The Mismeasure of Consciousness

A problem of coordination for the Perceptual Awareness Scale

Matthias Michel, Université Paris-Sorbonne, SND[1]

Author's version. Accepted for publication in **Philosophy of Science** (forthcoming).

Abstract: As for most measurement procedures in the course of their development, measures of consciousness face the problem of coordination, i.e., the problem of knowing whether a measurement procedure actually measures what it is intended to measure. I focus on the case of the Perceptual Awareness Scale to illustrate how ignoring this problem leads to ambiguous interpretations of subjective reports in consciousness science. In turn, I show that empirical results based on this measurement procedure might be systematically misinterpreted.

[1] Address for correspondence: Sciences, Normes et Décisions. Université Paris-Sorbonne, 1 Rue Victor Cousin, 75005, Paris, FRANCE. Email: Matthias.Michel.2@paris-sorbonne.fr

Introduction

The main research program that aims at a beginning of an explanation of consciousness is based on a quest for the "neural correlates of consciousness" (NCCs): the minimally sufficient neural states that lead to a subject being conscious of a content (Chalmers, 2000; Crick and Koch, 1990)[2]. Researchers typically rely on a comparison between two types of measurements for finding NCCs: measures of neuronal processes and measures of consciousness (Michel, 2017; Overgaard, 2015). By comparing measurement outcomes, one can judge which neural processes best correlate with consciousness. Establishing robust measures of consciousness is thus critical for this research program. Within this framework, the Perceptual Awareness Scale (PAS), a four-step scale used by subjects to report the degree of clarity of their experiences, has been widely used to assess participants' levels of consciousness of contents (Ramsoy and Overgaard, 2004). However, the development of the PAS in consciousness science faces the "problem of coordination", namely, the problem of knowing whether a measurement procedure actually measures the quantity that it is intended to measure (Chang, 2004; Mach, 1896; van Fraassen, 2008).

Most critiques addressed against the PAS do so on the ground that it is an inaccurate measure of consciousness (Irvine, 2012b; Wierzchoń et al., 2014). In this paper, I argue that both proponents and critics of the PAS rely on the assumption that measures using the PAS are at least *valid* measures of consciousness, namely, that the PAS measures what it is intended to measure. On my view, the PAS fails to coordinate the quantity that it intends to measure, i.e., levels of consciousness, with the main quantity that it intends to measure consciousness with, i.e., reported levels of clarity of experiences. I conclude by showing how the use of the PAS could lead to systematic mistakes in the search for NCCs.

1 From the response criterion problem to the PAS

Subjective and objective measures of consciousness are often distinguished (Irvine, 2012a). A measure of consciousness is objective if the experimenter ascribes a conscious experience to a subject based on her performance on a task. On the other hand, a measure

[2] A content of perception is what is conveyed to the subject through perception (Siegel (2016)). For example, perceiving a red apple conveys the content *there is a red apple*. In what follows, I will assume that contents can be conscious as well as unconscious. An unconscious perception of a red apple conveys the content *there is a red apple*, just as a regular experience of a red apple does. In this article, I will often talk about contents being, for example, intense or not. This should be taken as a shorthand for saying that, if the perception that includes this content is accurate, the perceived object, say, a flash of light, would have high luminous intensity.

of consciousness is subjective if the participant is required to produce a report about her current conscious experience. In this article, I focus on the latter kind of measures.

Subjective measurement procedures face a “response criterion problem”. When instructed to report whenever a stimulus is consciously visible, subjects must come up with a criterion to determine whether a subjective experience qualifies for the report. According to signal detection theory (SDT), a subject’s decision to report seeing a stimulus depends both on the accumulation of sensory evidence and on the subject’s response criterion determining whether to report seeing the stimulus or not (Macmillan & Creelman, 2004). Importantly, categories such as “seen” or “unseen” are interpreted by subjects to set a response criterion depending on their own interpretations of these categories, such that stimuli that are only vaguely perceived might be reported either as “seen” or “unseen” (Merikle and Reingold, 1998). Accordingly, it is difficult for scientists to determine report instructions that make sense for participants without introducing biases in the way in which they decide to report their experiences or not. Researchers must therefore make sure that subjects will report all the relevant experiences, and only those experiences, so that they can reliably infer the subjects’ conscious states from their reports. By doing so, scientists engage in a process that is similar to the process of calibration of measurement procedures observed in the natural sciences.

Calibration is the operation of "establishing a reliable relation between the indications of an instrument and relevant features of an object" or phenomenon (Tal (2017)). In the cognitive neuroscience of consciousness, researchers aim at establishing a reliable relation between a subject’s report and the presence or absence of consciousness of a given content in that subject. For that purpose, calibration is typically accomplished by giving a set of instructions allowing researchers to make sure that subjects report all, and only, those features that are relevant for the experiment. To this extent, the response criterion problem is essentially a problem of calibration (Spener, 2015).

One of the main subjective measure of consciousness, called the Perceptual Awareness Scale (PAS), was designed to solve the response criterion problem (Ramsoy and Overgaard (2004)). As Sandberg and Overgaard (2015) explain:

> Whereas the scientist has no external access to the contents of another person’s consciousness, one can still aim for a situation where reports stand in a "1:1 relationship" with the relevant inner states. Thus, even if the scientist cannot confirm such a relationship, a participant can inform the scientist that he/she can tell the difference between different degrees of visibility, and use this information to create the

> experimental categories for report. This can only be done, it seems, by involving the participants in the creation of those categories. (p.181)

The PAS was developed by asking participants to design their own scale for rating properties of simple geometric shapes (shape, color and precision), starting from "no experience at all" and ending with "a clear image" (although they were free not to follow these suggestions). Almost all subjects developed a scale with four categories: "no experience at all", "brief glimpse", "almost clear image" and "clear image". The "brief glimpse" category was used by subjects to report no awareness of form, color, or shape, but only a vague experience of having seen something. Since subjects designed these categories themselves, one can thus suppose that the PAS avoids the confound of imposing response categories that subjects do not correctly understand, or categories that would lead participants to rule out reports of some kinds of experiences.

The PAS is not only used to *detect* when subjects are conscious of given contents and when they are not, but also to *measure* the *level of consciousness* of a given content. Indeed, it is assumed that subjects can be more or less conscious of some contents. Some contents are perceived "clearly and distinctly", and are thus highly conscious, while other contents are perceived unclearly and in a somehow degraded way, and are therefore less conscious (Overgaard et al. (2006)). As such, subjects introspect to produce indications, allowing experimenters to *detect* consciousness of given contents and to *measure* their levels of consciousness of those contents. In summary:

> A measurement outcome using the PAS is a statement that a subject has a given conscious experience, with a given level of consciousness, such that this statement is inferred from an indication, i.e., a subjective report resulting from the subject's introspection calibrated by PAS-instructions.

The crucial point with this definition of a PAS measurement outcome is that measurement outcomes are different from indications. In order to get from an indication to a measurement outcome, a scientist has to interpret this indication as indicating something about what one intends to measure, usually by using a theory (Tal, 2017). For example, in order to measure temperature, one must interpret a thermometer indication, such as the height of a column of mercury in a thermometer, in light of a theory that relates the volume of mercury with temperature. An indication, by itself, does not tell us anything about the phenomenon we are interested in (van Fraassen, 2008). Consequently, two steps must be distinguished in the production of a measurement outcome using the PAS. The first step is the use of introspection by a subject, given some instructions, to produce a report: this is the process that leads to a specific indication. The second step is the researcher's interpretation of that indication as telling her something about the subject's current experience and level of

consciousness. This is only through this second, interpretative step, that subjective reports become measurement outcomes.

Most debates focus on the first step of the production of measurement outcomes. Critics of the role of introspection in consciousness science argue that it is difficult to calibrate introspection such that it could provide reliable indications (Schwitzgebel, 2011; Irvine, 2012a, 2012b). This argument focuses on the *accuracy* of measurement procedures, defined as the closeness of agreement between measurement outcomes and the true value of the quantity that one intends to measure (*International Vocabulary of Metrology*). On my view, these critiques are unnecessarily concessive to proponents of the PAS. The most important problem for the PAS is not a problem of accuracy but a problem of *validity*, or, as philosophers of science sometimes call it, a problem of coordination (Barwich and Chang, 2015; van Fraassen, 2008). A measurement procedure is valid if it measures the quantity that one intends to measure, and not a different quantity. The problem of coordination thus touches on the second step of the measurement procedure: what entitles researchers to interpret reports as indicating something about conscious experiences or levels of consciousness? While the problem of accuracy presupposes that reports are indicating relevant aspects of consciousness, the problem of coordination questions the assumption that reports should be interpreted as indications of consciousness in the first place.

2 Consciousness science and the problem of coordination

2.1 The problem of coordination

Empirical measurement procedures typically require the coordination of two quantities. For example, the height of a column of mercury in a thermometer, a quantitative property, can play the role of an indication that is then interpreted in light of a theory of temperature to generate a measurement outcome. To the extent that one can use a quantity such as the height of a column of mercury to measure another quantity, temperature, both quantities are *coordinated*.

Coordination requires some recognized regularities, so that one can discover the form of the functions that links two quantities. The problem is that it is difficult to know whether variations in thermometer indications correlate with variations in temperature, and how so, before one is able to measure temperature (Chang, 2004; van Fraassen, 2008). The problem of coordination, in other words, is that good measurement procedures require good theories, but good theories require good measurement procedures.

A first step towards solving this problem is the discovery of systematic correlations between variations in indications produced by measurement procedures and variations in the property that one intends to measure. This first step is crucial, since it is only through those correlations that one is able to assess the validity of the measurement procedure, i.e., whether the procedure actually measures the property that one intends to measure (Cronbach and Meehl, 1955). If variations in thermometer indications were to correlate with variations in properties other than temperature, such as, say, the color of measured objects, or their distance from the Eiffel Tower, thermometers would not provide valid measures of temperature. By analogy, a procedure for measuring consciousness is valid only if variations in indications produced by that procedure correlate with variations in the presence or absence of consciousness, or with levels of consciousness, and do not correlate systematically with other properties that are independent from consciousness.

The various methods that aim at measuring and detecting consciousness all try to coordinate given quantities, such as confidence ratings (Norman and Price, 2015), post-decision wagers (Persaud et al., 2007), or visibility ratings (Sergent and Dehaene, 2004), with the presence or absence of consciousness, or with levels of consciousness. As noted above, the main quantity concept used in the PAS is the "level of clarity of an experience". From this perspective, the main question is whether subjects categorize and report levels of clarity of their experiences depending on how conscious of these experiences they are. If a wide variety of factors other than levels of consciousness could produce variations in reported levels of clarity, it becomes at least open to question whether the PAS is a valid measure of consciousness.

I now argue that there are multiple plausible interpretations for all subjective reports resulting from the use of the PAS. As such, we have no reason to believe that the PAS is a valid measure of consciousness.

2.2 The PAS and the phenomenological fallacy

There are a variety of factors other than levels of consciousness that might lead to variations in reported levels of clarity. For example, Irvine (2012b) argued that the reported level of clarity of an experience depends on the subject's level of confidence that a content was present, and not on her level of consciousness of that content. However, it seems that levels of confidence and levels of clarity should be distinguished, because measures of consciousness using the PAS do not produce the same results as measures using confidence reports (Sandberg et al., 2010; Zehetleitner and Rausch, 2013).

Rather, I suggest that reported levels of clarity of experiences do not vary depending on levels of consciousness of contents, but depending on levels of intensity and precision of the contents of consciousness. Fazekas and Overgaard (2016, 2017) argue that reported levels of clarity of experiences depend on two main characteristics of experiences: intensity and precision. The more intensely and precisely a content of consciousness is experienced, the higher the reported level of clarity of that experience. However, levels of intensity and precision of the contents of consciousness are not necessarily identical with levels of consciousness of those contents. Intensity and precision could be properties of the *contents* of consciousness, and not properties of consciousness itself.

To understand this distinction, it is useful to go back to what Place (1956) called the 'phenomenological fallacy': the error of concluding that properties of experienced contents are properties of experiences themselves. As Smart (1959) wrote against a similar mistake: "Trees and wallpapers can be green, but not the experience of seeing or imagining a tree or wallpaper" (p.151). The same thing goes for clarity, intensity or precision: the contents of consciousness are clear, vague, intense or precise. But consciousness of these contents itself is neither clear nor unclear, it is not intense nor precise. To attribute intensity and precision to consciousness itself is to commit the phenomenological fallacy. Rosenthal (2018) made this point clear:

> Though variation in intensity of the content of a perception and in strength of awareness of the perception both affect overall experiential intensity, only variation that is due to the second source is strictly speaking a gradation of consciousness. It is plain that perceptual content can vary altogether independently of consciousness. Variation in intensity of perceptual content is not gradation of consciousness. One would think otherwise only on a theory that assimilates perceptual consciousness to the content of the perception, and so fails to distinguish subjective awareness of psychological states from the states themselves. (p.7)

The phenomenological fallacy, and the distinction between properties of contents and properties of consciousness are of fundamental importance if one is hoping to solve the problem of coordination. Indeed, measuring *a representational property of the contents of consciousness* is not the same as measuring *levels of consciousness of those contents*.

To be clear, I do not want to deny that there could be levels of consciousness, that levels of consciousness could contribute to experienced levels of clarity of contents, or that variations in reported levels of clarity could correlate partly with variations in levels of consciousness. All that is needed for my argument is that variations in reported levels of clarity do not correlate solely with levels of consciousness, but also correlate with the way in

which contents are built as clear or unclear by perceptual systems, independently of the fact that these contents are conscious or not.

Since variations in reported levels of clarity could correlate with both variations in properties of the contents of consciousness and variations in levels of consciousness, we do not know whether reported levels of clarity should be interpreted as delivering information about the former or the latter. Hence, for every subjective report using the PAS, two interpretations are available: either the report should be interpreted as indicating a given level of consciousness, or the report should be interpreted as indicating a given degree of clarity, precision or intensity of a subject's content of consciousness.

2.3 Validity and the PAS: an inference to the best explanation

An alternative way to provide support for the validity of PAS measures could be to appeal to the fact that PAS measurement outcomes correlate with those of other measurement procedures intended to measure consciousness. This strategy has been successful before. For instance, a measurement procedure relying on an electron-microscope can be shown to be valid if its results correlate with those of a different procedure, such as a procedure using a light microscope, which one intends at measuring the same phenomenon (Hacking, 1981). Here, correlations between outcomes of different measurement procedures are best explained by the fact that both instruments at least partly measure the same phenomenon.

I already argued that there are disagreements between results using the PAS and results using confidence ratings or post-decision wagering (Sandberg et al., 2010). Nonetheless, results of PAS measures correlate with so-called "objective measures" of consciousness, such as levels of performance in discrimination tasks. When subjects perform at chance level in sensory discrimination tasks, which is interpreted as indicating that they did not see the stimulus, they tend to use the lowest level on the PAS, and when they perform above chance, they tend to use higher levels (Overgaard et al., 2006). However, this would indicate that the PAS is valid only if an agreement between several measurement procedures was a sufficient condition to confirm their validity (Borsboom et al., 2004). In order to make a case in favor of the validity of the PAS, one has to argue that *the best explanation* of the correlation between measurement outcomes using the PAS and other measures is that the PAS measures levels of consciousness.

On my view, the best explanation of the correlation between measurement outcomes of objective measures of consciousness and the PAS is not that the PAS is a valid measure of consciousness, but that contents that are built by perceptual systems as being more precise and intense elicit better performances. Indeed, it does not really come as a surprise

that when subjects are asked to discriminate between two stimuli with low contrast, low brightness and low intensity, their performance is worse than when they have to discriminate between two stimuli with high contrast, high brightness and high intensity. The way in which a content is built by perceptual systems as clear or unclear could lead to an increased or decreased performance *and* to a report using high or low level categories of the PAS. As long as we are not provided with a reason for thinking that performance is better on discrimination tasks *in virtue of the fact* that subjects are more conscious of stimuli, it is hard to see why a correlation between PAS ratings and performance should be interpreted as providing support for the validity of the PAS. Rather, the simplest explanation of this correlation is that contents that are clearer elicit better performances, independently of subjects' levels of consciousness of these contents. Hence, one can explain the correlation between performance on discrimination tasks and PAS ratings without appealing to levels of consciousness.

Next, I show why we should care about the validity of measurement procedures in consciousness science by arguing that the uncritical use of the PAS might have led consciousness science astray in its quest for the neural bases of consciousness.

3 Neural correlates of what?

In the search for NCCs, the minimally sufficient neural conditions for consciousness, scientists typically rely on a method called "contrastive analysis" (Baars, 1988). This method consists in comparing neural activity on trials in which a subject either consciously or unconsciously perceives a stimulus, for example, thanks to visual masking or binocular rivalry (Breitmeyer, 2015), while holding the stimulus constant. As such, it is thought that consciousness of a content can be manipulated as a variable, and one can thus record the neural states that best correlate with consciousness of contents, and do not correlate with those contents remaining unconscious. As noted by Overgaard (2006), one obtains NCCs by matching two measures: a measure of the neural activity of the subject and a detection or measure of the subject's consciousness of a content. My view is that the use of the PAS might have led consciousness researchers to mistake the correlates of the perceptual processes responsible for building contents as clear or unclear with the correlates of consciousness.

The search for the event-related potentials (ERPs), or the measured electrophysiological activity of the brain, that best correlate with consciousness of contents provides a good example of the way in which the problem of coordination applies in consciousness science.

There are currently two main candidates for the electrophysiological signature of consciousness: the N200, or Visual Awareness Negativity (VAN), a negative component which peaks at 200 milliseconds and is recorded in occipital and parietal areas (Koivisto & Revonsuo, 2010; Rutiku et al., 2016); and a late positive component, the P300 (or P3), which peaks at 300 milliseconds and is recorded in frontal areas (Del Cul et al., 2007; Sergent et al., 2005).

In this debate, most studies using the PAS to investigate the NCCs bring support for the hypothesis that consciousness correlates with the VAN (Andersen et al., 2015; Fu et al., 2017; Koivisto et al., 2010; Tagliabue 2016). The claim that the VAN correlates with consciousness is based on the observation that variations in the amplitude of the VAN correlate with variations of consciousness as measured by the PAS. For example, Andersen et al. (2015) found that variations in PAS ratings could be predicted from magnetoencephalogram (MEG) signals originating in the occipital lobe in the VAN time range, while MEG signals originating in the prefrontal areas in the P3 time range could not be used to predict variations in PAS ratings. This result has been interpreted as indicating that "occipital activity seems a more likely candidate for a neural correlate of perceptual consciousness than does prefrontal activation." (p.9).

However, this interpretation of the data assumes that variations in reported levels of clarity are best interpreted as indicating variations in subjects' levels of consciousness. On the other hand, I argued that the different levels on the PAS might not correspond to several levels of consciousness, but to several levels of clarity of the contents of consciousness. On this alternative interpretation of reported levels of clarity, the VAN does not correlate with consciousness itself. Rather, it is the best correlate of the clarity of the contents of consciousness. Hence, different interpretations of what the PAS actually measures lead to radically different interpretations of experimental results relying on the PAS. The view according to which the VAN correlates with a property of the contents of consciousness is perfectly compatible with the P3 component being the true NCC. Indeed, one could interpret the results as follows: first, a content is built as clear or unclear by perceptual systems, and this event correlates with the VAN; second, this content becomes conscious, and this event correlates with the P3 wave. Such an interpretation of the data is not available if one assumes that the PAS is a valid measure of levels of consciousness.

A neural correlate of the clarity of the contents of consciousness is not the same as a correlate of consciousness of those contents, and measuring the former does not amount to measuring the latter. Analyzing what we measure when using the PAS is of crucial importance in order to interpret data, as it determines what purported NCCs actually

correlate with. Without measurement procedures that succeed at measuring consciousness, and demonstrably so, one cannot hope to establish correlations between neural states and consciousness, let alone between neural states and levels of consciousness. That's why measurement validity is so important.

Conclusion

While the accuracy of measures of consciousness is debated, their validity is often left unexamined. Here, I tried to lay the conceptual groundwork for analyzing measurement outcomes provided by measures of consciousness, focusing on the PAS. I argued that, as long as the problem of coordination is not solved, subjective reports can be interpreted in various ways. Consequently, I showed that experimental results that rely on the PAS for their measures of consciousness remain ambiguous. The problem of coordination should not be overlooked: assessing whether purported measures of consciousness actually measure consciousness is a crucial challenge for future research.

References

Andersen, L. M., Pedersen, M. N., Sand-berg, K., and Overgaard, M. (2016). Occipital MEG Activity in the Early Time Range (<300 ms) Predicts Graded Changes in Perceptual Consciousness. Cerebral Cortex, 26(6):2677–2688.

Baars, B. (1988). *A Cognitive Theory of Consciousness*. Cambridge University Press.

Barwich, A.-S. and Chang, H. (2015). Sensory Measurements: Coordination and Standardization. *Biological Theory*, 10(3):200–211.

Borsboom, D., Mellenbergh, G. J., and Van Heerden, J. (2004). The concept of validity. *Psychological Review*, 111(4):1061–1071.

Breitmeyer, B. G. (2015). Psychophysical "blinding" methods reveal a functional hierarchy of unconscious visual processing. *Consciousness and Cognition*, 35:234–250.

Chalmers, D. (2000). What is a Neural Correlate of Consciousness ? In Metzinger, T., editor, *Neural Correlates of Consciousness: Empirical and Conceptual Issues*, pages 1– 33. MIT Press.

Chang, H. (2004). *Inventing Temperature: Measurement and Scientific Progress*. Oxford University Press.

Crick, F. and Koch, C. (1990). Towards a neurobiological theory of consciousness. *Seminars in the Neurosciences*, 2:263–275.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*(4), 281–302.

Del Cul, A., Baillet, S. & Dehaene, S. (2007) Brain dynamics underlying the non- linear threshold for access to consciousness, *PLoS Biology*, **5**(10), pp. 2408– 2423.

Fazekas, P. and Overgaard, M. (2016). Multi-dimensional Models of Degrees and Levels of Consciousness. *Trends in Cognitive Sciences*, 20(10):715–716.

Fazekas, P., & Overgaard, M. (2017). A Multi-Factor Account of Degrees of Awareness. *Cognitive Science*, 1–27.

Fu, Q., Liu, Y.-J., Dienes, Z., Wu, J., Chen, W., and Fu, X. (2017). Neural Correlates of Subjective Awareness for Natural Scene Categorization of Color Photographs and Line-Drawings. *Frontiers in Psychology*, 08(February):1–13.

Hacking, I. (1981). Do we see through a microscope? *Pacific Philosophical Quarterly*, 62:305–322.

Irvine, E. (2012a). *Consciousness as a Scientific Concept*. Springer.

Irvine, E. (2012b). Old problems with new measures in the science of consciousness. *British Journal for the Philosophy of Science*, 63(3):627–648.

JCGM (Joint Committee for Guides in Metrology). (2008). *International vocabulary of metrology* (3rd ed.) Sèvres: JCGM. http://www.bipm.org/en/publications/ guides/vim.html.

Koivisto, M. and Revonsuo, A. (2010). Event-related brain potential correlates of visual awareness. *Neuroscience and Biobehavioral Reviews*, 34(6):922–934.

Mach, E. (1896). Critique of the concept of temperature. In Ellis, B., editor, *Basic Concepts of Measurement*, pages 183–196. Cambridge University Press.

Macmillan NA, Creelman CD. (2004). *Detection Theory: A User's Guide*. New Jersey: Lawrence Erlbaum Associates Inc.

Merikle, P. M. and Reingold, E. M. (1998). On demonstrating unconscious perception. *Journal of Experimental Psychology*, 127(3):304–310.

Michel, M. (2017). Methodological artefacts in consciousness science. *Journal of Consciousness Studies*, 24 (11-12):94-117.

Norman, E. and Price, M. C. (2015). Measuring consciousness with confidence ratings. In Overgaard, M., *Behavioral methods in consciousness science*. Oxford University Press.

Overgaard, M. (2006). Introspection in Science. *Consciousness and Cognition*, *15*, 629–633.

Overgaard, M. (2015). The challenge of measuring consciousness. In Overgaard, M., *Behavioral methods in consciousness science*. Oxford University Press.

Overgaard, M., Rote, J., Mouridsen, K., and Ramsoy, T. Z. (2006). Is conscious perception gradual or dichotomous? A comparison of report methodologies during a visual task. *Consciousness and Cognition*, 15(4):700–708.

Persaud, N., McLeod, P., and Cowey, A. (2007). Post-decision wagering objectively measures awareness. *Nature neuroscience*, 10(2):257– 61.

Place, U. T. (1956). Is consciousness a brain process? *British Journal of Psychology*, 47:44-50.

Ramsoy, T. Z. and Overgaard, M. (2004). Introspection and subliminal perception. *Phenomenology and the Cognitive Sciences*, 3(1-23).

Rosenthal, D. (2018). Consciousness and confidence. *Neuropsychologia*, In Press.

Rutiku, R., Aru, J. & Bachmann, T. (2016) General markers of conscious visual perception and their timing, *Frontiers in Human Neuroscience*, 10(Feb), pp. 1–15.

Sandberg, K. and Overgaard, M. (2015). Using the Perceptual Awareness Scale. In *Behavioral methods in consciousness science.* Oxford University Press.

Sandberg, K., Timmermans, B., Overgaard, M., and Cleeremans, A. (2010). Measuring consciousness: Is one measure better than the other? *Consciousness and Cognition*, 19(4):1069–1078.

Sergent, C. and Dehaene, S. (2004). Is consciousness a gradual phenomenon? Evidence for an all-or-none bifurcation during the attentional blink. *Psychological Science*, 15(11):720–728.

Sergent, C., Baillet, S. & Dehaene, S. (2005) Timing of the brain events underlying access to consciousness during the attentional blink, *Nature Neuroscience*, 8 (10), pp. 1391–1400.

Schwitzgebel, E. (2011) *Perplexities of Consciousness.* Cambridge, MA:MIT Press.

Siegel, S. (2016). The contents of perception. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy.* Stanford University, winter 2016 edition.

Smart, J. J. C. (1959). Sensations and brain processes. *The Philosophical Review*, 68(2): 141-156.

Spener, M. (2015). Calibrating introspection. *Philosophical Issues*, 25(1):300–321.

Tagliabue, C. F., Mazzi, C., Bagattini, C., and Savazzi, S. (2016). Early local activity in temporal areas reflects graded content of visual perception. *Frontiers in Psychology*, 7(APR):1–10.

Tal, E. (2017). Calibration: Modelling the measurement process. *Studies in History and Philosophy of Science Part A*, pages 1–13.

van Fraassen, B. (2008). *Scientific Representations: Paradoxes of Perspective*. Oxford University Press.

Wierzchoń, M., Szczepanowski, R., Anzulewicz, A., and Cleeremans, A. (2014). When a (precise) awareness measure became a (sketchy) introspective report. *Consciousness and Cognition*, 26(1):1–2

Zehetleitner, M. and Rausch, M. (2013). Being confident without seeing: What subjective measures of visual consciousness are about. *Attention, Perception, and Psychophysics* 75(7):1406–1426.