciousness reminds us that consciousness, and the qualia that make up that consciousness, are not static. The contents of consciousness are constantly changing and developing through our experiences and especially through our sharing of experiences with others. Such change and development does not cease after brain damage. Indeed it is the dynamic relationship between brain and consciousness that enables the recovery of lost experience.

REFERENCES

Alkire, M. T., Hudetz, A. G., and Tononi, G. (2008). Consciousness and anesthesia. *Science* 322, 876–880.

Allen, M., and Williams, G. (2011). Consciousness, plasticity, and connectomics: the role of intersubjectivity in human cognition. *Front. Psychol.* 2:20. doi: 10.3389/fpsyg.2011.00020

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

Bartoshuk, L. M., Duffy, V. B., Green, B. G., Hoffman, H. J., Ko, C. W., Lucchina, L. A., Marks, L. E., Snyder, D. J., and Weiffenbach, J. M. (2004). Valid across-group comparisons with labeled scales: the gLMS versus magnitude matching. *Physiol. Behav.* 82, 109–114.

Cappa, S. F., and Perani, D. (2010). Imaging studies of recovery from unilateral neglect. *Exp. Brain Res.* 206, 237–241.

Cleeremans, A. (2011). The radical plasticity thesis: how the brain learns to be conscious. *Front. Psychol.* 2:86. doi: 10.3389/fpsyg.2011.00086

Demertzi, A., Schnakers, C., Soddu, A., Bruno, M.-A. l., Gosseries, O., Vanhaudenhuyse, A., and Laureys, S. (2011). Neural plasticity lessons from disorders of consciousness. *Front. Psychol.* 1:245. doi: 10.3389/fpsyg.2010.00245

Ffytche, D. H., and Zeki, S. (2011). The primary visual cortex, and feedback to it, are not necessary for conscious vision. *Brain* 134(Pt 1), 247–257.

Frith, C., Perry, R., and Lumer, E. (1999). The neural correlates of conscious experience: an experimental framework. *Trends Cogn. Sci. (Regul. Ed.)* 3, 105–114.

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

Hurley, S., and Noë, A. (2003). Neural plasticity and consciousness. *Biol. Philos.* 18, 131–168.

Jameson, K. A., Highnote, S. M., and Wasserman, L. M. (2001). Richer color experience in observers with multiple photopigment opsin genes. *Psychon. Bull. Rev.* 8, 244–261.

Kuhl, P. K., Tsao, F. M., and Liu, H. M. (2003). Foreign-language experience in infancy: effects of short-term exposure and social interaction on phonetic learning. *Proc. Natl. Acad. Sci. U.S.A.* 100, 9096–9101.

Kupers, R., Pietrini, P., Ricciardi, E., and Ptito, M. (2011). The nature of consciousness in the visually-deprived brain. *Front. Psychol.* 2:19. doi: 10.3389/fpsyg.2011.00019

Lee, H. W., Hong, S. B., Seo, D. W., Tae, W. S., and Hong, S. C. (2000). Mapping of functional organization in human visual cortex: electrical cortical stimulation. *Neurology* 54, 849–854.

Li, W., Howard, J. D., Parrish, T. B., and Gottfried, J. A. (2008). Aversive learning enhances perceptual and cortical discrimination of indiscriminable odor cues. *Science* 319, 1842–1845.

Mancuso, K., Hauswirth, W. W., Li, Q., Connor, T. B., Kuchenbecker, J. A., Mauck, M. C., Neitz, J., and Neitz, M. (2009). Gene therapy for red-green colour blindness in adult primates. *Nature* 461, 784–787.

Mogensen, J. (2011). Reorganization of the injured brain: implications for studies of the neural substrate of cognition. *Front. Psychol.* 2:7. doi: 10.3389/fpsyg.2011.00007

Overgaard, M., and Mogensen, J. (2011). A framework for the study of multiple realizations: the importance of levels of analysis. *Front. Psychol.* 2:79. doi: 10.3389/fpsyg.2011.00079

Overgaard, M., and Overgaard, R. (2010). Neural correlates of contents and levels of consciousness. *Front. Psychol.* 1:164. doi: 10.3389/fpsyg.2010.00164

Rosenthal, D. (2009). "Concepts and definitions of consciousness," in *Encyclopedia of Consciousness*, ed. P. W. Banks (Amsterdam: Elsevier), 157–169.

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

Sandberg, K., Timmermans, B., Overgaard, M., and Cleeremans, A. (2010). Measuring consciousness: is one measure better than the other? *Conscious. Cogn.*19, 1069–1078.

Schwarzkopf, D. S., Song, C., and Rees, G. (2011). The surface area of human V1 predicts the subjective experience of object size. *Nat. Neurosci.* 14, 28–30.

Schwiedrzik, C. M., Singer, W., and Melloni, L. (2009). Sensitivity and perceptual awareness increase with practice in metacontrast masking. *J. Vis.* 9, 18 11–18.

Silvanto, J., and Rees, G. (2011). What does neural plasticity tell us about role of primary visual cortex (V1) in visual awareness? *Front. Psychol.* 2:6. doi: 10.3389/fpsyg.2011.00006

Smith, B. C. (ed.). (2007). *Questions of Taste: The Philosophy of Wine*. Oxford: Signal Books.

Zeki, S. (1990). A century of cerebral achromatopsia. *Brain* 113(Pt 6), 1721–1777.

Zeki, S., and Bartels, A. (1999). Toward a theory of visual consciousness. *Conscious. Cogn.* 8, 225–259.

Received: 08 March 2011; accepted: 26 April 2011; published online: 11 May 2011.

Citation: Frith CD (2011) What brain plasticity reveals about the nature of consciousness: commentary. Front. Psychology ***2****:87. doi: 10.3389/fpsyg.2011.00087*

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