Can Synchronization Explain Representational Content?

A Reply to Caspar M. Schwiedrzik

John-Dylan Haynes

Multivariate decoding provides an important tool for studying the representation and transformation of mental contents in the human brain. Specifically, decoding can be used to identify the neural correlates of contents of consciousness (NC-CCs). Decoding of functional magnetic resonance imaging (fMRI) signals has so far mostly revealed content-selectivity in sensory brain regions, but not in prefrontal cortex. The limitations of fMRI-decoding only permit cautious conclusions because fMRI signals are only indirectly related to neural coding. However, the role of prefrontal cortex in visual awareness is also questioned by other findings, reviewed in Schwiedrzik (this collection). Neural synchronization might offer an alternative to solving the binding problem by providing a computational means of integrating information encoded in distributed brain regions. However, it is unclear whether synchronization in itself serves as a coding dimension for visual features. Furthermore, other alternatives to synchronization, especially the role of spatial codes, need to be considered as potential solutions to the feature binding problem.

Author

John-Dylan Haynes
haynes@bccn-berlin.de
Charité – Universitätsmedizin Berlin
Berlin, Germany

Commentator

Caspar M. Schwiedrzik
cschwiedrz@rockefeller.edu
The Rockefeller University
New York, NY, U.S.A.

Editors

Thomas Metzinger
metzinger@uni-mainz.de
Johannes Gutenberg-Universität
Mainz, Germany

Jennifer M. Windt
jennifer.windt@monash.edu
Monash University
Melbourne, Australia

1 Introduction

Information-based approaches to brain function have been very successful in recent years (Pouget et al. 2000; Haynes & Rees 2006; Kriegeskorte et al. 2006). Most importantly, they allow to study how mental contents are represented and transformed during information processing in the brain. My target article in this volume (Haynes this collection) emphasized the importance of an information-based approach for the study of human consciousness, especially for understanding the neural mechanisms of visual awareness. Whereas many previous studies mainly aimed to establish which additional processing needs to occur for a stimulus to reach awareness, a second question is equally important: how and where the brain encodes the specific contents of consciousness. Research on these neural correlates of the contents of consciousness (NCCCs; Chalmers 2000; Block 2007; Koch 2004) has been sparse. For identifying NCCCs, simply establishing that a brain area responds stronger under conscious than un-

der unconscious processing is not sufficient, because this could merely reflect unspecific processes such as attention or memory (Corbetta & Shulman 2002; Goldman-Rakic 1995). Instead, for identifying the neural code of contents several specific questions need to be addressed: Which brain regions encode sensory information in a representational space that exactly matches perception? And under which circumstances does a crossing of the threshold to awareness involve changes of representations in these specialized coding spaces?

2 The role of prefrontal cortex

One example of the importance of considering content-based processing is the global workspace theory (Dehaene & Naccache 2001; Baars 2002). In specific readings of the theory, consciousness involves a distribution of sensory information from sensory cortices to parietal and prefrontal cortex (Haynes this collection). Increased activity in frontoparietal regions under conscious perception is seen as evidence for such a "broadcasting" of sensory information (Baars 2002). However, without additional support by information-based or representational analyses, increased activity in frontoparietal regions with increased awareness might simply reflect unspecific processes, say as in detecting or reporting a change in perception, rather than coding the sensory information itself. In several studies with functional magnetic resonance imaging (fMRI) we found no evidence for changes in prefrontal representation of sensory information under increased levels of awareness (reviewed in Haynes this collection). Thus, we found no evidence that sensory information is re-represented in prefrontal or parietal cortex.

At this important point, the comment by Schwiedrzik (this collection) adds further important details on the potential role of prefrontal cortex (PFC) in visual awareness (Crick & Koch 1995; Dehaene & Naccache 2001). In a first line of arguments Schwiedrzik provides more detail on a point briefly sketched in the original article (Haynes this collection), whether absence of decodable information in PFC might reflect limitations of fMRI-based pattern classi-

fication. FMRI decoding will only be able to access neural information that is encoded in specific formats and topologies (Chaimow et al. 2011). For example, if neurons with different tuning properties are randomly distributed within a voxel, then the voxel will not be able to pick up any information about these properties. Thus, a macroscopic clustering of cells with similar tuning properties is required for fMRI-decoding to pick up information. So information could be present in prefrontal cortex, but in a format that is not accessible to fMRI, not even with the increased sensitivity of multivariate analyses. This might explain the discrepancy between the absence of information in many fMRI studies and differential responsivity to stimulus features of cells in PFC in non-human primates (Pasternak & Greenlee 2005). Please note, however, that most of the evidence for sensory tuning in PFC is obtained under working memory paradigms, which also includes temporal two-alternative forced choice tasks (Romo et al. 1999). Thus, it is unclear whether this generalizes to "realtime" perceptual experience.

There are further challenges in accessing neural information. Schwiedrzik (this collection) brings forward an important point already raised previously (Duncan & Owen 2000): coding in prefrontal cortex might be dynamic and thus the code might change across time. Such dynamically changing coding spaces might again not be detectable in classification analyses that assume a constant population code across the period analysed (Stokes et al. 2013). These points raised by Schwiedrzik are fully valid: It is highly important to consider these limitations when interpreting the results of fMRI decoding studies. To some degree these challenges might be alleviated with future technical developments. For example, columnar-level information can be accessed following recent advances in high-resolution fMRI (Yacoub et al. 2008). However, many limitations of fMRI will remain due to its vascular origin that only samples neural information indirectly. Please note that the limitations go far beyond the points raised by Schwiedrzik (this collection). For example, fMRI might not only miss information, but it might

also tap into information that is not available at the level of single neurons. Say, if an fMRI voxel samples a homogenous group of cells with highly similar tuning properties, the voxel might reflect a degree of averaging that is not available at the level of single neurons.

Please note, that the target article was not restricted to decoding approaches in fMRI alone. Instead, the aim was to outline a more general approach to studying the neural correlates of the contents of consciousness. If suitable recording techniques were available the information could be assessed based on a whole family of potential representational signals, including especially axonal and dendritic population activity. Please further note, that any recording technique has its blind spots. For example single-cell electrophysiology is biased towards large cells (Bartels et al. 2008), or optical imaging with voltage sensitive dyes is restricted to the surface of the brain (Grinvald & Hildesheim 2004). Thus, only a combination of techniques will be able to provide a full picture of the changes in neural coding with varying levels of visual awareness. Importantly, in a second line of argumentation Schwiedrzik (this collection) provides additional support and plausibility to the finding that the absence of information in PFC is real. For example the effects of lesions in PFC on visual recognition can be strikingly weak, which would not be expected if representation or routing of information in PFC were a necessary condition for awareness.

A third line of argumentation brings neural integration between spatially separated brain regions into play as a different potential mechanism of visual awareness (Engel & Singer 2001). The basic idea is that representation might involve a dynamic binding involving a synchronization of neural activity. According to this model, as Schwiedrzik points out, a distribution of sensory information into prefrontal brain regions might not be necessary. Many studies have related changes in neural synchronization to changes in visual awareness (see e.g., Engel & Singer 2001; Uhlhaas et al. 2009). However, it is important to look more closely at the explanation that can be obtained by changes in synchronization, specifically if the aim is to provide an explanation of the neural correlates of contents of consciousness (NCCCs). Typically, synchronization is not viewed as a coding dimension for contents, but as a code for binding and integration of features that are distributed across multiple content-specific regions (von der Malsburg 1999). The example provided in figure 1 of the comment by Schwiedrzik illustrates this nicely. A person views two superimposed clouds of moving dots, one green cloud moving left and a red cloud moving right. The features are encoded in content-specific fashion with two different activation patterns in the color area coding the two colors and two different activation patterns in the motion area coding the two motion directions. Synchronization between the neural representations of "green" and "left" on the one hand and "red" and "right" enables a separate binding of these two distributed features and also allows them to jointly be more effective in activating any downstream brain regions (König et al. 1996). Here, the contents are represented as differential activation states in the content-specific regions and their binding is achieved by synchronization. In this example representation and binding are separable problems based on separate computational mechanisms. However, Schwiedrzik (this collection) also goes one step further by suggesting that the large-scale connectivity patterns between brain regions might themselves code for different conscious contents. It has already been shown with fMRI that connectivity patterns between remote brain regions reflect changes in visual awareness (e.g., Haynes et al. 2005; Imamoglu et al. 2012). It has also been shown that connectivity matrices obtained with fMRI can be used to classify cognitive states (Richiardi et al. 2011; Heinzle et al. 2012). However, to date I am not aware of any evidence that fine-grained perceptual contents are encoded in differential patterns of brain connectivity. Furthermore, synchronization sometimes fails to explain perception (Thiele & Stoner 2003) and there other solutions to the binding problem besides synchronization. For example, high-level content-selective brain regions that achieve a certain degree of tolerance to variations in spatial location, still have con-

siderable information about the spatial location of features (e.g., Cichy et al. 2011). Thus, the spatial maps and their associated differential anatomical (as opposed to functional) connectivity patterns provide a plausible alternative hypothesis to synchronization (Treisman & Gelade 1980).

3 Conclusions

Prefrontal cortex is often considered vital for visual awareness (Crick & Koch 1995; Dehaene & Naccache 2001), however multivariate decoding studies have revealed a marked absence of sensory information in prefrontal cortex (Haynes this collection). Neural synchronization (Schwiedrzik this collection) might provide an alternative account for feature binding and selective routing of information. However, it is currently unclear whether any form of functional connectivity can itself code specific sensory contents.

References

Baars, B. J. (2002). The conscious access hypothesis: Origins and recent evidence. *Trends in Cognitive Sciences*, *6* (1), 47-52.

Bartels, A., Logothetis, N. K. & Moutoussis, K. (2008). fMRI and its interpretations: An illustration on directional selectivity in area V5/MT. *Trends in Neurosciences*, *31* (9), 444-453. 10.1016/j.tins.2008.06.004

Block, N. (2007). Consciousness, accessibility and the mesh between psychology and neuroscience. *Behavioral and Brain Sciences*, *30* (5-6), 481-499. 10.1017/S0140525X07002786

Chaimow, D., Yacoub, E., Ugurbil, K. & Shmuel, A. (2011). Modeling and analysis of mechanisms underlying fMRI-based decoding of information conveyed in cortical columns. *NeuroImage*, *56* (2), 627-642. 10.1016/j.neuroimage.2010.09.037

Chalmers, D. (2000). What is a neural correlate of consciousness? In T. Metzinger (Ed.) *Neural correlates of consciousness: Conceptual and empirical questions* (pp. 17-40). Boston, MA: MIT.

Cichy, R. M., Chen, Y. & Haynes, J. D. (2011). Encoding the identity and location of objects in human LOC. *NeuroImage*, *54* (3), 2297-2307. 10.1016/j.neuroimage.2010.09.044

Corbetta, M. & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, *3* (3), 201-2015.

Crick, F. & Koch, C. (1995). Are we aware of neural activity in primary visual cortex? *Nature*, *375* (6527), 121-123.

Dehaene, S. & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: Basic evidence and a workspace framework. *Cognition*, *79* (1-2), 1-37.

Duncan, J. & Owen, A. M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in Neurosciences*, *23* (10), 475-483.

Engel, A. K. & Singer, W. (2001). Temporal binding and the neural correlates of sensory awareness. *Trends in Cognitive Sciences*, *5* (1), 16-25.

Goldman-Rakic, P. S. (1995). Cellular basis of working memory. *Neuron*, *14* (3), 477-485.

Grinvald, A. & Hildesheim, R. (2004). VSDI: A new era in functional imaging of cortical dynamics. *Nature Reviews Neuroscience*, *5* (11), 874-885.

Haynes, J. D. (2015). An information-based approach to consciousness: Mental state decoding. In T. Metzinger & J. M. Windt (Eds.) *Open MIND*. Frankfurt a.M., GER: MIND Group.

Haynes, J. D., Driver, J. & Rees, G. (2005). Visibility reflects dynamic changes of effective connectivity between V1 and fusiform cortex. *Neuron*, *46* (5), 811-821.

Haynes, J. D. & Rees, G. (2006). Decoding mental states from brain activity in humans. *Nature Reviews Neuroscience*, *7* (7), 523-534.

Heinzle, J., Wenzel, M. A. & Haynes, J. D. (2012). Visuomotor functional network topology predicts upcoming tasks. *Journal of Neuroscience*, *9960* (9968), 475-483. 10.1523/JNEUROSCI.1604-12.2012

Imamoglu, F., Kahnt, T., Koch, C. & Haynes, J. D. (2012). Changes in functional connectivity support conscious object recognition. *NeuroImage*, *63* (4), 1909-1917. 10.1016/j.neuroimage.2012.07.056

Koch, C. (2004). *The quest for consciousness: A neurobiological approach.* Englewood: Roberts.

Kriegeskorte, N., Goebel, R. & Bandettini, P. (2006). Information-based functional brain mapping. *Proceedings of the National Academy of Sciences of the USA*, *103* (10), 3863-3868.

König, P., Engel, A. K. & Singer, W. (1996). Integrator or coincidence detector? The role of the cortical neuron revisited. *Trends in Neurosciences*, *19* (4), 130-137.

Pasternak, T. & Greenlee, M. W. (2005). Working memory in primate sensory systems. *Nature Reviews Neuroscience*, *6* (2), 97-107.

Pouget, A., Dayan, P. & Zemel, R. (2000). Information processing with population codes. *Nature Reviews Neuroscience*, *1* (2), 125-132.

Richiardi, J., Eryilmaz, H., Schwartz, S., Vuilleumier, P. & Van De Ville, D. (2011). Decoding brain states from fMRI connectivity graphs. *NeuroImage*, *56* (2), 616-626. 10.1016/j.neuroimage.2010.05.081

Romo, R., Brody, C. D., Hernández, A. & Lemus, L. (1999). Neuronal correlates of parametric working memory in the prefrontal cortex. *Nature*, *399* (6735), 470-473.

Schwiedrzik, C. (2015). What's up with prefrontal cortex? – A commentary on John-Dylan Haynes. In T. Metzinger & J. M. Windt (Eds.) *Open MIND*. Frankfurt a.M., GER: MIND Group.

Stokes, M. G., Kusunoki, M., Sigala, N., Nili, H., Gaffan, D. & Duncan, J. (2013). Dynamic coding for cognitive control in prefrontal cortex. *Neuron*, *78* (2), 364-375.

Thiele, A. & Stoner, G. (2003). Neuronal synchrony does not correlate with motion coherence in cortical area MT. *Nature*, *421* (6921), 366-370.

Treisman, A. M. & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12* (1), 97-136.

Uhlhaas, P. J., Pipa, G., Lima, B., Melloni, L., Neuenschwander, S., Nikolic, D. & Singer, W. (2009). Neural synchrony in cortical networks: History, concept and current status. *Frontiers in Integrative Neuroscience*, *3* (17). 10.3389/neuro.07.017.2009

von der Malsburg, C. (1999). The what and why of binding: The modeler's perspective. *Neuron*, *24* (1), 95-104.

Yacoub, E., Harel, N. & Ugurbil, K. (2008). High-field fMRI unveils orientation columns in humans. *Proceedings of the National Academy of Sciences of the USA*, *105* (30), 10607-10612. 10.1073/pnas.0804110105