Erin Shaver, Brian Maniscalco, Hakwan Lau
Columbia University

AWARENESS AS CONFIDENCE

Abstract Lesion to the primary visual area in the brain abolishes visual awareness. And yet, patients with such lesions can perform forced-choice visual detection or discrimination tasks better than chance. In performing these tasks, however, they claim that they are only guessing, that is they have little confidence that they are correct. This paper considers whether this reported lack of confidence could help us to characterize the apparent lack of visual awareness. In other words, do confidence and awareness always go together? Or are there other consistent relationships between the two such that we could use confidence ratings as measures of visual awareness? We highlight why the approach of using confidence ratings to assess the level of visual awareness is better than the currently dominant approach, which is to use forced-choice performance itself as an index of awareness. But we also discuss potential pitfalls, both technical and conceptual, in interpreting confidence ratings in relation to visual awareness.

Key Terms Consciousness, Awareness, Confidence, Blindsight, Signal Detection Theory, Phenomenology.

1. Introduction

The history of using confidence ratings as a measure of awareness dates back to the 1880's, when Peirce and Jastrow were examining minute changes in pressure sensation. These authors asserted that a slight tactile pressure change could be detected consciously or unconsciously. The confidence rating measures not the detection of sensation but the subject's "feelings" about the sensation change. Peirce and Jastrow's experiment drew attention to the discordance between awareness and forced-choice responses, and they pointed to the fact that performance judgments may be correct at a rate above chance even when confidence ratings are at zero. In their conclusions, Peirce and Jastrow speculated that confidence might measure a "secondary sensation of a difference between the primary sensations compared". Forced choice judgments allow us to test those primary comparisons (change vs. no change), but a subjective report is necessary to ascertain the degree to which such judgments are consciously made (Peirce and Jastrow, 1885).

This is different from the current mainstream approach to studying visual awareness (Rees *et al.*, 2002). In these studies, we typically require subjects to make forced-choice judgments on some aspects of the visual presentation (whether the target is there or not, or whether stimulus A or B is presented). Researchers typically take the accuracy of such forced-choice judgments as an index of visual awareness. If subjects can perform such tasks, we assume that they are visually aware of the stimuli. If they cannot, i.e. performance is at chance level, we assume that they are visually unaware. The approach we are concerned with here is different, in that the experimenters collect subjective confidence ratings in addition to forced-choice judgments, and awareness is assessed based on these subjective ratings. This article evaluates whether this confidence-based approach is better than the current mainstream approach, and also highlights certain potential pitfalls in using it.

2. Two approaches to using confidence as a measure of awareness

To start, let us distinguish two ways of how one could use confidence ratings to measure awareness (Dienes, in press): (1) high confidence ratings (as determined by subjective report) can be equated with awareness directly; (2) a high correlation between confidence and performance can be used to determine level of awareness. The first approach is vulnerable to the subjects' overstating or understating their level of awareness, depending upon their personal criteria for confidence. Whereas one subject might report a high level of confidence based on a weak hunch, another might report low confidence due to discomfort with even a minor lack of certainty. The confidence rating taken alone is highly subjective and difficult to apply when comparing across subjects.

Proponents of the second approach attempt to address the above issue by measuring the correlation between confidence ratings and the performance of a forced-choice task. This approach focuses on whether subjects are able to place high subjective ratings at the right time, i.e. in trials where their judgments are correct, rather than whether subjects on the whole give high subjective ratings or not. It thereby somewhat alleviates the problem of uncontrolled overall biases that accompany using a liberal vs. conservative strategy. Essentially, the correlation approach focuses on a kind of metacognitive capacity, i.e. the ability to place the subjective ratings appropriately such that they reflect knowledge of when one is likely to be correct.

It is perhaps important to note here that by adopting either approach, one does not necessarily commit oneself to conceptually equating awareness and confidence, or otherwise reducing one in terms of the other. These should be taken as operational definitions, in the sense that if these conditions would capture the necessary and sufficient conditions for awareness, that would satisfy our purpose.

3. Motivations for using confidence as a measure of awareness

3.1. Forced-choice task performance measures are poor

It has previously been argued that performance in forced-choice tasks measures the capacity for visual processing, but not necessarily awareness (Lau, in press). Blindsight, for instance, is a classic example of a situation in which there appears to be a disassociation between performance ability and awareness (Weiskrantz, 1986, 1999). When forced to report on whether the stimulus (shown to the damaged visual field) is present or not, 80-90% of the time the subject will report correctly, indicating that the brain has processed the visual information. The same subject, however, will report that he has no visual awareness in the damaged field, and that he does not perceive what he is seeing. In, other words, we can have good forced-choice task performance and a complete lack of visual awareness at the same time, which means the former cannot be a good measure of the latter.

3.2. Confidence measures of awareness are not necessarily subjective

One apparent attraction of forced-choice task performance is that it can be objectively measured. Confidence ratings, on the other hand, seem to be entirely subjective. One could argue that this subjective nature may well be necessary; after all, consciousness is supposed to be subjective. However, the fact that confidence is subjective somewhat undermines its value as a good scientific measure.

In one recent experiment, subjects were asked to place monetary bets on their responses to forced-choice tasks in order to introduce an "objective measure of awareness" through post-decision wagering. (Persaud *et al*, 2007). After each trial, subjects were required to place a bet as to whether they are correct or incorrect in the forced-choice task. If they are confident that they are correct, presumably they would place a large bet, and if they lack confidence, the bet would presumably be small. The advantage of this approach is, unlike ordinary confidence ratings, now there is an objective answer as to how one should bet, i.e. how one could maximize the amount of money earned towards the end of the experiment. In this sense then, wagering can be used as a substitute for confidence ratings, thereby objectively measuring awareness.

It should be noted that although the approach of directly taking confidence ratings as an index of level of awareness might seem susceptible to subjective biases, the correlation approach as described in the last section essentially assesses a form of metacognitive capacity. One could argue that while the ratings are still subjectively made, the ability to place these ratings on the right trials is determined by objective cognitive factors. For this reason, some have even applied signal detection theory to refine this correlation approach (Kunimoto *et al.*, 2001), to make it more objective.

Signal detection theory is a formalism that allows one to separate subjective decision criteria from cognitive capacity. For instance, in a visual detection task, individual differences in target detection rates could be explained by differences in overall willingness to give positive identification responses, in addition to differences in actual perceptual ability to detect a visual signal. Signal detection theory allows one to disentangle these two factors and assess the actual perceptual or cognitive performance without any confound arising from response bias. Kunimoto *et al.* (2001) applied the standard signal detection model to confidence judgments. They considered confidence judgments to be analogous to correct responses when they were congruent with actual performance (i.e. high confidence for correct responses and low confidence for incorrect responses). The idea was that this approach would allow one to assess a subject's ability to discriminate between his own correct and incorrect responses, independently of that subject's idiosyncratic tendency to give high confidence judgments. If successful, this approach could capture the "subjective" aspects of confidence ratings in a measure more methodologically "objective" than taking subjective reports at face value.

4. Potential pitfalls

4.1. Wrongly applying standard signal detection models to the analysis of confidence ratings

Galvin *et al*. (2003) provided a theoretical framework for how to apply signal detection theory to confidence judgments about one's own performance. They referred to the primary task (e.g. to detect a visual stimulus) as the type 1 task, and the secondary task of judging one's confidence in one's type 1 response as the type 2 task. In their analysis they described what is the theoretically maximal type 2 performance, given a certain type 1 performance. Their analysis revealed that the type 2 task is theoretically constrained in ways different from the type 1 task.

For instance, signal detection theory posits that stimuli are evaluated along a decision axis, which quantitatively characterizes the perceptual or cognitive evidence or signal used to make judgments about the stimuli. In type 1 tasks, stimuli representations are modeled as being normally distributed along the decision axis. That is, members of a given stimulus class tend to occur most frequently at a particular location on the decision axis (the mean of the distribution), with some variance about this mean (the spread of the distribution) in such a way that the distribution is Gaussian. However, Galvin *et al*. (2003) showed that if type 2 tasks are performed using the same decision axis as type 1 tasks, then the type 2 "stimuli" (correct and incorrect responses) are not normally distributed along the decision axis. In fact, they are always systematically and radically different from standard normal distributions. Thus, any standard signal detection theory approaches that assume normal distribution of stimulus representations (such as the standard measure d') cannot be applied to type 2 tasks.

However, Kunimoto *et al*. (2001) assumed in their signal detection treatment of confidence judgments that correct and incorrect responses are distributed normally. If the analysis of Galvin et al. is correct, then Kunimoto et al.'s approach rests on theoretically dubious foundations. Indeed, Evans and Azzopardi (2007) discovered theoretical and empirical flaws in Kunimoto's application of d' analysis to confidence judgments. They showed that given a constant type 1 d', the type 2 d' as computed by Kunimoto et al. varies systematically with type 2 criterion. This theoretically predicted pattern was subsequently found to exist in actual human observers. In a visual target detection task, Evans and Azzopardi manipulated type 2 response criteria of their subjects in one experiment by changing the relative frequency of target stimuli, and in another by allowing subjects a predetermined number of responses at each level of confidence. Across both cases, Kunimoto et al's type 2 d' increased with increases in type 2 response criterion. Thus, it would seem that Kunimoto et al.'s analysis does not deliver the intended bias-free measure of type 2 performance. This is likely due at least in part to Galvin et al.'s observation that the assumptions signal detection theory makes with respect to type 1 tasks do not generalize to type 2 tasks.

Likewise, Galvin *et al.* (2003) showed that for an ideal observer, the range of values that type 2 performance could take is theoretically constrained by type 1 d', and further that the precise level of type 2 performance within that range that actually obtains is determined by the type 1 decision criterion. Because type 2 performance depends on the type 1 d' and decision criterion, it is misleading to straightforwardly evaluate type 2 performance. Rather, Galvin *et al.* (2003) endorse a more sophisticated approach wherein one evaluates type 2 performance by comparing it to the predicted type 2 performance of an ideal observer, given the type 1 d' and decision criterion of the actual observer.

All the technical specifics aside, perhaps the take home message is that signal detection theoretic analysis of confidence ratings can be complicated, with the theoretical details formally worked out only recently (Galvin et al. 2003). Some previous attempts have been shown to be seriously flawed. And research adopting such flawed analysis must be interpreted with caution.

4.2. Validity of the correlation approach

Even if we apply the correct formal model to assess the association between confidence ratings and performance, it is still questionable whether a high correlation between the two should mean a high level of awareness. In some sense, it is trivially true that if one's subjective ratings reflect different levels of effectiveness in information processing, one demonstrates some level of metacognitive awareness, i.e. one is aware of one's own performance. But what we are primarily concerned with here is sensory awareness, in particular whether a visual percept gives conscious phenomenology.

For instance, Kolb and Braun (1995) have studied different masking conditions under which a visual target is not phenomenologically registered, yet subjects could still locate the target at well above chance level. This result is interesting because this is essentially a demonstration of "blindsight in normal observers". The authors used a correlation approach to show that in these conditions, performance and confidence are uncorrelated. Later researchers (Robichaud and Stelmach, 2003 and Morgan *et al.*, 1997), however, have failed to replicate the result; they found instead that confidence and performance remained correlated across trials. This has led some authors to call this line of research utilizing visual masking a "fruitless line of inquiry". However, if the visual phenomenology associated with seeing the targets is actually absent, or at least nearly absent or much reduced compared to normal conditions (Morgan *et al.*, 1997), perhaps what we should challenge is the validity of the correlation approach, but not the masking phenomenon itself.

Indeed, it seems as if there are at least some cases where the correlation approach incorrectly predicts the presence or absence of awareness. According to the correlation approach, when confidence and performance are well correlated, the subject should be aware of the stimulus. However, in one recent experiment we have found that blindsight patient GY can have similar levels of correlation between confidence and performance in his good and blind fields, while still retaining the characteristically stark difference in visual awareness between the two fields (Persaud, Cowey, Maniscalco, Mobbs, Passingham and Lau, manuscript in preparation). This demonstrates that a good correlation between confidence and performance is not sufficient for awareness.

Furthermore, it seems likely that the converse relationship need not hold either: a good correlation between confidence and performance is not necessary for awareness. Visual awareness can occur alongside a poor correlation between confidence and performance in instances where a subject's visual experience systematically fails to represent the external world faithfully in some respect. For instance, a subject administered a psychoactive, hallucinatory substance such as psilocybin while performing a motion detection task might visually experience false motion alongside actual motion. On this basis the subject might endorse incorrect responses with high confidence (mistaking hallucinatory motion for actual motion). The subject might also endorse correct responses with low confidence (not clearly perceiving the difference between actual and hallucinatory motion). Even cognitively normal subjects might fail to give confidence ratings in agreement with actual performance, e.g. if performing judgments on stimuli exploiting known visual illusions. In both scenarios, the failure to give appropriate confidence ratings arises from non-veridical visual awareness, rather than from a lack of awareness.

In sum, it seems as if awareness can be absent even when confidence and performance are well correlated, and likewise that awareness can be present even when confidence and performance are not well correlated. This calls into question the conceptual validity of the correlation approach, as such a correlation seems neither necessary nor sufficient for awareness.

4.3. Confidence may capture too much

The approach of taking confidence level directly as an index of awareness is not without its problems either. People in general vary in how much stock they put into their ability to act on a hunch or to feel their way through a situation. It is possible to imagine that some subjects may be quite used to "trusting their guts" even while not having vivid phenomenological basis for such intuitive feelings. One could argue that such gut feelings or intuitions are still part of the conscious experience. However, at the very least this highlights a problem for using confidence ratings as a measure for awareness in a particular perceptual modality, such as vision. High confidence in a visual task may not reflect high conscious visibility, because such confidence could be driven by non-visual experiences.

5. Closing remarks

The fact that conscious awareness is an essentially subjective phenomenon means that there is a need to take seriously subjective reports like confidence ratings. However, there are caveats in analyzing and interpreting such data. The possibility that one could use sophisticated analysis to make such measures more objective may seem appealing, but as we have highlighted, there are formal and conceptual problems associated with this approach. Interestingly, perhaps this indicates that one of the most basic ways of using subjective ratings, i.e. taking the subjects' word for it as to whether they actually see (Weiskrantz, 1986), may well be the most straightforward and useful approach. If one can demonstrate a difference because this subjective report of visibility between two conditions where forced-choice performance is matched, one can be fairly confident that there is a real difference in terms of the level of awareness there (Lau and Passingham, 2006).

References

Dienes, Z. (in press). "Subjective Measures of Unconscious Knowledge", In R. Banerjee, B. Chakrabarti (Eds.) *Models of Brain and Mind: Physical, Computational and Psychological Approaches*. Elsevier.

Evans, S., Azzopardi, P. (2007). Evaluation of a 'bias-free' measure of awareness. *Spatial Vision*, 20 (1-2), 61-77.

Galvin, S.J., Podd, J.V., Drga, V., Whitmore, J. (2003). Type 2 tasks in the theory of signal delectability: discrimination between correct and incorrect decisions. *Psychonomic Bulletin & Review*, 10 (4), 843-876.

Kolb, C., Braun, J. (1995). Blindsight in normal observers. *Nature*, 377 (6547), 336-338.

Kunimoto, C., Miller, J., Pashler, H. (2001). Confidence and accuracy of near-threshold discrimination responses. *Consciousness and Cognition*, 10, 294-340.

Lau, H.C., Passingham, R.E. (2006). Relative blindsight in normal observers and the neural correlate of visual consciousness. *PNAS*, 103 (49), 18763-18768.

Lau, H.C. (in press). "Are we studying consciousness yet?". In M. Davies, L. Weiskrantz (Eds.) *Frontiers of Consciousness*.

Macmillan, N., Creelman, C. (1991). *Detection theory: a user's guide*. Cambridge [England]; New York, Cambridge University Press.

Morgan, M.J., Mason, A.J.S., Solomon, J.A. (1997). Blindsight in normal subjects? *Nature*, 385, 401-402.

Peirce, C.S., Jastrow, J. (1885). On small differences in sensation. *Memoirs of the National Academy of Sciences*, 3, 317-329.

Persaud, N., McLeod, P., Cowey, A. (2007). Post-decision wagering objectively measures awareness. *Nature Neuroscience*, 10, 257-261.

Pessoa, L. (2007). Fear Perception: Can objective and subjective awareness measures be dissociated? *Journal of Vision*, 7 (4), 10, 1-17

Rees, G., Kreiman, G., Koch, K. (2002). Neural correlates of consciousness in humans. *Nature Reviews Neuroscience*, 3, 261-270.

Robichaud, L., Stelmach, L.B. (2003). Inducing blindsight in normal observers. *Psychonomic Bulletin & Review*, 10 (1), 206-209.

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs*, 74, 1-29.

Weiskrantz, L. (1986). *Blindsight: A Case Study and Implications*. Oxford, Oxford University Press.

Weiskrantz, L., Barbur, J.L., Sahraie, A. (1995). Parameters affecting conscious versus unconscious visual discrimination with damage to the visual cortex (V1). *PNAS*, 92, 6122-6126.

Wesikrantz, L. (1999). *Consciousness Lost and Found.* Oxford, Oxford University Press.