Causal Considerations in Experimental Studies on Consciousness*

Borysław Paulewicz†
bpaulewicz@swps.edu.pl

Michał Wierzchoń‡
michal.wierzchon@uj.edu.pl

ABSTRACT

The consciousness studies are unavoidably linked to the mind-body problem, as most of the researchers are trying to investigate how a physical system (namely the brain) generates mental property (consciousness). To get around this problem the researchers often seem to endorse physicalism and identity theory but it is not clear to what extent, if any, does the research practice or the evidence based theories of consciousness depend on such assumptions. Here, we provide a few examples of studies investigating the so-called Neural Correlates of Consciousness (NCC). Based on these examples and the contents of the methodological papers addressing the problem of identifying the NCCs, we provide a novel, operational definition of the NCC. This allows us to shed some light on the meaning of various claims about the causal role of consciousness within the field of empirical studies. We also provide an inference rule for identifying the NCCs and the exhaustive list of the NCC confounders in the typical research paradigms.

Keywords: consciousness, awareness, neural correlates of consciousness (NCC), neural causes of consciousness, NCC confounders

1. Introduction

Identifying the neural underpinnings of consciousness remains one of the

* M.W. was supported by a SONATA BIS grant from the National Science Centre of Poland (2012/07/E/HS6/01037).

† University of Social Sciences and Humanities, Faculty in Katowice, Katowice, Poland.
‡ Consciousness Lab, Institute of Psychology, Jagiellonian University, Krakow, Poland.

main challenges within consciousness studies. Starting with the influential paper by Francis Crick and Christof Koch (1990) researchers are trying to locate the so called Neural Correlate of Consciousness (NCC), that is the brain activation that in some preferred sense reflects the difference between conscious and unconscious processing of stimuli (see also Boly, Seth, Wilke, Ingmundson, Baars, Laureys, Edelman, & Tsuchiya, 2013; Rees, Kreiman, & Koch, 2002 for the review). Here, we will be solely interested in the neural correlates of conscious content, rather than the level or the mode of consciousness. Perhaps the most influential definition of the NCC to date was given by Chalmers (2000), who observed that it is not enough for the neural correlate to be necessary for consciousness or to be sufficient for it. There are many correlates that seem to be necessary, but not sufficient. For example, certain kinds of visual experiences probably require primary visual cortex, but activation of the primary visual cortex is not sufficient for visual perception (see also the next sections of this paper). On the other hand, the sufficiency condition is too liberal since activation of the whole brain is sufficient for conscious experience, and yet the whole brain activation is clearly not what the researchers searching for the NCC have in mind. According to the Chalmers' definition the NCC is a minimal neural system sufficient for consciousness. It is minimal in a sense that no proper part of this system is also an NCC. Later in the paper we will argue that this definition could be replaced by an operational definition that does a better job of capturing what the researchers seem to have in mind when designing the experiments and interpreting the data. First however we will describe some important experiments that are often used to search for the NCCs.

2. Examples of the NCC Studies

In a typical experiment investigating the neural correlates of consciousness of some content the main source of information on the localization of the NCC is the observed difference between the brain activation when the participants declare to be aware of some content and the brain activation when they declare not to be aware of that content, or declare to be aware of some other content (Rees, Kreiman, & Koch, 2002). This activation can be measured with electroencephalography (EEG), functional magnetic resonance imaging (fMRI) or other electrophysiological or neurobiological methods. If some neural process, however defined, is different in those two conditions then the two events (neural process and verbal report) are correlated. This is called the contrastive

method (Baars, 1989) and it can be used to search for the neural correlates of any events of interest to the researcher.

Interestingly, in most of the studies the differences in verbal reports that are used to classify the brain activation patterns are not brought about by manipulating the stimuli. That is because most of the processing occurs unconsciously, and so most of the neural correlates of *stimuli* processing cannot be directly related to consciousness. Thus, the researchers usually fix the stimulus, i.e., try to match the stimulus presentation conditions so that, ideally, even though the stimulus is present in both cases, the only difference between the conditions is whether the participants are aware of the stimulus or not. This requires using the stimuli that are near-threshold, degraded or otherwise difficult to perceive or unstable. This way the contents of conscious experience can vary with time even though the stimulus remains approximately the same.

When the differences in verbal report or other content measure are observed in the absence of significant stimulus variability then, it is hoped, the true NCC can be revealed. For example, one may compare the visibility of a stimulus under bistable perception, where for the same stimulus, perception alternates between two percepts, or under binocular rivalry, where the perception alternates between two percepts of two stimuli that are constantly presented, each to a separate eye field, or under motion induced blindness, where the movement of the items appearing in the background results in the temporal changes in perception of the target stimuli presented around the fixation point. Based on the results of the studies with fixed stimuli several candidates for the NCC were found so far. For example, activation related to conscious processing of visual stimuli was found mainly in higher-order sensory areas and in parietal and prefrontal cortical areas (see Boly et al., 2013). Analogical effects were observed with event-related potentials (ERPs) registered with EEG, where consciousness correlated with the long-latency components (c.a. 300 ms after the stimulus presentation) involving frontal and parietal areas (Dehaene & Changeux, 2011).

Another way of dealing with the problem of unconscious processing is to use both the so-called objective and the subjective measures of conscious experience (see e.g. Timmermans & Cleeremans, 2015 for the review). The objective measures are responses indicating the ability to discriminate between the stimuli, such as stimulus identification or recognition. The subjective reports, also called subjective awareness measures, do not reflect the stimulus identity but are judgements about ones own experience, such as the clarity of percep-

tion or the level of confidence about the accuracy of the (objective) decision. It is often assumed that the objective measures may be sensitive to stimulus processing that precedes and may be independent of consciousness. Thus, only the subjective reports of visibility reflect the conscious processing of the stimuli (but see Block, 2011; Lamme, 2003; 2010). This rationale is clearly illustrated by the dissociations between objective performance and subjective reports, one striking example being the blindsight patients who lose conscious vision after lesions of the primary visual cortex but who still show various behavioural responses to visual stimuli (Ko & Lau, 2012).

The classical NCC paradigm has been criticized for revealing mainly the brain mechanisms related to a limited range of stimuli (mostly visual) and a limited range of tasks. Because of this limitation the activation of a certain brain area revealed by the contrastive method might simply be related to the fact that the brain is processing a certain kind of content. For example, activation of the fusiform face area (FFA) observed when participants claim to be aware of faces could be observed not due to the awareness of the stimuli but just due to the processing of facial content by the FFA regardless of any awareness of them (Rees, Kreiman, & Koch, 2002).

In response to the criticism of the classical NCC approach, researchers try to identify the brain activation patterns that can be observed in a larger set of experimental situations involving conscious perception. Multiple theories of the non-specific neural correlates of content have been proposed so far. For example, Crick and Koch (1990) suggest that consciousness is related to coherent and synchronous oscillations of all neurons engaged in the perception of a given object responsible for feature binding. Authors of the neural global workspace theory (Dehaene & Naccache, 2001; Dehaene, Changeux, Naccache, Sackur, & Sergent, 2006; Dehaene & Changeux, 2011) propose that the visual stimuli are brought to awareness due to the large-scale activation of neurons in the parietal, prefrontal and cingulate cortices that are internally synchronized in a reciprocal and long-lasting manner. Finally, according to Lamme (2003; 2010) visual awareness is an effect of recurrent interactions between the early visual and the late visual as well as the higher frontal and parietal brain areas.

It is beyond the scope of this paper to give a more detailed description of the studies investigating the NCCs and the theories proposed to account for the results. The interested reader will find good reviews elsewhere (Boly et al., 2013; Dehaene & Changeux, 2011; Rees, Kreiman, & Koch, 2002). For our immediate purposes it is important to note that most of the NCC studies

investigate brain activation that *co-occurs* with content reports and this correlation is often interpreted as evidence that some particular brain activation is among the *causes* of the contents of conscious experience. Several authors have pointed out recently (Wessel, 2012; Neisser, 2012) that the real intention of the researchers searching for the NCCs is to find the causes, not just the correlates, of consciousness. We will now take a look at what could this actually mean and to what extent do the classical paradigms allow for searching for the causes of consciousness.

3. Direct Evidence of the Influence of the Neural Correlates on Conscious Experience.

There are three basic ways in which evidence that goes beyond the correlations of brain activation with the reports of the participants can be collected (Wessel, 2012): observation of patients with structural brain damage, temporal lesions resulting from transcranial magnetic stimulation (TMS) and introduction of additional neural activity using the TMS, transcranial altering current stimulation (tACS), transcranial direct-current stimulation (tDCS) or deep brain stimulation techniques. Because the non-invasive intervention techniques allow for the experimental manipulation of the brain, in principle they make it possible to directly investigate the influence of brain activation on consciousness. Even structural brain damage can be thought of as a special case of experimental manipulation, because usually it can be safely assumed that no event that caused that damage is a common cause of the damage and the dependent variable.

The most widely discussed clinical condition in this context is blindsight, caused by the damage to the primary visual cortex. Blindsight patients are able to detect, discriminate and localize visual stimuli presented in their cortically blind field despite a denial of subjective awareness (see Kentridge, Heywood, & Weiskrantz, 1999; Ko & Lau, 2012). The condition was extensively studied both empirically (see e.g. Overgaard, Fehl, Mouridsen, Bergholt, & Cleeremans, 2008) and theoreticaly (see e.g. Block, 1995). It was interpreted as the evidence for the primary visual cortex being necessary, but not sufficient for consciousness of the visual stimuli (but see Block, 2011). This is because participants can detect, but cannot consciously perceive the stimuli that are presented in the cortically blind field (but see Overgaard et al., 2008 for alternative interpretations). This was further confirmed by the TMS studies by Rounis, Maniscalco, Rothwell, Passingham and Lau (2010) who showed that

when the TMS theta-burst stimulation was targeted at the dorsolateral prefrontal cortex (DLFPC), the subjective visibility of the stimulus was selectively decreased without any changes to the objective detection performance. The authors concluded that it is the activation of the prefrontal areas, rather than the primary visual cortex, that is a candidate for the NCC.

Another study directly showing the effect of brain activation on subjective reports was conducted by Fleming and colleagues (2015). Using the TMS stimulation targeted at the premotor cortex (SMA) they were able to show that the stimulation of the motor cortex associated with the response not chosen disrupts the subjective confidence of the response, but not the ability to discriminate the stimuli. This again suggests that awareness is associated with the activation of higher-level cortices and it highlights the role of action generation in metacognitive judgements. All of these results clearly show the fundamental role of experimental manipulations in identifying the directional effects of brain activation on consciousness. The observed NCCs are not only co-occurring with the conscious reports but they are shown to be the causes of the changes in verbal reports.

4. Two Notions of the Neural Cause of Consciousness

Several recent papers and conference presentations indicate that it came to be recognized that the real motivation for searching for the NCCs is not to find the neural correlates but the neural causes of consciousness. According to Wessel (2012) «the ultimate step for a neuroscientific theory [is to] go from a number of neural correlates to a definite set of biological causes» (p. 299). In the case of consciousness it is supposed to be «a unified, general mechanism underlying consciousness, without which consciousness cannot emerge, and which inevitably leads to consciousness when present» (p. 300). In a similar vein, Neisser (2012) points out that «NCC research is best characterized as an attempt to locate a causally relevant neural mechanism and that the first C in "NCC" should stand for "causes" rather than "correlates"» (p. 682).

The emphasis on causal relations is clearly visible in several recent critical essays (Aru, Bachmann, Singer, & Melloni, 2012; Bachmann, 2009; Gamez, 2014; de Graaf, Hsieh, & Sack, 2012) concerned with conceptual and methodological problems related to identifying the proper NCCs among the mere correlates. A major theme in all of those papers is the issue of confounders in the NCC studies. Roughly, a confound is a correlate that is not an element of the causal

relation of interest to the researcher. For example, when investigating the causal effect of health on height the common cause of both is a confound, because it provides an alternative causal explanation of the correlation between the height and the health. A great many neural events correlate with what is reported, but only a few such events are the NCCs. For example, Bachmann (2009) points out that a neural correlate of conscious content could actually be a cause (sometimes labeled as NCCpr where the suffix *pr* stands for predecessor) or a consequence (NCCco) of the proper NCC. Once the possibility of the NCCprs and the NCCcos is acknowledged they are immediately recognized as confounders that have to be dealt with when interpreting the results of the studies.

We believe that while it is true that the search for the NCCs is in a way the search for the causes of consciousness, there are two very different meanings of "causes of consciousness" at play here. One is related to what Hohwy and Bayne (2014) calls the horizontal relation between conscious experience and other processes or events. The elements of the horizontal relation are neural events or other physical events and the relation is a causal one. *Some neural event is horizontally related to consciousness if there is a causal relation between this event and the proper neural correlate of the mental event of interest.* The vertical relation is a relation between consciousness as a mental event and its proper neural correlate (the NCC). Being realized by, emerging from or supervening on are all examples of vertical relations between body and mind. The relations of emerging from, supervening on or being realized by do not have to be causal at all.

Both Neisser (2012) and Wessel (2012) seem to be mostly concerned with the *vertical* relation. Somewhat confusingly, according to these authors this vertical relation can and should be understood as a causal one. When Neisser claims that what the researchers are really interested in are the NCCs that are the causes of consciousness he seems to assume that the NCCs in question are already the proper ones and the question to investigate is how these proper NCCs cause conscious experiences. When Wessel talks about the biological causes of consciousness what he has in mind is a rather peculiar notion of causality; the biological causes of consciousness are the biological processes necessary and sufficient for consciousness. Clearly, the neural event A is necessary and sufficient for the mental event B only if A is in some sense the vertical equivalent of B. Wessel finds the prospect of finding such "causes" particularly challenging because he is primarily concerned with the vertical relation, albeit interpreted causally, and the vertical mind-body relation is obviously much more problematic than the horizontal relation between the neural events.

The causal problems associated with the search for the NCCs addressed in the methodologically oriented papers are not concerned with the vertical relation at all. Fo example, Bachmann (2009), Aru and colleagues (2012) and de Graaf and colleagues (2011) assume, either explicitly or implicitly, that there exist the proper NCCs and the task left to deal with is to find the proper NCCs among the mere correlates. According to all the aforementioned authors the mere correlates are either the NCCprs or the NCCcos, but only those two. The practice of identifying confounds of the NCCs depends on the underlying notion of the NCC. Once we agree on what a confound is we can use it to try to formulate the weakest operational definition that is descriptively accurate in a sense that it agrees with the research practice of identifying the confounds.

5. The Operational Definition of the NCC

The notion of causal predecessors and consequences of consciousness presupposes that there exists a causal chain, in normal circumstances originating at something that could be called a stimulus (either external or internal). This causal chain was explicitly considered by Gamez (2014). It leads through a series of intermediate neural cause and effect events to the event of changing the contents of consciousness. The change of content in turn has its own effects, ultimately leading to the changes in verbal responses or other reactions. This leads us to our operational definition of the NCC:

> The neural correlate of consciousness is the last stage in the neural causal path from the stimulus to the content based response that still causes changes in the content of the experience and through this change affects the change in the response.

The idea that something similar to this definition is tacitly assumed by the researchers studying consciousness explains why in the methodologically oriented literature the causal predecessors and consequences of the NCCs are considered problematic but the existence of the NCCs is not. The very idea of studying consciousness experimentally requires that there were some proper NCCs and that these neural equivalents of conscious experience had neural causes (Gamez, 2014). In fact, it could be argued that our operational definition follows from the assumption that there are any NCCs at all. The researchers studying the NCCs routinely ask causal questions about the relations of neural processes and experience but they rarely if ever explicitly address the

issue of vertical relation between neural events as neural events and conscious experiences as mental events. That is why we believe that our operational definition is both inevitable and minimal.

While there is a deep analogy between the Chalmers' definition and ours these two are certainly not equivalent. The sufficiency condition in the Chalmers' definition is replaced in our definition with the non-equivalent condition of being an element of a certain causal path. The minimality condition is replaced with the non-equivalent condition of being in a certain sense the last element of a certain causal path. Both definitions require that there was a more or less well defined set of candidates for the NCC and according to both definitions the proper NCC is the one that is in some way as close to consciousness as possible. However, the definition proposed by Chalmers seems to be at least partially concerned with the vertical relation. It entails that among the neural elements that are in some vertical (necessary, sufficient) sense candidates for the NCC there are the ones that are vertically related in the best way (minimally sufficient). The focus on vertical relations seems to be the reason why Chalmers introduced the notions of sufficiency and necessity in his analysis.

Thinking about the vertical mind-body relations as causal relations is problematic, as witnessed by the long history of philosophical debate on mental causation. Not only is it hard to reconcile the idea of mental events as mental events being causally related to neural events with the idea of conservation laws, the classical candidates for the vertical relation such as being realized by, emerging from or supervening on have a property that is not shared by most if not all real world causal relations; formally, the vertical relation between neural properties and mental properties is usually a function (perhaps relative to the enabling conditions) from neural properties to mental properties. That is, once the neural properties are fixed, all the mental properties are fixed as well. Real world causal relations are hardly ever like that, because the only way to have a physical causal relation that is a deterministic function from physical properties of the cause to the physical properties of the effect is to assume determinism and think of the cause as the state of the universe. Causally related events have a degree of autonomy, otherwise they wouldn't really be separate events. For certain purposes physical causal relations can be modelled by functions but they are inherently non-functional, because the target of causal relation is always influenced by the omitted factors, i.e., it is underdetermined with respect to the cause.

Both ours and Chalmers' definition is operational in the sense that each can be used as a guide in searching for the NCCs by empirical means. According to

the Chalmers' definition the researcher should look for a neural process such that when this process is operating, some conscious content *has to be* present (sufficiency). Once this sufficient neural process is found, the researcher should look for the ways of isolating the smallest part of this process that still corresponds to the conscious content always being present. Somewhat similar guidelines follow from our definition; the researcher should look for the causes of conscious content as reflected in the reports or any other measures and among these find the last effect of such causes that still affects the content.

Our operational definition is purely causal and it makes use of a weak notion of causality, one that doesn't require that the cause was sufficient for the effect. The causal definition immediately suggests an inference rule: one way to test if a given neural correlate of the report is an element of the right causal path is to manipulate the stimulus and, using virtual lesions, fix the neural correlate. If this breaks the causal connection between the stimulus and the response then the neural correlate of the response is probably an element of the causal path from the stimulus to the response. The only thing left to investigate is whether it is a predecessor of the NCC, a consequence of the NCC or the NCC itself. The NCC should not be located too early in the causal chain but also care must be taken not to locate it too far. There is always the risk that the neural cause of the response will be found, but this possibility can be tested by measuring different responses.

Neither definition requires that there was just one NCC for the given kind of conscious experience but both can be used as a guide in searching for the *non-specific* NCCs. The basic strategy is to find as many NCCs of the same content as possible by varying the conditions under which the correlates of the report are observed. The causal definition indicates that the primary candidates for the mere correlates are the neural processes closely associated with the stimuli or the responses. This means that once a candidate NCC is found the conditions of the task should be changed at least with respect to the stimuli or the responses to test if the candidate NCC is not a confound. Hopefully, the non-specific NCC will be revealed eventually as the neural processes common to all the specific NCCs. A prime example of an attempt at varying the response is recent work on the non-verbal reports of consciousness (see for example Einhäuser, Stout, Koch, & Carter, 2008).

Upon closer examination the causal definition also indicates that the idea, explicit in several recent methodological papers mentioned earlier, that the only confounders in the NCC research are the predecessors (NCCpr) or the

consequences (NCCco) of the NCCs is wrong. Perhaps it was Gamez (2014) who was the first to discover that in this context there are other kinds of confounders possible, but here we want to show how it follows from our definition. The causal definition has two parts, one concerned with there being a right kind of causal path, and the other concerned with there being a special place on that path. To say that the only correlates of the report are either the predecessors of the NCC, the consequences of the NCC or the NCC itself is to say that every correlate of the report is an element of the right causal path, i.e., the unidirectional causal path leading to the report and containing the NCC. However, this is not guaranteed by the fact that the responses observed reflect the contents of conscious experience.

Correlation between observed variables can result from one variable causing the other, from some common cause or from any combination of those two possibilities. The only kind of causal relation between two variables that does not induce the correlation is when the two variables are independent causes of a common third variable. It follows that, contrary to the already widespread belief, among the correlates of the report or the stimulus there are not two, but exactly five different kinds of confounders with respect to the NCC. There are 1) the causes of the NCC (NCCprs), 2) the consequences of the NCC (NCCcos), but there are also 3) the consequences of events causing the NCC (connected to the NCC by the common causal path), 4) the independent causes of the report and 5) the consequences of the independent causes of the report. Typically the consequences of the report are not among the NCC confounders, because the report is the last thing measured. All that is needed to prove that these possibilities are real is the simple fact that the correlation can result from one variable causing another or from some common causes. It is worth noticing that the confounders 4 and 5 are neural correlates of the report that by definition do not even have to be correlated with the NCC!

We do not claim to be original here, since with some care all the five possibilities can be found in the causal diagram in the Gamez's 2014 paper (all the five possibilities can be found there if the bidirectional arrows in that diagram are interpreted as common causes). What we do claim is that the possibility of there being such confounders follows naturally from our operational definition. This demonstrates that being armed with the operational definition cast in terms directly relevant to the practice of doing research helps in identifying new classes of confounders not recognized by some authors of the methodological papers concerned with the very problem of identifying the NCC confounders.

Our causal definition is more neutral with respect to the vertical relation than the one proposed by Chalmers. Obviously, as was pointed out by Gamez (2014), this is a desirable property. In fact, Gamez explicitly considers the causal paths between the stimuli and the content reports, but he provides the definition of the NCC according to which the conscious mental event is perfectly correlated with the NCC, «consciousness nomologically supervenes on the correlates of consciousness» and there is «a functional connection between consciousness and correlates of consciousness» (p. 5). This is all very interesting but also quite far from vertical neutrality.

The definition proposed in this paper does not say anything about the relation between the NCC as a neural event and the conscious experience as a mental event. In particular, in contrast to the Chalmers' definition, our definition does not say that the neural element is sufficient for the conscious experience. Saying that some neural event causes some conscious experience is to be interpreted as a shorthand for saying that some neural event causes some proper neural correlate of conscious experience, whatever being a vertically proper correlate is supposed to mean in the given context. All that our definition requires is that there were some neural causes of conscious content *in that sense* and that there was a unidirectional causal path such that this path corresponds to the stimulus causing the changes in the report by changing the contents of the experience. By virtue of what they do the researchers searching for the NCC are forced to accept it as a working definition but nothing seems to force them to accept the sufficiency condition, or the perfect correlation condition, or the functional connection condition, or any other vertical relation. That is why we believe that the definition proposed in this paper is purely operational while Chalmers' definition is not. We also think that the examples provided earlier (i.e., the inference rule for the effect of virtual lesions and stimulus manipulations or the exhaustive list of the NCC confounders) show that the purely causal definition proposed in this paper can be quite useful when interpreting the results of the studies or designing new ones. Since research should proceed without introducing unnecessary, untested and possibly confusing assumptions we believe that as far as the research practice is concerned our definition should be considered the recommended one.

REFERENCES

Aru, J., Bachmann, T., Singer, W., & Melloni, L. (2012). Distilling the Neural Correlates of Consciousness. *Neuroscience and Biobehavioral Reviews*, 36(2), 737–746.

Bachmann, T. (2009). Finding ERP-signatures of Target Awareness: Puzzle persists because of experimental co-variation of the objective and subjective variables. *Consciousness and cognition*, 18(3), 804–808.

Baars, B.J. (1989). *A Cognitive Theory of Consciousness*. New York: Cambridge University Press.

Block, N. (1995). On a Confusion about a Function of Consciousness. *Behavioral and Brain Sciences*, 18, 227–247.

Block, N. (2011). Perceptual Consciousness Overflows Cognitive Access. *Trends in Cognitive Sciences*, 15(12), 567–575.

Boly, M., Seth, A.K., Wilke, M., Ingmundson, P., Baars, B., Laureys, S., Edelman, D.B., & Tsuchiya, N. (2013). Consciousness in Humans and Non-Human Animals: Recent advances and future directions. *Frontiers in Psychology*, 4, 625.

Chalmers, D.J. (2000). What is a Neural Correlate of Consciousness. In T. Metzinger (Ed.), *Neural Correlates of Consciousness: Empirical and Conceptual Questions*. Cambridge, MA: MIT Press, 17–39.

Crick, F., & Koch, C. (1990). Towards a Neurobiological Theory of Consciousness. *Seminars in the Neuroscience*, 2, 263–275.

Dehaene, S., & Changeux, J.P. (2011). Experimental and Theoretical Approaches to Conscious Processing. *Neuron*, 70(2), 200–227.

Dehaene, S., Changeux, J.P., Naccache, L., Sackur, J., & Sergent, C. (2006). Conscious, Preconscious, and Subliminal Processing: A testable taxonomy. *Trends in Cognitive Sciences*, 10(5), 204–211.

Dehaene, S., & Naccache, L. (2001). Towards a Cognitive Neuroscience of Consciousness: Basic evidence and a workspace framework. *Cognition*, 79(1–2), 1–37

Einhäuser, W., Stout, J., Koch, C., & Carter, O. (2008). Pupil Dilation Reflects Perceptual Selection and Predicts Subsequent Stability in Perceptual Rivalry. *Proceedings of the National Academy of Sciences*, 105(5), 1704–1709.

Fleming, S.M., Maniscalco, B., Ko, Y., Amendi, N., Ro, T., & Lau, H. (2015). Action-Specific Disruption of Perceptual Confidence. *Psychological Science: A Journal of the American Psychological Society / APS*, 26(1), 89–98.

Gamez, D. (2014). The Measurement of Consciousness: A framework for the scientific study of consciousness. *Frontiers in Psychology*, 5, 714.

de Graaf, T.A., Hsieh, P.J., & Sack, A.T. (2012). The 'Correlates' in Neural Correlates of Consciousness. *Neuroscience & Biobehavioral Reviews*, 36(1), 191–197.

Hohwy, J., & Bayne, T. (2014). The Neural Correlates of Consciousness: Causes, confounds and constituents. In S.M. Miller (Ed.), *The Constitution of Phenomenal Consciousness: Toward a Science and Theory*. Amsterdam: John Benjamins, 155–176.

Kentridge, R.W., Heywood, C.A., & Weiskrantz, L. (1999). Attention without Awareness in Blindsight. *Proceedings of the Royal Society B. Biological Sciences*, 266(1430), 1805–1811.

Ko, Y., & Lau, H. (2012). A Detection Theoretic Explanation of Blindsight Suggests a Link between Conscious Perception and Metacognition. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 367(1594), 1401–1411.

Lamme, V.A. (2003). Why Visual Attention and Awareness are Different. *Trends in Cognitive Sciences*, 7(1), 12–18.

Lamme, V.A. (2010). How Neuroscience Will Change Our View on Consciousness. *Cognitive Neuroscience*, 1(3), 204–220.

Neisser, J. (2012). Neural Correlates of Consciousness Reconsidered. *Consciousness and cognition*, 21(2), 681–690.

Overgaard, M., Fehl, K., Mouridsen, K., Bergholt, B., & Cleeremans, A. (2008). Seeing without Seeing? Degraded conscious vision in a blindsight patient. *PloS One*, 3(8), e3028.

Rees, G., Kreiman, G., & Koch, C. (2002). Neural Correlates of Consciousness in Humans. *Nature Reviews. Neuroscience*, 3(4), 261–270.

Rounis, E., Maniscalco, B., Rothwell, J.C., Passingham, R.E., & Lau, H. (2010). Theta-Burst Transcranial Magnetic Stimulation to the Prefrontal Cortex Impairs Metacognitive Visual Awareness. *Cognitive Neuroscience*, 1(3), 165–175.

Timmermans, B., & Cleeremans, A. (2015). How Can We Measure Awareness? An overview of current methods. In M. Overgaard (Ed.), *Behavioural Methods in Consciousness Research*. Oxford: Oxford University Press, 21–46.

Wessel, J.R. (2012). From "Neural Correlates of Consciousness" to "Neural Causes of Consciousness": A commentary on "Consciousness, Biology and Quantum hypotheses", by B.J. Baars and D.E. Edelman. *Physics of Life Reviews*, 9(3), 299–300.