

The Perceptual Awareness Scale—recent controversies and debates

Morten Overgaard^{1,2,*†} and Kristian Sandberg²

¹Aarhus Institute of Advanced Studies, Aarhus University, Nørrebrogade 1A, Aarhus 8000, Denmark; ²Department of Clinical Medicine, Center for Functionally Integrative Neuroscience, Aarhus University, Universitetsbyen 3, Building 1710, Aarhus 8000, Denmark

†Morten Overgaard, <http://orcid.org/0000-0002-1215-5355>

*Correspondence address. Aarhus Institute of Advanced Studies, Aarhus Universitet, Nørrebrogade 1A, Aarhus 8000, Denmark.

Tel: +45 20783154; E-mail: morten.storm.overgaard@cfin.au.dk

Abstract

Accurate insight into subjective experience is crucial for the science of consciousness. The Perceptual Awareness Scale (PAS) was created in 2004 as a method for obtaining precise introspective reports for participants in research projects, and since then, the scale has become increasingly popular. This does not mean, of course, that no critiques have been voiced. Here, we briefly recapitulate our main thoughts on the intended PAS usage and the findings of the first decade, and we update this with the latest empirical and theoretical developments. We focus specifically on findings with relevance to whether consciousness is gradual or all-or-none phenomenon, to what should be considered conscious/unconscious, and to whether PAS is preferable to alternative measures of awareness. We respond in detail to some recent, selected articles.

Keywords: consciousness; methodology; subjective report; the Perceptual Awareness Scale; PAS; perception

In the late 1990s, Thomas Zoega Ramsøy and Morten Overgaard invented the Perceptual Awareness Scale (PAS) and published the first scientific paper explaining the method in 2004 (Ramsøy and Overgaard 2004). At the time, we did not imagine that the idea would turn into one of the most widely used measures of consciousness over the following decade. Several recent publications have critically debated PAS, and below, we will address the raised issues and discuss some of the articles in detail. First, we summarize how PAS was created and how we recommend that it is used as a general method.

PAS was constructed based on evidence from collaborating participants as an attempt to tackle what we found was one of the most challenging methodological problems related to consciousness research—the lacking possibility to externally calibrate subjective content. We thought that even though the scientist has no external access to the contents of another person's consciousness, she 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 externally confirm such a relationship, a participant can inform the scientist that she can tell the difference between different degrees of visibility and use this information to create the experimental categories for report. We have proposed that this is best achieved by involving the participants in the creation of those categories. The alternative would be that the scientist came up with the number of scale points and labelling of the categories—which would require external access.

There is no external way of validating—if a participant reports a particular experience—whether that report is given because the participant had that exact experience, or rather because, for instance, the participant was biased to give this report. Such arguments have paved the way for a broad conviction in cognitive science that one should avoid subjective reports as much as possible and rather rely on objective measures of correctness or reaction time only. In consciousness research, views on subjective reports are rarely this pessimistic. It has however been argued that all subjective measures of consciousness are conflated by the reporting itself (e.g. Irvine 2012) and that no-report paradigms can be used as alternative—i.e. paradigms where we measure some objective behaviour instead of using a report (Tsuchiya et al. 2015).

There are, however, other problems with objective measures that may be considered even more difficult. Essentially, if one wishes to study subjective experience, it is not at all clear which objective measure to use. How can we know, for instance, that any measure such as correct identification or any other measure of performance is actually about the subjective experience of interest—and more so than the subjective report? It seems the only knowledge we could have comes from a prior correlation with introspective observation and report and, accordingly, cannot have any higher precision than the introspective observation/report. In other words, in order to associate a particular behaviour or cognitive function with consciousness, one seems forced to base this decision on information from introspective

Received: 4 July 2021; Revised: 10 November 2021; Accepted: 25 November 2021

© The Author(s) 2021. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License

(<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

reports—e.g. the idea that working memory is closer related to consciousness than, say, activity in the appendix derives from introspection only. That said, any argument for why subjective reports seem a *sine qua non* for consciousness research is not an argument for any subjective reporting being precise or trustworthy. Regardless of one's position on whether consciousness research needs to sometimes replace subjective methods with objective methods, the need to refine subjective methods seems an important endeavour.

In our first study, participants were asked to identify the shape, colour, and location of briefly presented and masked geometric figures (thus ratings different stimulus aspects separately), and the purpose of the study was for the participants to find out how many and which kinds of report categories were necessary and sufficient to describe experienced differences in perceptual clarity. During the course of the experiment, each participant created an awareness scale, and the participants were subsequently to report the clarity of each stimulus property (shape, colour, and position) using this scale. 'Clarity' is itself a complicated concept and is here used to denote the fact that we sometimes see objects clearly, sometimes not at all, and possibly, we sometimes see objects less clear—in between the other two extremes. In the study, the participants were told that they could give any kind of labelling or description of the scale points they preferred and that it would be fine if they ended with a two-point scale, a 100-point scale, or anything in between. End points were, however, proposed to be 'nothing at all' and 'completely clear' in order to underline our understanding of 'clarity', yet participants were encouraged to also define the end points themselves. Participants were suggested to think aloud or discuss their thoughts with the experimenters, even though they were not given any feedback or suggestions in order to make sure that results were not confounded.

In this initial experiment, participants made many changes to their preferred definitions and number of scale points. Some participants started with six points, yet found out that they did not themselves understand the definitional borderlines between all scale points. This made them re-evaluate both numbers and definitions until they in the end all had scales, they knew how to use. In the end, all our participants ended up using a 4-point scale as described below. Once the scale was established, the participants thus used this scale inspired by their own wording. Most subsequent experiments that have used PAS have applied the 4-point scale with the original wording without re-doing the calibration phase. As discussed elsewhere (Sandberg and Overgaard 2015), we consider this to be in full accordance with the logic of PAS as long as there are not too many contextual (e.g. task-related) differences. It may be noted that if the scale calibration procedure was to be applied again in a large sample of participants, or with a different stimulus material, some variation in the number of scale steps would not be surprising. We treat this aspect in greater detail elsewhere (Sandberg and Overgaard 2015).

The potentially most crucial aspect of PAS is the meaning of the individual scale points, as this information helps the scientist to understand the subjective state of the participant and makes it possible—at least to some degree—to compare reports. For this reason, we decided to emphasize the potential importance of thorough and flexible instructions rather than rigid, standardized instructions. Although the idea of very 'trained participants' can be criticized as well, the aim of the instructions should be to ensure that participants interpret the meaning of the scale points in the same way. Below, we list the meaning of the PAS scale points in accordance with the original version (Ramsøy and Overgaard 2004):

(i) 'No experience' (NS): No subjective experience of the stimulus, not even the 'faintest sensation' that anything was presented at all. Not even a feeling that something might have been presented.

(ii) 'Brief glimpse' (BG): A variation in subjective experience that is 'stimulus related'. One does not have any clue at all what the stimulus was (e.g. a geometric shape, a natural scene, or a red dot), just an experience of 'something being there'.

(iii) Almost clear experience (ACI): A somewhat blurry and not very clear experience of a stimulus, however with some idea about its nature. One is typically less confident about the stimulus than if one has had a clear experience.

(iv) Clear experience (CE): An experience of seeing the entire stimulus without problems.

The distinction between 'NS' and 'BG' is typically the one that most participants confuse in the beginning of experiments. It seems most people are used to labelling perceptions as 'unconscious' if they have no idea about what they saw, even if they had a feeling of seeing 'something'. The distinction is very important as previous PAS experiments specifically point out that the crucial difference is between those two categories with regard to 'subliminal perception': at 'BG', participants are typically well above chance, whereas this is rarely the case at 'NS' (Sandberg and Overgaard 2015).

We suggest starting every experiment using PAS with a thorough instruction phase, explaining all scale points. Experience so far indicates that it is important to combine the instruction about how to use PAS with an open discussion with the participants about how they understand the individual categories. It is rarely enough to just ask the participant whether she understood the instructions—it is more effective to ask her to repeat the definitions, possibly with her own words.

After the initial instruction, we suggest spending time on pilot trials. It is well-known that participants as a result of getting tired or bored perform worse on objective measures (e.g. correctness or reaction time) over several trials, but, at the same time, they learn still more effective strategies to complete tasks. In this way, the intense learning typical of early trials in experiments continue, although typically less intense, throughout the experiment. How such observations apply to subjective measures within an experimental session is currently unknown although Schwiedrzik et al. (2011) observed that subliminal perception at the subjective threshold was present in their first experimental session but disappeared in subsequent sessions. It should be noted, however, that it is difficult to judge whether this is a result of experiences (and the task accuracy for specific experiences) actually changing or a result of participants learning to relate their experience to objective accuracy better: One cannot easily show a learning curve when the scientist, in the absence of direct access to the participant's experiences, cannot evaluate the correctness of the report. One could, however, examine for example, the development in mean awareness ratings over trials and compare this to the development of mean accuracy, or one could plot full psychometric functions for accuracy and awareness. Both these methods would allow a researcher to identify the periods of greatest change in their experimental paradigm (initial learning and eventual fatigue) so that these might be avoided.

In our own previous studies, we have used 40–50 pilot trials and observed that a large part of the change in the accuracy–awareness relationship took place across these trials, and we therefore suggest using at least 40 unless specifically examining the learning effect. More training might be optimal, but practical

issues typically prevent testing participants across two or more days and only using the data for the last day.

An effective practical method to ensure fast learning of the appropriate use of PAS categories is to interrupt participants during the pilot trials in order to ask them why they chose a particular PAS rating, and/or to recall the definition of the rating, they just used. In pilot experiments, we have seen how the use of PAS as a 4-point scale with labels, but without thorough descriptions as shown above, gives markedly different results. We have found that the difference typically relates to the NS/BG difference, as described above, so that results indicate more subliminal perception before than after the correct instruction. For this reason, an experiment that makes use of the PAS categories but without proper instructions, and thus effectively works as any other '4-point scale', may behave differently than a PAS with in-depth instructions. In other words, PAS is not defined as 'a 4-point scale of awareness', but as the methodological approach of allowing participants to label their experienced 'levels' of subjective clarity. This is one of the most frequent misunderstandings of PAS. It was never the intention that the scale should be simply presented to participants on a screen with the labels—and no further instructions (Sandberg et al. 2013; Szczepanowski et al. 2013). In one study, feedback from participants during the training session in a pilot study even changed the PAS used in the main experiment in line with the original PAS framework. Specifically, Christensen et al. (2006) presented participants with simple geometric shapes for durations between 33 and 100 ms and asked them to report on the four-step PAS. Pilot participants reported difficulties distinguishing weak glimpses from almost clear experiences and indeed used these two categories on average only on one in six trials each, whereas the other two PAS categories were used on one in three trials each. Consequently, the two middle categories were collapsed to a joint 'vague percept' category in the main experiment.

PAS has also been used in paradigms with somewhat different visual stimuli than in the original study. For instance, it has been used to rate the awareness of a number or its colour (Windey et al. 2013). It has also been used to rate the experience of a stimulus as a whole even when task accuracy may be influenced by the perception of several subcomponents of the target. In one study, participants were asked to discriminate between fearful and neutral faces and rate their experience of the face using PAS (or rate their confidence in being correct or place a wager on being correct) (Szczepanowski et al. 2013).

Although the original version of PAS was designed to study clarity of visual consciousness, we have used the same approach to create a scale for auditory consciousness that we directly compare with the 'visual scale' (Overgaard et al. 2013). In another experiment, we created a 'sense of control scale' using the very same approach (Dong et al. 2015). The latter example is rather far from the original version of PAS, as the reports here refer to the sense of control over an action rather than clarity of perception. This is to illustrate that PAS—in spite of the name 'Perceptual Awareness Scale'—should fundamentally be seen as an approach that can be applied to any aspect of consciousness that potentially comes in degrees.

As all other methods, PAS is based on certain assumptions. One may agree or disagree with those assumptions, but one cannot use a method to investigate the validity of its own assumptions. In the case of PAS, it is assumed that we have a privileged access to our own experiences. From this perspective, there is no 'invisible consciousness', and knowledge about consciousness is derived from our own 'introspective' access to it. This assumption may obviously be false, although, to us at least, it is difficult to see how one

can say anything about consciousness, including how to study it objectively, if one believes not to have any access to its character or content.

PAS, consciousness and unconsciousness

By far, PAS has primarily been used to investigate to which degree conscious and unconscious processes contribute to a particular performance. The original PAS study revealed that the amount of measured unconscious influence on a discrimination task depended on the subjective rating scale (Ramsøy and Overgaard 2004). At 'No experience' rating, participants were at base chance, whereas when participants performed the same task using a dichotomous scale, we found massive unconscious influence. This result—that the measured unconscious influence is much lower and typically non-existing—has been replicated many times (Overgaard et al. 2006; Sandberg et al. 2010; Timmermans et al. 2010). The effect has been found in many different paradigms and settings that typically have been used to argue in favour of the existence of unconscious processes. In a blindsight patient, we found that the blindsight phenomenon relied on vague perception rather than unconscious perception (Overgaard et al. 2008; Overgaard 2011; Overgaard and Grünbaum 2011; Overgaard and Mogensen 2015), which was replicated in another blindsight patient by another research group (Mazzi et al. 2016). Various experimental paradigms claimed to be objective approaches to consciousness and that have been used to argue a massive amount of unconscious influence on behaviour have been revealed to indicate the exact opposite using PAS. As one example, exclusion tasks seem to require weak glimpses of the stimulus (Sandberg et al. 2014). In another study, we showed that emotional priming only works when there is some degree of experience (weak glimpses) of the prime (Lohse and Overgaard 2019). Other experiments using PAS found that auditory affective processing requires consciousness (Overgaard et al. 2013; Lähteenmäki et al. 2019).

Despite the number of experiments pointing in the same direction, the validity of PAS has been challenged. A somewhat surprising take on this is put forward by Michel (2019). Like Overgaard (2006), he argues that no method can measure consciousness directly, but we must instead ensure that it is as well related to consciousness as possible—that it is a valid measure. He refers to this process as 'calibration', and his main claim is that PAS is in fact not entirely (or maybe not at all) related to so-called levels of consciousness, but instead measures the quality of perceptual contents, which he considers as different. He extends his argument into the subtopic of the search for the neural correlates of consciousness and puts forward that the problem of PAS's validity could have led to wrong conclusions in this field. There are quite a few things to unpack and respond to in his criticism.

As has been pointed out by others (Skóra et al. 2021), Michel's preferred term 'levels of consciousness' is undefined by him, and despite how central it is to his claim, it is not entirely clear what it covers. The term appears to be adopted from an older article of ours (Overgaard et al. 2006), in which we remain relatively agnostic about how to characterize consciousness in general and use 'levels' interchangeably when discussing different categories/types/grades of experience and even the PAS scale steps themselves. We have since then moved away from this broad usage of the term and generally attempt to specify which of the aspects we refer to. Another clue to Michel's interpretation of 'levels of consciousness' comes from his reference to Rosenthal (2019), who distinguishes between the intensity of a representation and the strength of the awareness of that representation.

One could thus be very aware of an unclear experience—for example, when looking at someone from a long distance. In such a case, one might identify that it is indeed a person but lack the perceptual details to judge who it is. As we understand Michel, he claims that PAS measures primarily the intensity of representation, but that we should actually be interested in the awareness as a separate phenomenon.

To accept this criticism at the general level, it seems necessary to accept some version of Higher-Order Theory (HOT), which Michel appears to ascribe to (at least to some extent, as was also pointed out by Skóra et al. (2021)). While there is some intuitive appeal in the statements, we believe that once unpacked, however, they merely reflect a poorer choice of what the field of consciousness should explain. We do not reject the distinction between representational strength in general and awareness in general but view it somewhat differently. For example, independently of whether one ascribes to an all-or-none or gradual/graded view of consciousness, one typically accepts some kind of threshold of awareness (the distinction between ‘no experience’ and something else), and below that point, representational strength could vary. In this sense, the degree of representational strength can be different in two ‘unaware’ cases, and representational strength would thus not always be the same as the degree of awareness.

If we dive deeper into the imagined case of a strong experience of something perceptually weak, some differences in our views arise. From a HOT perspective, this case reflects a strong higher-order representation of a weak perceptual first-order representation. If instead, one views an experience as composed of or built from its perceptual qualities, it can be described differently. In the case of viewing someone from a distance, for example, we could imagine that the defining features of the person’s face are now at spatial frequencies outside the spectrum that can be processed even in the fovea, and the face area appears blurry or unclear, and there is little or no representation of facial features. Yet, other aspects of the representation are clear, such as the spatial position, the overall shape, perhaps the colours of the clothes and even the blurred face is available over time making it different from a glimpse-like experience. In such a view, awareness can thus overall be described as ‘strong’ because it is rich in temporal and spatial stability, but at the same time, it can be described as being ‘of something perceptually weak’ because the key telling feature as not represented. This description is not only highly compatible with the multi-factor account of degrees of awareness, we have presented earlier (Fazekas and Overgaard 2018), but appears to also be compatible with Integrated Information Theory where the experience would be reflected in the cortical areas that are part of the dynamic core (Tononi et al. 2016), and at least the Partial Awareness Hypothesis version of Global Workspace Theory where partial experiences like these are discussed explicitly (Kouider et al. 2010). One may further note that any and all of the aspects of the conscious perception could be probed by PAS or other awareness scales.

While Michel’s criticism is specifically directed at PAS, it applies in fact to a greater extent to most of the current alternatives to PAS. For example, confidence and post-decision wagering are built on the same assumption that insight into the strength of the sensory (or generally first-order) representation is a relevant aspect (perhaps the most important one), but both these types of scales are expected to integrate other aspects (e.g. knowledge about the task or cues that affect decision accuracy without affecting awareness) into the rating to a greater extent than PAS does. More indirect measures such as exclusion tasks have even

more assumptions while the intended outcome is whether a task was performed consciously (overtly) or not, which depends on the perceptual or otherwise informative features of the experience.

With this in mind, it is valuable to take a further look at what the field of consciousness has attempted to identify historically and how well that aligns with Michel’s view as compared to what PAS measures. From the Global Workspace Theory perspective, Dehaene and Naccache (2001) argue that ‘Subjective reports are the key phenomena that a cognitive neuroscience of consciousness purport to study’, and Sergent et al. (2005) used graded visibility reports when presenting evidence for the theory. Other researchers have argued that the Visual Awareness Negativity (VAN)—an Electroencephalography (EEG) difference wave frequently observed in consciousness studies—is the most consistent correlate of consciousness (Koivisto and Revonsuo 2010; Förster et al. 2020). Proponents of this view also accept and use graded awareness reports (e.g. Koivisto and Grassini 2016), and so have even HOT theorists. For example, Lau and Rosenthal (2011) emphasize a study from Lau’s research team distinguishing between unclear and clear experiences (Rounis et al. 2010) as evidence in support for HOT. Overall, it would seem that researchers from a wide range of theoretical backgrounds have accepted graded awareness ratings as valid for measuring awareness. Accepting Michel’s criticism would thus appear to mean that the field must abandon not just evidence from PAS studies but huge bulks of studies. Given the large number of studies in recent years using awareness/confidence scales and supporting VAN/posterior cortex views (for reviews, see Koch et al. (2016) and Förster et al. (2020)), it is not hard to predict who will be tempted to accept the criticism and who will say that the other camp is suddenly moving the goal post.

Michel continues to draw consequences of his criticism to the NCC debate and points out that the P3 wave could be the true correlate of consciousness, whereas VAN might only be a correlate of perceptual clarity. To us, this does not make sense as—for example—in Sergent et al. (2005), the P3a/b are absent for awareness ratings of 10 or lower on a 21-point scale. If there is no awareness in half of the experimental trials, why are participants then still reporting different clarities of experience overtly and how could these reports meaningfully index perceptual clarity and correlate with distinct neural activity (such as the N2 component in the same experiment)? If there is no access to the perceptual state as indexed by the absent P3a/b, how can the states be reported as distinct in consciousness?

In summary, with his criticism, we believe that Michel has presented a case where—in order to avoid a minor, widely accepted premise—we have to accept an even bigger premise (the framework of HOT), accept counterintuitive instances of unconscious reports of weak experiences, and reject huge bulks of the consciousness literature. We would further have to accept that we do not have an even reasonably valid measure of consciousness, and Michel provides little or no cue as to what such a measure might look like. It would probably not be a very intuitive measure for participants to use as they would have to understand not to report the clarity of their experience, but instead the strength of their awareness independently of whether what they see is clear or not, but at the same time be careful not to confuse this with their introspective process which is something different as well.

Persuh (2018) has argued that reports about consciousness are behaviour in the same right as all other actions, and thus, they are not anymore ‘subjective’ than what is considered ‘objective’ measures in experiments. She concludes that subjective reports are vulnerable to bias and that they consequently are illusory.

The argument is thus that subjective reports say nothing over and above how participants perform in a task. The position is essentially the opposite of the fundamental claim behind PAS that we have privileged access to our own experiences and that subjective experiences are real. If such experiences are real, if we can have access to them and talk about them, then experiences can be the content of a report. The fundamental claims underlying PAS are not that the content of consciousness is without influence from contextual factors—on the contrary. The claim is rather that a report of consciousness is about nothing over and above the content of that experience—regardless of its causal history. Any experience, according to this view, is influenced by numerous biological and contextual/social factors, some of which could be called bias, yet as long as the report accurately represents what is experienced, then that report is a true representation about an aspect of the world, i.e. that content itself. For Persuh to argue that objective and subjective measures should be evaluated in the same way, she must argue that they are about the same ‘type’ of object. In contrast, the logic behind PAS would be that bias may influence and challenge the validity of an objective measure. A metacognitive report is vulnerable to bias only under the condition that some contextual factor makes a person lie or falsely report what she experiences, which exactly is an argument to propose more elaborate methods for report.

The discussion could be said to echo classical debates in the history of psychology. Particularly the influential contribution of Nisbett and Wilson (1977) has sometimes been taken as a still undisputed argument that introspective reports have no place in scientific research. They present evidence that people have little introspective knowledge about the reasons behind their own opinions—for example, that people systematically preferred objects presented to the right of them, while reporting other reasons for this choice. It could be argued, however, that this conclusion misconstrues the evidence of Nisbett and Wilson. Subjects giving an introspective report about liking objects presented to the right for some other reason than the object’s location in space may be giving a perfectly good and scientifically usable report of what they experienced. Nisbett and Wilson correctly rejected introspection as a methodology to learn about (some aspects of) choice and decision-making, as the behavioural data suggested a very different explanation from the one that subjects themselves reported. Another interpretation of the results could be, however, that in some unknown (but probably vast) number of situations, people do not have introspective access to their own cognitive processes. However, not surprisingly, they still have some experience and interpretation of their own actions. Thus, a conflict in data between subjective report and behaviour could be interpreted to show that the subject’s experience differs from what can be analysed from his or her behaviour, and, thus, it does not automatically follow that the introspective report is invalid (Overgaard 2006).

In one recent experiment, we used a false feedback paradigm to investigate whether confidence ratings (reports on the performance of a task) and PAS relate to different processes (Skewes et al. 2021). Participants were asked to perform a standard psychophysical detection task and report using either PAS or confidence ratings (both presented as comparable four-point scales). We used feedback to selectively intervene either on PAS or confidence ratings and measured the effects of these interventions on response accuracy, on reports of perceptual awareness, and on response confidence. We found that false feedback based on PAS responses reliably reduced not only the PAS responses themselves but also their accuracy on the task. False feedback based

on confidence ratings did not reduce objective performance. The results suggest that different processes underlie different types of metacognitive reports (as previously predicted by Mogensen 2017 and discussed by Overgaard and Sandberg 2012). In other words, if confidence ratings, subjective reports, and other metacognitive measures can be separated conceptually and empirically, there seems to be strong evidence that there are in fact ontologically different ‘types’ of metacognitive access to the content of subjective experience.

We believe that the findings and arguments above suggest that consciousness research should not just discuss ‘for and against’ subjective reporting, but also differences and similarities between types of report, types of access, and how such varieties of access relate to conscious content and behaviour.

A further complexity of this debate is its conceptual side. In the major part of consciousness research, it is assumed that a cognitive state is unconscious if there are no traces of phenomenal experience related to it at all. Another option, inspired by the PAS steps, would be that a cognitive state can be considered unconscious in spite of experience under the condition that participants cannot specify the exact content of that experience. In most experiments using PAS, participants are unable to correctly report a stimulus when reporting not to see it at all—yet they are often above chance level when reporting vague experiences. Such a discussion would not need to question the validity of awareness measures nor the concept of consciousness at all, but rather discuss how our concepts of consciousness determine in which cases experimental results may demonstrate unconscious perception.

PAS and gradual consciousness

Does PAS reveal that consciousness is gradual? Several publications argue that this is the case because participants naturally categorize their experiences as gradual and because objective behaviour correlates with gradual differences in consciousness (Aru and Bachmann 2017). Furthermore, neural correlates of visual experience have been found to correlate with gradual differences in consciousness just like behavioural measures of correctness using a variety of methods, e.g. functional Magnetic Resonance Imagery (fMRI) (Christensen et al. 2006; Binder et al. 2017), Evoked Response Potential (ERP) (Derda et al. 2019), and Magnetoencephalography (MEG) (Andersen et al. 2016).

All experiments using PAS have found that participants use the scale in a gradual manner, and that the scale generally correlates with objective measures of behaviour and neural activity. The question is of course whether reports of gradualness prove metaphysical gradualness. Or, in other words, whether there is something more to consciousness than what we have phenomenal access to. According to the most fundamental assumption underlying PAS, people understand what they experience better than anyone else can. We would personally add to this that consciousness has no other definition than what is experienced—because what would justify the need of that? But this is a conceptual question bigger than the methodological point of PAS.

Bayne et al. (2016) argue that reports of vague perceptions can be accounted for without supposing that consciousness is gradual. Taking a perspective of Predictive Coding theory, they argue that when subjects report gradual consciousness, they are in fact only talking about uncertainty and not actual perceptual gradualness. They believe that the descriptions of the PAS ‘levels’ support this view (mentioning ‘almost clear experience’ as an example, which is the only ‘level’ actually referring to uncertainty). Obviously, this

interpretation is difficult to rule out principally—yet equally difficult to defend. However, in our work with PAS through many years, we have never been under the impression that participants are in fact talking about confidence. Three of four PAS ‘levels’ do not refer to certainty, and participants can be highly confident saying that something was vaguely experienced. The most direct evidence to our knowledge is the experiment presented above by [Skewes et al. \(2021\)](#) showing that a manipulation of PAS gives different results than the same manipulations of confidence ratings. As discussed below, there is some further evidence that gradual reports of confidence yield different results than PAS ([Sandberg et al. 2010, 2011](#); [Rausch et al. 2015](#); [Rausch and Zehetleitner 2016](#)). These findings, we believe, make it difficult to defend that PAS is only about confidence. To our knowledge, no experiment directly supports their claim.

Bayne, Hohwy, and Owen present another interpretation of PAS that, for a different reason, suggests PAS does not demonstrate gradual consciousness. They argue that participants misunderstand the PAS categories so that they in fact do not report about gradualness at all. Again, whereas this cannot be ruled out, it seems very speculative and without any empirical support. On the contrary, we have the clear impression that participants describe their experiences in much detail and with much help and instruction. One recent experiment found evidence that visual perception is gradual in a holistic rather than fragmented way (i.e. the entire percept is reduced)—a level of detail and consistency in results that to us does not suggest that participants all misunderstand what they are talking about ([Del Pin et al. 2020](#)).

This experiment introduced a manipulation aiming to disentangle two prevalent positions: so-called ‘rich views’ that posit that people virtually represent the external world with unlimited capacity. Alternatively, ‘sparse views’ state that representations are reconstructed from expectations and information.

Eight objects were each presented in a box arranged in a circular array. After the offset of the objects, the boxes remained on the screen, and the frame of one of them turned red with a line pointing towards it. The participants then had to select between two images or words, one corresponding to the target object and one corresponding to an object not presented in the array. Subsequently, participants reported their awareness of the object using the PAS. Some theoretical positions would expect that participants are equally accurate, no matter whether they are probed with the same image of the object again or a word naming the object as all aspects are represented. Other theories would predict better performance for image probes than word probes as a word is only one of many potential representational levels, whereas an image contains them all. The results supported the first view. If vague perceptions were vague in the sense that they are fragments, these results would be difficult to explain: How would fragments of a visual figure relate as much to a word representing the image as a more complete version of that image? That a word and image probe led to similar results seem more in line with the view that vague perceptions are holistically degraded.

A very recent article by [Kim and Chong \(2021\)](#) has approached this topic from an interesting angle, and we here dedicate some space to commenting on their study. They specifically take a starting point in the Partial Awareness Hypothesis (PAH) ([Kouider et al. 2010](#)), which states that awareness is all-or-none, but at different stages of sensory processing. In this way, awareness can be graded overall if, for example, the spatial position is clear, but the object identity is absent in awareness. They write that ‘the hypothesis has a limitation in that it cannot explain partial awareness

experienced within each level, especially graded visual experience in low-level stimuli’, and they continue to propose separate processing of low and high spatial frequencies (LSF/HSF) as crucial to explaining such observations. In other words, the distinction between LSFs and HSFs might be key to understanding graded awareness in low-level stimuli, which would otherwise pose a challenge to the PAH.

They contrasted perception of LSF and HSP, arguing that graded experiences could be a consequence of all-or-none perception of different frequencies. The authors point out that LSF and HSF are likely processed separately and in parallel (and are thus not an example of levels of processing), with LSF providing course/global information (overall shapes and gist of scenes) and HSF providing fine/local information (e.g. specific stimulus features or details). In their Experiment 1a, for example, stimuli were two superimposed gratings of different frequencies presented in a continuous stream with varying inter-stimulus intervals (ISIs) to modulate visibility where increased ISIs decreased visibility. They found that when stimulus visibility decreased, the proportion of reports of HSF decreased faster than the proportion of LSF. In Experiment 2, more complex stimuli were used in combination with PAS ratings, and categorizations based on local information worsened faster with increased ISIs than categorizations based on global information. PAS ratings were generally associated with higher performance for global categorizations (less information is needed to perform the task), and this was especially the case for the two middle PAS ratings (which is not overly surprising as ‘no experience’ is typically related to chance or near-chance performance while ‘clear experience’ is typically related to ceiling or near-ceiling performance, leaving the middle ratings to carry most of the performance difference).

Overall, the authors interpret the findings as support that all-or-none access to different frequency content could be behind graded awareness reports. As the authors themselves note, however, the findings do not in fact rule out that gradedness can exist even within a single stimulus feature although they note that this would be a more complex explanation. While we value the focus on parallel processing, we do not find it convincing evidence for all-or-none processing being behind graded awareness ratings. For example, it is surprising that PAS ratings were not used with the simple grating stimuli as these might show that even for perception of a single grating frequency, graded ratings would be used as they have indeed been in experiments using similar stimuli ([Hobot et al. 2020](#)). Hypothetically, the ratings could in principle reflect all-or-none aspects of various aspects of a grating stimulus of which only some are diagnostic, leading us back to previous discussions on how to interpret the ratings ([Dienes and Seth 2010](#); [Timmermans et al. 2010](#)). Another aspect is that—as the authors mention—LSF information reaches the PFC rapidly, yet studies using PAS typically show longer reaction times for low than high PAS ratings ([Andersen et al. 2016](#)), which speaks more for a general evidence accumulation view where the incoming information is sampled/analysed longer if it is weak. If only LSF information were responsible for reports of a ‘brief glimpse’, it is surprising that it would take longer to report this fast information than to fully process and report HSF information. For these reasons, we believe that there is not yet conclusive evidence in one direction, and yet more work is needed to fully disentangle the two possibilities. Given the basic commitment underlying PAS that consciousness cannot be accessed from the outside and that participants describe their experiences as being gradual, we believe the burden of evidence lies with those who claim that this is in fact not true.

Is PAS superior to other subjective measures?

Sandberg et al. (2010) compared three different four-point scales representing the most common subjective methods at the time—PAS, confidence ratings and post-decision wagering. The study found that participants used all scales in a gradual manner, although PAS showed the least amount of unconscious perception as indicated by a better correlation between accuracy and scale rating as well as lower accuracy at the lowest scale step (which indicated absence of awareness/confidence). The authors suggested that PAS is the better measure because it correlates more with objective correctness than the other scales. Several other publications have replicated these findings (Sandberg and Overgaard 2015; Overgaard 2017). The finding that visibility ratings and confidence ratings behave differently has been replicated many times, e.g. in Rausch et al. (2015) and Rausch and Zehetleitner (2016). While Szczepanowski et al. (2013) reported confidence ratings to correlate better with correctness than PAS, a following re-analysis of the data with a different statistical method did not indicate this (Sandberg et al. 2013).

It may be noted that following the publications of most of the articles supporting PAS, some weaknesses have been pointed out regarding the use of correlation measures when examining the relationship between task accuracy and awareness—in particular the dependency of the correlation strength on the particular criteria used to report awareness (Fleming and Lau 2014). Various methods based on signal detection theory have been proposed as alternatives, and among those, meta-d' (Maniscalco and Lau 2012) is perhaps the most popular today as it handles this so-called type II bias well. It would be interesting to see comparisons of PAS to other measures using meta-d', but we are not aware of any such studies.

The possibly most important challenge to PAS is a challenge to all proposed measures of consciousness: How do we decide what constitutes a good measure? We have previously defended that PAS is a good measure of consciousness because it correlates well with performance—knowing that it in a certain sense is a difficult argument because it might add the unwarranted assumption that consciousness and performance always correlate in nature (but see Michel (2021) for an analysis of why such a strong assumption is not necessary—or even ideal—in practice). Given the fundamental assumption for PAS, mentioned above, the strongest argument in favour of PAS is that participants say the scale represents how they experience clarity of perception.

PAS never intended to be defended as 'the one and only measure of consciousness'. It represents a standpoint, and it takes the consequence of its assumption—that a measure of consciousness must be grounded in the conscious subject rather than a theoretical abstraction. As it can be seen in this discussion, we consider most arguments against the use of PAS to be flawed or at least problematic—although there are certainly several problems with PAS as well. It is our hope that the methodological development and the attempt to take subjective reports seriously will inspire future consciousness research rather than get stuck in an endless sandpit of being 'for or against'.

Data availability

This is a review/perspective paper with no empirical data.

Funding

The work of KS on this article is based upon work from COST Action CA18106, supported by COST (European Cooperation in Science and Technology).

Conflict of interest statement

None declared.

References

- Andersen LM, Pedersen MN, Sandberg K et al. Occipital MEG activity in the early time range (<300 ms) predicts graded changes in perceptual consciousness. *Cereb Cortex* 2016;**26**:2677–88.
- Aru J, Bachmann T. In and out of consciousness: how does conscious processing (D)evolve over time? *Front Psychol* 2017;**8**:1–4.
- Bayne T, Hohwy J, Owen AM. Are there levels of consciousness? *Trends Cogn Sci* 2016;**20**:405–13.
- Binder M, Gociewicz K, Windey B et al. The levels of perceptual processing and the neural correlates of increasing subjective visibility. *Conscious Cogn* 2017;**55**:106–25.
- Christensen MS, Ramsøy TZ, Lund TE et al. An fMRI study of the neural correlates of graded visual perception. *NeuroImage* 2006;**31**:1711–25.
- Dehaene S, Naccache L. Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition* 2001;**79**:1–37.
- Del Pin SH, Skóra Z, Sandberg K et al. Comparing theories of consciousness: object position, not probe modality, reliably influences experience and accuracy in object recognition tasks. *Conscious Cogn* 2020;**84**:102990.
- Derda M, Koculak M, Windey B et al. The role of levels of processing in disentangling the ERP signatures of conscious visual processing. *Conscious Cogn* 2019;**73**:102767.
- Dienes Z, Seth AK. Measuring any conscious content versus measuring the relevant conscious content: comment on Sandberg et al. *Conscious Cogn* 2010;**19**:1079–80.
- Dong MY, Sandberg K, Bibby BM et al. The development of a sense of control scale. *Front Psychol* 2015;**6**:1–14.
- Fazekas P, Overgaard M. A multi-factor account of degrees of awareness. *Cogn Sci* 2018;**42**:1833–59.
- Fleming SM, Lau HC. How to measure metacognition. *Front Hum Neurosci* 2014;**8**:443.
- Förster J, Koivisto M, Revonsuo A. ERP and MEG correlates of visual consciousness: the second decade. *Conscious Cogn* 2020;**80**:102917.
- Hobot J, Koculak M, Paulewicz B et al. Transcranial magnetic stimulation-induced motor cortex activity influences visual awareness judgments. *Front Neurosci* 2020;**14**:580712.
- Irvine E. Old problems with new measures in the science of consciousness. *Br J Philos Sci* 2012;**63**:627–48.
- Kim C, Chong SC. Partial awareness can be induced by independent cognitive access to different spatial frequencies. *Cognition* 2021;**212**:104692.
- Koch C, Massimini M, Boly M et al. Neural correlates of consciousness: progress and problems. *Nat Rev Neurosci* 2016;**17**:307–21.
- Koivisto M, Grassini S. Neural processing around 200 ms after stimulus-onset correlates with subjective visual awareness. *Neuropsychologia* 2016;**84**:235–43.
- Koivisto M, Revonsuo A. Event-related brain potential correlates of visual awareness. *Neurosci Biobehav Rev* 2010;**34**:922–34.
- Kouider S, de Gardelle V, Sackur J et al. How rich is consciousness? The partial awareness hypothesis. *Trends Cogn Sci* 2010;**14**:301–7.

- Lähtenmäki M, Kauramäki J, Sauter DA et al. Auditory affective processing requires awareness. *Emotion (Washington, D C)* 2019;**19**:53–69.
- Lau H, Rosenthal D. Empirical support for higher-order theories of conscious awareness. *Trends Cogn Sci* 2011;**15**:365–73.
- Lohse M, Overgaard M. Emotional priming depends on the degree of conscious experience. *Neuropsychologia* 2019;**128**:96–102.
- Maniscalco B, Lau H. A signal detection theoretic approach for estimating metacognitive sensitivity from confidence ratings. *Conscious Cogn* 2012;**21**:422–30.
- Mazzi C, Bagattini C, Savazzi S. Blind-sight vs. degraded-sight: different measures tell a different story. *Front Psychol* 2016;**7**:901.
- Michel M. The mismeasure of consciousness: a problem of coordination for the Perceptual Awareness Scale. *Philos Sci* 2019;**86**:1239–49.
- Michel M. Calibration in consciousness science. *Erkenntnis* 2021;1–22.
- Mogensen J, Overgaard M. Reorganization of the connectivity between elementary functions – a model connecting conscious states to neural connections. *Front Psychol* 2017;**8**:1–21.
- Nisbett RE, Wilson TD. Telling more than we can know: verbal reports on mental processes. *Psychol Rev* 1977;**84**:231–59.
- Overgaard M. Introspection in science. *Conscious Cogn* 2006;**15**:629–33.
- Overgaard M. Visual experience and blindsight: a methodological review. *Exp Brain Res* 2011;**209**:473–9.
- Overgaard M. The status and future of consciousness research. *Front Psychol* 2017;**8**:1719.
- Overgaard M, FehI K, Mouridsen K et al. Seeing without seeing? Degraded conscious vision in a blindsight patient. *PLoS One* 2008;**3**:e3028.
- Overgaard M, Grünbaum T. Consciousness and modality: on the possible preserved visual consciousness in blindsight subjects. *Conscious Cogn* 2011;**20**:1855–9.
- Overgaard M, Lindeløv J, Svejstrup S et al. Is conscious stimulus identification dependent on knowledge of the perceptual modality? Testing the ‘source misidentification hypothesis’. *Front Psychol* 2013;**4**:1–9.
- Overgaard M, Mogensen J. Reconciling current approaches to blindsight. *Conscious Cogn* 2015;**32**:33–40.
- Overgaard M, Rote J, Mouridsen K et al. Is conscious perception gradual or dichotomous? A comparison of report methodologies during a visual task. *Conscious Cogn* 2006;**15**:700–8.
- Overgaard M, Sandberg K. Kinds of access: different methods for report reveal different kinds of metacognitive access. *Philos Trans R Soc Lond B* 2012;**367**:1287–96.
- Persuh M. Measuring perceptual consciousness. *Front Psychol* 2018;**8**:2320.
- Ramsøy TZ, Overgaard M. Introspection and subliminal perception. *Phenomenol Cognit Sci* 2004;**3**:1–23.
- Rausch M, Müller HJ, Zehetleitner M. Metacognitive sensitivity of subjective reports of decisional confidence and visual experience. *Conscious Cogn* 2015;**35**:192–205.
- Rausch M, Zehetleitner M. Visibility is not equivalent to confidence in a low contrast orientation discrimination task. *Front Psychol* 2016;**7**:591.
- Rosenthal D. Consciousness and confidence. *Neuropsychologia* 2019;**128**:255–65.
- Rounis E, Maniscalco B, Rothwell J et al. Theta-burst transcranial magnetic stimulation to the prefrontal cortex impairs metacognitive visual awareness. *Cogn Neurosci* 2010;**1**:165–75.
- Sandberg K, Bibby BM, Overgaard M. Measuring and testing awareness of emotional face expressions. *Conscious Cogn* 2013;**22**:806–9.
- Sandberg K, Bibby BM, Timmermans B et al. Measuring consciousness: task accuracy and awareness as sigmoid functions of stimulus duration. *Conscious Cogn* 2011;**20**:1659–75.
- Sandberg K, Del Pin SH, Bibby BM et al. Evidence of weak conscious experiences in the exclusion task. *Front Psychol* 2014;**5**:1–6.
- Sandberg K, Overgaard M. Using the Perceptual Awareness Scale (PAS). In: Overgaard M (ed.), *Behavioral Methods in Consciousness Research*. Oxford: Oxford University Press, 2015, 181–96.
- Sandberg K, Timmermans B, Overgaard M et al. Measuring consciousness: is one measure better than the other? *Conscious Cogn* 2010;**19**:1069–78.
- Schwiedrzik CM, Singer W, Melloni L Subjective and objective learning effects dissociate in space and in time. *Proc Natl Acad Sci* 2011;**108**:4506–11.
- Sergent C, Baillet S, Dehaene S. Timing of the brain events underlying access to consciousness during the attentional blink. *Nat Neurosci* 2005;**8**:1391–400.
- Skewes J, Frith C, Overgaard M Awareness and confidence in perceptual decision-making. *Brain Multiphys* 2021;**2**:100030.
- Skóra Z, Del Pin SH, Derda M et al. No validity without a theory— a critical look at subjective measures of consciousness. *Neurosci Conscious* 2021;**2021**:niab009.
- Szczepanowski R, Traczyk J, Wierchoń M et al. The perception of visual emotion: comparing different measures of awareness. *Conscious Cogn* 2013;**22**:212–20.
- Timmermans B, Sandberg K, Cleeremans A et al. Partial awareness distinguishes between measuring conscious perception and conscious content: reply to Dienes and Seth. *Conscious Cogn* 2010;**19**:1081–3.
- Tononi G, Boly M, Massimini M et al. Integrated information theory: from consciousness to its physical substrate. *Nat Rev Neurosci* 2016;**17**:450–61.
- Tsuchiya N, Wilke M, Frässle S et al. No-report paradigms: extracting the true neural correlates of consciousness. *Trends Cogn Sci* 2015;**19**:757–70.
- Windey B, Gevers W, Cleeremans A. Subjective visibility depends on level of processing. *Cognition* 2013;**129**:404–9.