Consciousness and Cognition 20 (2011) 1659–1675

Contents lists available at SciVerse ScienceDirect

Consciousness and Cognition

journal homepage: www.elsevier.com/locate/concog

Measuring consciousness: Task accuracy and awareness as sigmoid functions of stimulus duration

Kristian Sandberg [a,*], Bo Martin Bibby [b], Bert Timmermans [c], Axel Cleeremans [d], Morten Overgaard [a,e]

[a] *Cognitive Neuroscience Research Unit, Hammel Neurorehabilitation and Research Center, Aarhus University, Voldbyvej 15, Hammel, Denmark*
[b] *Department of Biostatistics, Aarhus University, Aarhus, Denmark*
[c] *Department of Psychiatry, Neuroimaging, University Hospital of Cologne, Cologne, Germany*
[d] *Consciousness, Cognition, & Computation Group, Université Libre de Bruxelles, Brussels, Belgium*
[e] *Cognitive Neuroscience Research Unit, Department of Communication and Psychology, Aalborg University, Denmark*

ARTICLE INFO

Article history:
Received 16 March 2011
Available online 29 September 2011

Keywords:
Consciousness
Awareness
Vision
Masking
Unconscious
Subliminal
Regression

ABSTRACT

When consciousness is examined using subjective ratings, the extent to which processing is conscious or unconscious is often estimated by calculating task performance at the subjective threshold or by calculating the correlation between accuracy and awareness. However, both these methods have certain limitations. In the present article, we propose describing task accuracy and awareness as functions of stimulus intensity (thus obtaining an accuracy and an awareness curve) as suggested by Koch and Preuschoff (2007). The estimated lag between the curves describes how much stimulus intensity must increase for awareness to change proportionally as much as accuracy and the slopes of the curves are used to assess how fast accuracy and awareness increases and whether awareness is dichotomous. The method is successfully employed to assess consciousness characteristics on data from four different awareness scales.

1. Introduction

The phenomenon of conscious experience has received increasing scientific attention within the last decades, but in spite of numerous findings within the field, the methodology for studying consciousness is still under development (Overgaard, 2006; Overgaard, Gallagher, & Ramsøy, 2008). At present, several useful methods are employed to assess various aspects of consciousness, including how to estimate the amount of unconscious perception in the most accurate manner. Overall, at least four different behavioral methods can be distinguished, differing in the extent to which they rely on the subjective experiences and reports of the subject: subjective threshold (/rating) approaches, dissociation procedure approaches, type I signal detection theory (SDT) approaches, and type II SDT approaches. As we will argue in the following, none of these methods has (yet?) become the golden standard, and even within each method, a large number of statistical methods are employed. In the present article, we briefly examine the statistical methods used in the subjective threshold (/rating) approach and point out some unresolved issues. We then go onto suggest a novel statistical approach to be used within the approach. First, however, we will briefly introduce the four overall approaches and mention the advantages/disadvantages associated with each.

* Corresponding author.
E-mail address: krissand@rm.dk (K. Sandberg).

1053-8100/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.concog.2011.09.002

1.1. Methods for examining unconscious perception

Subjective threshold approaches have been used since Sidis (1898) presented his participants with single letters or digits at various distances. The participants made a judgement as to whether or not they could see what was on the card followed by an attempt at identification. Sidis observed that even when subjects claimed not to see the letter or digit, they performed above chance on the identification task. This kind of paradigm has also been referred to as the subjective threshold or dissociation approach (Cheesman & Merikle, 1984, 1986; Merikle & Joordens, 1997). Unconscious processing, in this case, is presumed to be responsible for any above-chance performance found when stimuli are below the so-called subjective criterion (i.e. when subjects claim to have no experience of the identity of the stimulus or no confidence in their response) (Snodgrass & Shevrin, 2006). If participants are asked to report their confidence in being correct instead of their experience, this method has also been referred to as establishing unconscious processing by "the guessing criterion" (Dienes, Altmann, Kwan, & Goode, 1995).

The subjective threshold approach has been criticized on different grounds. The most common criticisms are related to the exhaustiveness and exclusiveness assumptions. For the conclusions to be valid, it is required that the subjective measure is able to detect all relevant conscious knowledge (Merikle, 1982; Merikle & Joordens, 1997; Reingold & Merikle, 1988, 1990), i.e. the measure has to be exhaustive. If some conscious processing goes undetected, the report of absence of experience might be invalid, and the above-chance accuracy could be caused by undetected conscious processing. However, if one attempts to tackle the exhaustiveness issue simply by using the most sensitive subjective measure (i.e. the measure indicating the smallest amount of unconscious perception), another problem arises, for there is now a risk that some unconscious perception is erroneously classified as conscious perception (Reingold & Merikle, 1990). In other words, the measure also needs to be exclusive – it should not only measure all of conscious perception (exhaustiveness), but also only conscious perception (exclusiveness).

Different alternative behavioral measures have been proposed to overcome the unrealistic requirement that a single measure can be simultaneously exhaustive and exclusive. A first method—the relative sensitivity approach—was introduced by Reingold and Merikle (1988), who suggested that the search for absolute measures of awareness should simply be abandoned in favor of approaches that seek to compare the sensitivity of direct measures and indirect measures of some discrimination. Direct measures involve tasks in which the instructions make explicit reference to the relevant discrimination, and include objective measures such a recognition or recall. In contrast, indirect measures, such as stem completion in memory tasks, make no reference to the relevant discrimination. By assumption, direct measures should exhibit greater or equal sensitivity than indirect measures to consciously held task-relevant information, for subjects should be expected to be more successful in using conscious information when instructed to do so than when not. Hence, demonstrating that an indirect task is more sensitive to some information than a comparable direct task can only be interpreted as indicating unconscious influences on performance. Jiménez, Méndez, and Cleeremans (1996), for instance, successfully applied the approach to implicit learning, demonstrating that an indirect measure of sequence learning exhibited greater sensitivity to the learned knowledge than a comparable direct measure, which can only be interpreted by assuming that the learned knowledge was implicit.

Another method to overcome the unrealistic assumptions of dissociation logic consists of estimating the relative magnitude of conscious and unconscious influences by using the exclusion tasks as developed by Debner and Jacoby (1994), Jacoby (1998), Jacoby, Toth, and Yonelinas (1993). In this paradigm, the experimental task is to be solved *without* using information from a particular stimulus. The reasoning is that this is only possible as long as the stimulus is consciously perceived. Crucially, if the stimulus (or some aspects of it) is perceived unconsciously, it will influence and produce above-chance performance. For instance, a (potentially unconsciously perceived) word is presented briefly followed by a three-letter word stem. The task is then to complete the word without using the word that was just presented. This is then compared to a baseline at which no word is presented. If the primed words are used more often than in the baseline condition in spite of instructions not to do so, the results are said to demonstrate unconscious perception. A somewhat similar method is the process dissociation procedure (PDP), proposed by Jacoby (1991) in which the exclusion task is compared to an inclusion task for which the instruction is to use the primed word as often as possible. In the more specific context of implicit learning, the process dissociation procedure has been applied to sequence learning by Destrebecqz and Cleeremans (2001), with the same logic: If people show evidence of having learned the elements of a sequence by being able to generate them when so instructed (inclusion), but are also found unable to refrain from generating the same trained sequence elements when specifically instructed to do so (exclusion), one must conclude that they have no conscious control over the learned knowledge and hence that the knowledge is best described as implicit.

Although exclusion tasks and PDP do not seem to require subjective measures, they have also been criticized in several different ways. PDP, in particular, has been questioned from a statistical perspective ((Curran & Hintzman, 1995; Dodson & Johnson, 1996; Graf & Komatsu, 1994), but see also (Jacoby, 1998; Toth, Reingold, & Jacoby, 1995)), while both PDP and exclusion tasks in general have been questioned from the perspective of SDT. The criticism that advocates of SDT raise of exclusion tasks is very similar to the criticism exclusion task proponents have raised against subjective threshold approaches: If the sensory strength of a stimulus is below the criterion for which a participant will categorize it as conscious or seen (but still very vaguely perceived in some manner), the subject might not trust the experience enough to exclude the word (Snodgrass, 2002; Snodgrass & Shevrin, 2006). In this sense, vague conscious experiences might explain the results of exclusion tasks in the same way as for subjective threshold tasks (see Kouider and Dupoux (2004) for a similar critique from a partial awareness perspective).

The use of SDT in consciousness research has been proposed as a method to overcome the unresolved issues of subjective threshold and exclusion paradigms. Explained very briefly SDT is a method to quantify the ability of a particular system to distinguish a given signal from random noise (Peterson, Birdsall, & Fox, 1954; Tanner & Swets, 1954) and has been suggested as a general method to be used in psychophysics (Green & Swets, 1966). The main advantage of SDT is that it calculates discrimination performance independently of the signal intensity needed for the participant to categorize a stimulus as present (i.e. their criterion). SDT is typically applied to consciousness research in one of two ways. The first is a highly conservative method, in which unconscious perception is examined only for stimulus conditions for which the participant has no discriminatory abilities at all ($d' = 0$ or the objective threshold). As there is no ability to discriminate, the results cannot be explained in terms of suboptimal exhaustiveness, thus solving the main issue it set out to solve. This method is often used in priming experiments to demonstrate that a prime was not seen.

The approach ($d' = 0$), however, has, just like its competitors, been criticized in a number of ways (Merikle & Reingold, 1998). One important point is that SDT in itself remains agnostic as to whether discriminations are performed unconsciously or consciously – it simply assesses the discriminatory ability. For this reason, it can be argued that the assumption that SDT (or any other objective threshold approach) makes that all discriminatory abilities are conscious is as questionable as the assumptions made by subjective threshold or exclusion approaches, and hence that $d' = 0$ approaches might categorize much unconscious processing as conscious (i.e. they are not exclusive) (Merikle & Daneman, 2000). In other words, a participant might be sensitive to some information yet remain unaware of it.

Another manner of applying SDT to the study of consciousness is by making use of subjective reports of confidence in one's answer to the experimental task. In this setup, both performance (i.e., discrimination, a "type I" task) as well as confidence ratings (a "type II" task) are analyzed using SDT, and awareness is usually considered to be the result of the type II analysis. This effectively solves the issue that SDT might not index consciousness as such, but only performance, since subjective ratings are explicitly asked for. However, it has been emphasized that great care should be taken when interpreting the results as (for mathematical reasons) type II d' will always be lower than the type I d', type II d' will be dependent upon type I criterion, and criterion jitter can cause invalid results (Galvin, Podd, Drga, & Whitmore, 2003).

The conclusion therefore seems to be that no matter which method is selected for measuring the conscious experience of a participant, there are potential issues to be aware of (no pun intended). For this reason, consensus is not to be expected on which method to use, and a brief look at the literature does indeed reveal that recent studies have been published using both subjective rating approaches (Boyer, Harrison, & Ro, 2005; Koivisto, Mäntylä, & Silvanto, 2010; Persaud & McLeod, 2008; Persaud, McLeod, & Cowey, 2007), exclusion tasks (Visser & Merikle, 1999), type I SDT (Snodgrass & Shevrin, 2006), and type II SDT (Kanai, Walsh, & Tseng, 2010). With all types of measures in use, the important question is then how to treat the data obtained. As mentioned, the present article focuses on making a contribution to how awareness data obtained by subjective ratings can be treated.

1.2. Currently used statistical methods

Although Dienes and Seth (2009) and Sandberg, Timmermans, Overgaard, and Cleeremans (2010) have recently attempted to minimize the issue of suboptimal exhaustiveness and exclusiveness of subjective rating approaches by systematically comparing the exhaustiveness of various scales used for measuring consciousness, the criticism is still valid to some extent. However, even if one accepts the criticism, measures of unconscious processes based on subjective ratings will still be highly useful in comparison across experimental conditions in which accuracy and awareness are suspected to be influenced to different extents. This way, the development of methods examining the relationship between accuracy and awareness is relevant independently of whether one believes that the subjective threshold is an appropriate measure of unconscious processes.

At present, subjective rating approaches commonly estimate unconscious processing at the subjective threshold. The influence of unconscious processes on task performance, however, does not necessarily say anything about whether conscious processes also have an influence, and neither do high ratings of consciousness in themselves, as such ratings in theory might simply reflect a subjective bias for using precisely those ratings. For this reason, the correlation between task performance and the measure of consciousness is often examined as well since a positive correlation would demonstrate awareness (Chan, 1992; Dienes et al., 1995).

Both examination of unconscious/conscious processing using the subjective threshold and correlations between accuracy and awareness have proved very useful so far in the study of consciousness and have been employed in a number of studies (Dienes, 2008; Dienes & Seth, 2009; Dienes et al., 1995; Kuhn & Dienes, 2006; Scott & Dienes, 2008), but they are poorly suited to examine conscious and unconscious processes in certain conditions. When testing for the presence of unconscious processes using the subjective threshold/guessing criterion, only answers in which the subject has very little confidence are used in the analysis. As stimulus intensity increases, subjects become more confident on average, and the frequency of "no confidence" ratings decrease proportionally. This means that when subjects perform somewhat above chance, statistical power is lost when employing this type of analysis. This problem is exemplified in Fig. 1 where the confidence intervals for the lowest awareness ratings in a previous study (Sandberg et al., 2010) are plotted for three scales (PAS, where participants report awareness directly, CR, where participants report their confidence in being correct, and PDW, where participants wager on being correct). In this graph, it is clear that the size of the confidence intervals increase drastically as a function of stimulus duration. Confidence intervals are calculated from binomial distributions without taking into account that each participant is tested repeatedly and contributes with a different number of trials. This makes the confidence intervals much smaller than they would be if

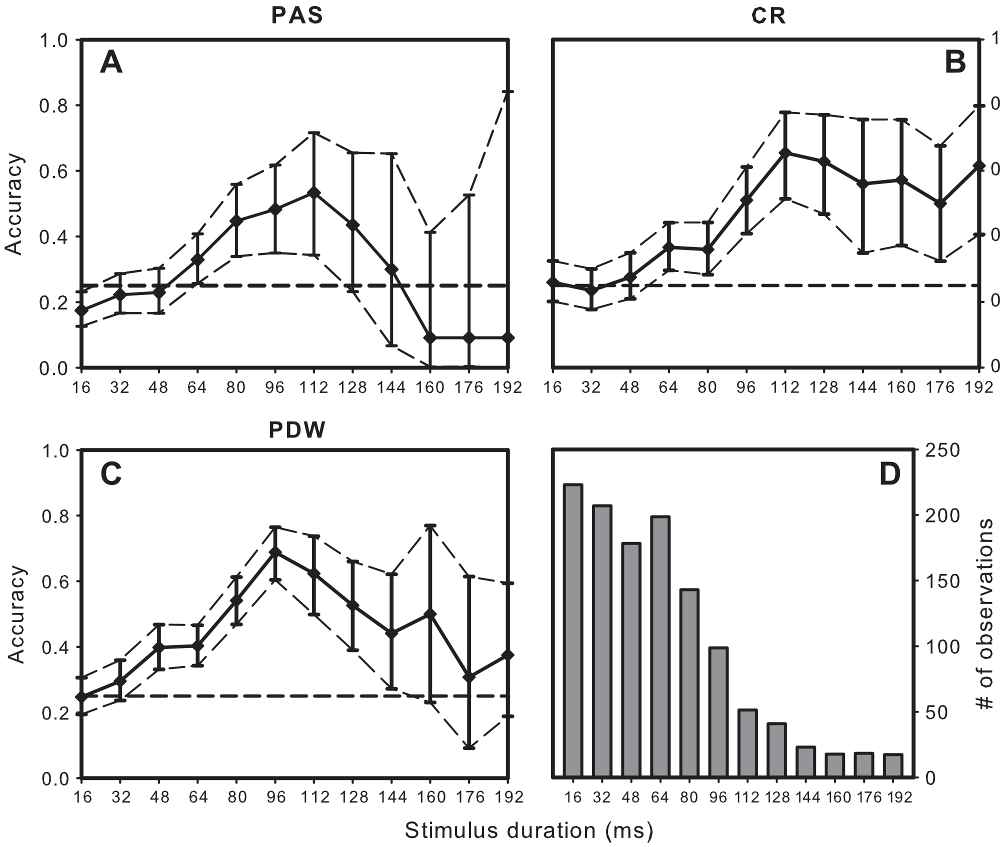

Fig. 1. Subjective threshold performance. (A–C) Performance in a masked identification task at the subjective threshold as indicated by three scales (A: perceptual awareness scale, PAS, B: confidence ratings, CR, and C: post-decision wagering, PDW. Chance (25%) is indicated by the horizontal dotted line. Confidence intervals are calculated from binomial distributions without taking into account that each participant is tested repeatedly and contributes with a different number of trials. This makes the confidence intervals much smaller than they would be if more appropriate, but highly complex, estimations were made. Note that the size of the confidence intervals increase as a function of stimulus duration. (D) Average number of times the lowest awareness rating (no experience, no confidence, or smallest wager) were used for each group of 12 participants. Note how the use of this scale step decreases as a function of stimulus duration. In fact, around half of the participants in each group did not use scale step 1 at all at the highest stimulus durations. Estimating unconscious processing using a subjective threshold approach is thus highly problematic for high stimulus intensities.

more appropriate, but highly complex, estimations were made. In fact, 2 out of 36 participants contributed 47% of the data points, thus making it impossible to meaningfully compare this part of the results between groups. Additionally, the methods do not inform us about whether unconscious processes can account for some of the performance when a subject reports more confidence in his answer (or reports a clearer conscious experience). In order to do this, we would have to examine if the confidence or conscious experience reported by the subject can account for the entire performance pattern. One method that potentially resolves both of these issues is briefly described by Koch and Preuschoff (2007).

1.3. Accuracy and awareness as sigmoid functions

Koch and Preuschoff (2007) suggested that both task accuracy and awareness ratings be plotted as functions of stimulus intensity (see Fig. 2 for an image of such curves). If unconscious processing is responsible for (some) task performance, the awareness curve is shifted along the x-axis compared to the accuracy curve. This means that stimulus intensity has to be higher for awareness ratings to have changed proportionally as much as task accuracy (independently of what the exact minimum and maximum values are). In other words, awareness is lagging accuracy in such a situation. An estimate of this lag would say something about the degree to which unconscious processes influence task performance. If no lag is present, conscious perception follow accuracy perfectly (unconscious processes could, of course, still play a part in the formation of the conscious percept). If a lag is present, consciousness can be said to be lagging awareness (if the exhaustiveness of the reports is trusted), or at the very least the size of the lag can be examined across experimental conditions in order to test for a differential impact on accuracy and awareness by an experimental manipulation.

Koch and Preuschoff (2007) further suggested that information is present in the slopes of the curves; according to them, the steepness of the slope of the awareness curve gives information about whether consciousness is an all-or-none phenomenon or whether it increases gradually. If the awareness curve is close to vertical, i.e. the slope is close to infinite,

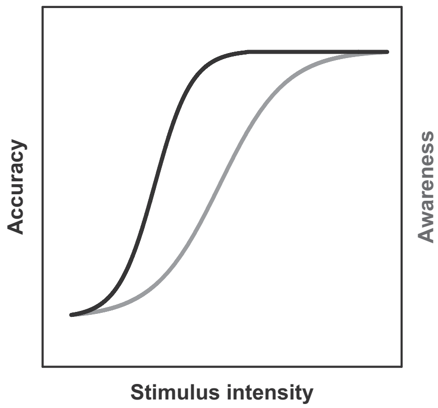

Fig. 2. Example of task accuracy (black curve) and awareness rating (gray curve) as a function of stimulus intensity. If the right stimulus intensities are used, accuracy ranges from chance to close to 100%, whereas awareness ranges from minimum to maximum awareness (whatever the values might be).

consciousness can be said to be an all-or-none phenomenon. It is also possible to compare the slopes of the awareness and accuracy curves in order to examine whether awareness builds up more or less abruptly than the ability to report correctly. If the slopes are not significantly different, awareness and accuracy increase in an identical manner.

Previous studies (Overgaard, Rote, Mouridsen, & Ramsøy, 2006) have used another method to examine whether consciousness can be described as a gradual or dichotomous phenomenon. They compared task accuracy across awareness ratings on a 4-point scale, reasoning that if the probability of correctly identifying a stimulus increases as a function of awareness rating, there is reason to believe that subjects are correct when they claim to have an experience that is neither completely clear nor completely unclear. The method thus examines whether descriptions of various levels of perceptual clarity are meaningful, whereas curve comparisons examine how fast *average* awareness ratings increase across stimulus durations. It should be noted that these are two separate theoretical understandings of graded awareness. A gradual increase of the average awareness could occur even if a scale is used completely dichotomously, and/or if there is no real correlation between the individual awareness ratings and specific accuracies. This last case could occur if subjects use higher awareness ratings in general when stimulus duration increases, but they do not exclusively use the individual high awareness ratings on the trials in which they are correct.

Apart from the uses of the method suggested by Koch and Preuschoff (2007), at least two additional uses can be thought of. First, the estimated awareness function can be subtracted from the estimated accuracy function (taking into account that the scales are different) in order to obtain a difference function. This function will indicate at which stimulus intensity unconscious processing peaks and will provide an estimate of the amount of processing explained by unconscious processing. Again, the validity of the exact values of such a method is dependent on the degree to which the subjective report is valid in general, but the method can be used to compare conditions even if the individual values are invalid. The stimulus durations for which the largest amounts of unconscious processing is found using such difference curves can be compared to the results obtained using the subjective threshold approach in order to examine if the two methods indicate that unconscious processing occurs at similar stimulus intensities. As we argue above and show in Fig. 1, the confidence intervals of the amount of unconscious processing estimated by the subjective threshold approach are quite large at high stimulus intensities. This should not be a problem for the results obtained by difference curves as the confidence intervals at high stimulus intensities do not increase in the same way as for subjective approaches. In fact, the confidence intervals around the upper and lower plateaus are very small.

Second, the estimated accuracy and awareness functions can be used to predict an expected accuracy/awareness relationship after some experimental manipulation. The hypothesis could be that accuracy is affected more than awareness by this manipulation. Imagine, for instance, that performance is inhibited in a visual discrimination task using transcranial magnetic stimulation (TMS), causing a drop in accuracy of 10%. Knowing that accuracy and awareness are described by different functions, we cannot simply expect that awareness is reduced 10% as well if the two are inhibited to the same extent. A traditional method around this problem has been to test accuracy and awareness before the manipulation, and then adjust stimulus intensity between blocks after the manipulation in order to get the same accuracy level on average. Then awareness can be compared before and after the manipulation as accuracy is believed to be stable across the two conditions. In effect, however, an accuracy level a is obtained with a certain degree of variance before the manipulation, and the same accuracy a is obtained after the manipulation, but most likely with a larger variance since stimulus intensity is varied across blocks. Since accuracy and awareness vary differently as a function of stimulus intensity, the results obtained with such a method might be invalid as any difference between measured and expected awareness could simply be an effect of averaging. As an alternative to this method, we suggest simply obtaining the accuracy and awareness functions before the experimental manipulation and then look for the new accuracy level on our accuracy curve and see what the corresponding awareness level should be. It can then be examined if the observed and expected awareness values are significantly different.

Using data from two experiments, we developed the suggested method and examined whether it can be used to estimate awareness. In Experiment 1, we examined whether awareness can be described by a sigmoid function as is commonly the case for task accuracy in identification tasks. We estimated the relationship between accuracy and awareness for three different awareness scales. In Experiment 2, we applied the notions of exclusiveness and exhaustiveness to the curve comparisons in order to increase the validity of the obtained results.

2. Experiment 1

To assess whether the suggested curve comparison can be conducted in a statistically meaningful manner and whether the newly suggested statistics are able to provide insight into the characteristics of consciousness, we tested the method on previously gathered experimental data (Sandberg et al., 2010). We compared three sets of accuracy and awareness curves obtained using different awareness scales in terms of lag and slope as suggested by Koch and Preuschoff (2007) in order to test whether the results are similar. In the following we describe the experiment briefly. For a more detailed description, see Sandberg et al. (2010).

2.1. Materials and methods

2.1.1. Procedure

Thirty six subjects performed a simple masked visual identification task consisting of 336 trials divided into 5 blocks. On each trial, one of four possible geometric shapes appeared on a screen on which subjects were fixating. Stimulus durations varied from 16 ms to 192 ms. The task was to identify the figure (four-alternative forced choice). After attempted identification, the subjects were requested to supply information about their conscious experience using one of three possible scales (each scale was used by 12 subjects). When using the Perceptual awareness scale (Ramsøy & Overgaard, 2004), PAS, subjects were asked to describe the quality of their visual experience when they saw the stimuli. Using confidence ratings, CR, subjects reported their confidence in their performance on the identification task. Using Post-decision wagering (Persaud et al., 2007), PDW, subjects placed a wager on the correctness of their response in the identification task.

All scales were displayed as a bar divided into 4 equally large segments. A number and a description were displayed below each segment. For PAS, the descriptions were: (1) no experience, (2) a vague experience, (3) an almost clear experience, and (4) a clear experience. For CR, the descriptions were: (1) not confident at all, (2) slightly confident, (3) quite confident, and (4) very confident. For PDW, the descriptions were: (1) €5, (2) €10, (3) €15, and (4) €20. The only difference between groups was thus the scale-specific instruction and the descriptions that appeared on screen when using the 4-point scale. Imaginary money was used for PDW participants (as has been done previously by Persaud et al. (2007) so as not to introduce additional differences between conditions.

2.1.2. Curve fitting (nonlinear regression)

We chose to estimate the lag and slope differences between curves directly from the differences in curve parameters by describing the data using nonlinear functions. From a theoretical perspective, sigmoid functions were expected, and the sigmoid shape was clearly identified by inspecting plots of the data. We chose the 4-parameter sigmoid function for all the analyses because it gave a good fit to all individual data. Furthermore, the function is very flexible and its parameters all have interpretations relevant to the problem at hand.

The 4-parameter sigmoid function is given by the following expression:

$$F(x) = a + \frac{b-a}{1+e^{\frac{c-x}{d}}}$$

where a denotes the lower plateau, b denotes the upper plateau, c denotes the x-value of the center point of the slope, and d is a measure of the steepness of the slope (see Fig. 3). Using this model, the lag between accuracy and awareness can be based on the difference in the values of the c parameter between the two functions. The steepness of the slope increases as the parameter d approaches 0. Thus, awareness should be taken to be all-or-none if the parameter d is close to 0. If the d parameters of the awareness and accuracy curves are identical, awareness and accuracy increase in a similar manner. The two most interesting parameters are thus c and d. The absolute values of the lower and upper plateau (a and b) are not important as both c and d are estimated on the basis of x-axis values corresponding to proportional changes of values on the y-axis, i.e. c is defined as the x-value that corresponds to the y-value that is exactly half of the difference between a and b. This way, comparisons between the c and/or d values can be made between an accuracy curve and an awareness curve without the two curves sharing the same a and/or b value.

2.2. Results

4-Parameter sigmoid functions were fitted to the accuracy data (the accuracy curve) and to the awareness data (the awareness curve) for each subject using a nonlinear mixed effects regression model with random effects corresponding to subject and curve (accuracy and awareness) within subjects included for each of the parameters a, b, c, and d. The larger

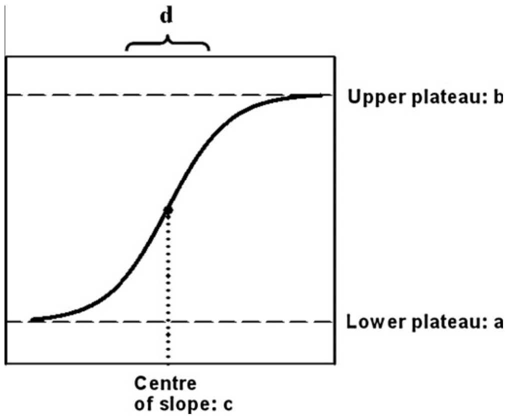

Fig. 3. The 4-parameter sigmoid function. (a) Denotes the lower plateau, (b) denotes the higher plateau, (c) denotes the center of the slope, and (d) is a measure of the steepness of the slope.

Table 1
Experiment 1 (Sandberg et al., 2010): Sigmoid parameters (with 95% confidence intervals).

Parameter		*a*	*b*	*c*	*d*
Scale					
PAS	Accuracy	0.25(0.21; 0.29)	0.96(0.94; 0.98)	68.5(60.8; 76.2)	11.0(8.23; 13.8)
	Awareness	1.27(1.17; 1.37)	3.60(3.52; 3.69)	96.1(88.1; 104.0)	22.1(18.8; 25.4)
Confidence	Accuracy	0.25(0.21; 0.29)	0.96(0.94; 0.98)	79.2(72.2; 87.6)	11.1(8.35; 13.9)
	Awareness	1.27(1.17; 1.37)	3.60(3.52; 3.69)	94.1(86.4; 101.9)	17.7(14.8; 20.7)
Wagering	Accuracy	0.25(0.21; 0.29)	0.96(0.94; 0.98)	74.4(66.5; 82.3)	14.1(11.1; 17.1)
	Awareness	1.27(1.17; 1.37)	3.60(3.52; 3.69)	98.3(90.5; 106.1)	18.1(15.2; 21.0)

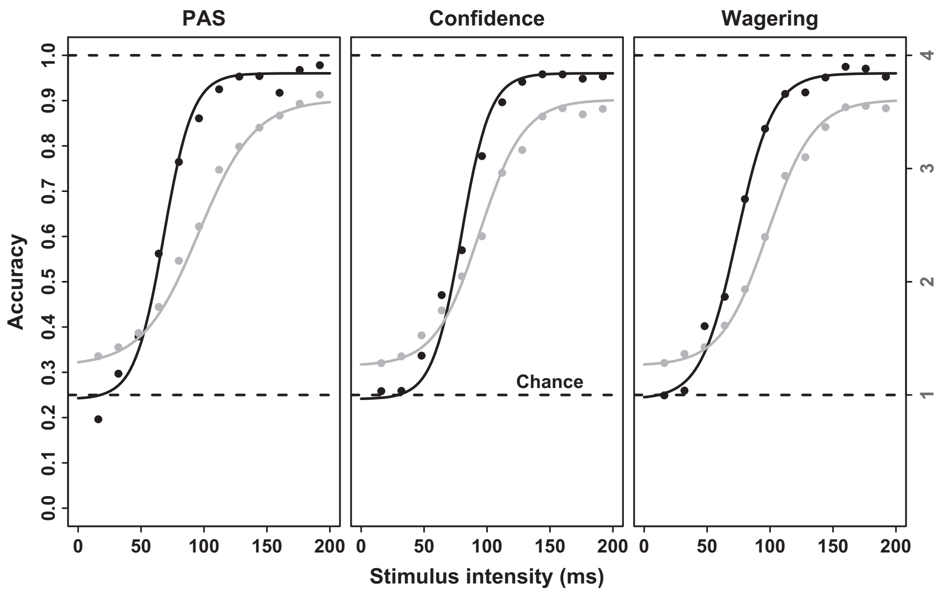

Fig. 4. Data from Experiment 1. Accuracy (black) and awareness ratings (gray) plotted as functions of stimulus duration for each group. 4-Parameter sigmoid functions are fitted to the results.

variation in the accuracy scores[1] was taken into account in a heterogeneous within-subject variance model. The data were analyzed using R version 2.7.2 (R Development Core Team).

[1] The larger variation is caused by the accuracy variable having only two steps (correct/incorrect) whereas the awareness variable has four steps.

The parameters of the nonlinear regression models are shown with 95% confidence intervals in Table 1 and the group curves are shown in Fig. 4 (see Fig. A1 to inspect the individual curves). The parameters *a* and *b* (the lower and upper plateaus in the graphs) did not differ between groups for either awareness curves or accuracy curves ($p = .98$), and common values were estimated whereas for the parameters *c* (center of slope) and *d* (steepness of slope) group estimates were obtained.

2.2.1. Unconscious processing

For all scales, a significant horizontal shift was found between the accuracy and awareness curves. For PAS, the shift was 27.6 ms (95% CI: 20.1; 35.2, $t = 7.14$, $p < .0001$), for CR, the shift was 14.2 ms (95% CI: 6.9; 21.5, $t = 3.77$ $p = .0002$), and for PDW, the shift was 23.9 ms (95% CI: 16.3; 31.4, $t = 6.14$, $p < .0001$). The scales did not indicate awareness to be lagging accuracy to the same extent ($p = .0316$). The only significant difference in lag size was found between PAS and CR, where PAS indicated a 13.2 ms (95% CI: 3.2; 23.2) larger lag ($t = 2.58$, $p = .0101$). We also found a tendency for PDW to indicate a larger lag than CR: 9.68 ms (95% CI: −0.30; 19.7, $t = 1.89$ $p = .0597$). PAS and PDW did not differ significantly ($t = 0.67$, $p = .501$).

2.2.2. Gradual increase of average awareness

No scale indicated that consciousness is an all-or-none phenomenon. A *d* value of 0 would give an undefined slope so the hypothesis that $d = 0.1$ was tested and rejected for each scale with $p < .0001$ for all scales. For PAS, the value of *d* for the awareness curve was 1.93 times higher (95% CI: 1.55; 2.41) than for the corresponding accuracy curve ($t = 5.78$, $p < .0001$). For CR, it was 1.54 times higher (95% CI: 1.24; 1.92) ($t = 3.82$, $p = .0001$), and for PDW 1.28 times higher (95% CI: 1.04; 1.59) ($t = 2.26$, $p = .0242$). All scales thus indicated that the awareness function (averaged over many trials) increased more gradually. The differences in steepness between the accuracy curve and the awareness curve were not identical between groups ($p = .0436$), although the only significant difference found was between PAS and PDW ($t = 2.62$, $p = .0090$), the PDW curve being the steepest.

2.3. Discussion

From a statistical perspective our analysis was highly successful. We found that both accuracy and awareness ratings in the data from Sandberg et al. (2010) were well described as 4-parameter sigmoid functions of stimulus duration, even on an individual subject level. We found that all three scales indicated awareness to be lagging accuracy in the masked identification task. The slope analyses indicated that conscious experience increases more gradually than task performance for all scales. However, as with the traditional methods of analysis, the results obtained differed somewhat between the scales. Before mentioning the differences, it should be noted that these differences between scales are mainly relevant if one should wish to draw conclusions about the absolute values indicated in the analyses (e.g. the amount of subliminal processing). If one wishes only to test the impact of an experimental manipulation upon accuracy and awareness, the specific estimates of a scale is not crucial.

The lags between awareness and accuracy differed between scales in an interesting manner. PAS, along with PDW, seemed to indicate a larger lag than CR both when examining differences between the *c* parameter of the curves. For PDW, the result is not surprising as the large lag is probably caused by risk aversion causing subjects to withhold high wagers (Dienes & Seth, 2009; Schurger & Sher, 2008). For PAS, the result is more interesting as previous analyses (Sandberg et al., 2010) showed that the relationship between a specific awareness rating and a specific accuracy level was better for PAS than for CR and PDW, and that PAS indicated less unconscious processing by the guessing criterion. One explanation could be that CR is used more dichotomously than PAS: A rating of 1 on the CR scale will thus be used across a larger range of stimulus durations (and hence accuracy levels) than when using PAS, which will cause a worse correlation between accuracy and awareness and at the same time more unconscious processing by the guessing criterion.

In order to explain the results of the scales in the present analysis, it is useful to look at the preconditions for using the traditional methods and the novel method. When analyzing data by the guessing criterion, two aspects are of key importance: exhaustiveness and exclusiveness (Reingold & Merikle, 1988). A measure is maximally exhaustive when it detects all conscious performance, and it is maximally exclusive when no unconscious performance is misclassified as conscious performance. The most valid results are thus obtained when the subject uses the lowest awareness rating only when they are actually guessing or have no experience of the stimulus and they do not increase their rating until they experience the stimulus in a clearer manner. With the proposed analyses, these preconditions also apply, but since all awareness ratings are used in making the awareness curve, there should also be some restraint on when the highest awareness rating is used.

Applying the notions of maximum exhaustiveness and exclusiveness to this end of the scale, a logical precondition is that the top scale step should be used as soon the subject sees the stimulus fully or has full confidence (exhaustiveness) but not before (exclusiveness). This would be easily achieved with a CR scale, as the highest step could simply be termed "full confidence" or something similar. However, a problem arises when using a scale asking about the clarity of conscious experience. As Kouider, de Gardelle, Sackur, and Dupoux (2010), Kouider and Dupoux (2004) have previously pointed out, complete awareness of relevant stimulus characteristics for performing a classification is not necessarily complete awareness of *all* stimulus characteristics. So when subjects rate the clarity of their conscious experience, it would thus seem logical that they would rate it according to how unclear and clear it could be in general, unless specifically instructed to use the top scale step as soon as they have the relevant information to make the discrimination.

If we look at the scales as used in the experiment, we find that PDW does not meet at least two preconditions: subjects do not wager low only when having no experience, and they do not necessarily wager high as soon as and only when they see the relevant features. For PDW to meet the preconditions would require that subjects be instructed precisely how to wager, in which case it would become very similar to a CR scale. Subjects using PAS seemed to use the lowest scale step when they had no experience, and they seemed to increase their ratings as soon as and only when their experience became clearer (although see Dienes and Seth (2010) for a critique of this, and Timmermans, Sandberg, Cleeremans, and Overgaard (2010) for the reply). However, as subjects were not instructed to use the highest scale step as soon as they could see the stimulus clearly enough to identify it, it is highly probable that PAS did not meet this precondition.

Subjects using CR seemed somewhat more reluctant to raise their confidence than PAS subjects, thus making it probable that CR does not meet the precondition that ratings should increase as soon as some conscious information is present. As mentioned above, a top scale step with the description "full confidence" would seem to meet the precondition that the top step should be used as soon as the subject has awareness of relevant stimulus characteristics, but not before. However, in the present experiment, the top step of the CR scale had the description "Very confident", thus not meeting the precondition, and in fact making it more likely that the awareness curve would peak too early (the scale thus not being fully exclusive). Thus, the difference in lag between PAS and CR found in the present experiment might be explained by PAS indicating too little awareness at high stimulus durations while CR is indicating too much.

Summing up, this first experiment allowed us to establish that the statistical methods we applied were highly successful as such, but due to certain preconditions not being met, the results obtained might not be as valid as possible. We were interested in examining whether better instructions would change the rating behavior of the subjects. In a second experiment, we thus tested a single group of subjects using PAS with modified instructions on how to use the scale. The new instructions were designed to meet all the relevant preconditions, and the goal was thus to explore whether such instructions would reveal reduced unconscious processing as indicated by the analysis.

3. Experiment 2

3.1. Materials and methods

Experiment 2 was identical to Experiment 1 with a few exceptions. 11 new subjects were recruited and all of them used a 3-step version of the PAS scale. The first two steps were very similar to the ones used in Experiment 1. The descriptions of the first two steps were (1) no experience, and (2) a brief glimpse. Subjects were instructed to use the second category when they had a vague experience of the figure or when they saw a brief glimpse of it. The third scale step was a combination of steps 3 and 4 in Experiment 1. The description was (3) an almost clear or completely clear experience. The key difference between experiments was that subjects were instructed to use this third category when they experienced the figure clearly enough to identify it, thus meeting the proposed preconditions for the analysis. The experiment consisted of 384 trials grouped into 4 blocks. 12 different stimulus durations were used as in Experiment 1, and these appeared 8 times in each block. The order of the stimuli was pseudorandom.

3.2. Results

We first observed that the categories on the new 3-step PAS scale were used similarly to the categories on the 4-step scale (see Fig. A2). 4-Parameter sigmoid functions were then fitted to the data as in Experiment 1. The parameters of the nonlinear regression models are shown in Table 2 and the group curves are shown in Fig. 5 (see Fig. A3 for the individual subject curves).

3.2.1. Unconscious processing

As in Experiment 1, a lag between accuracy and awareness was found. The lag was 18.4 ms (95% CI: 13.3; 23.5, $t = 7.02$, $p < .0001$). Importantly, however, the size of the lag was found to be 9.2 ms (95% CI: 0.02; 18.4) smaller than in Experiment 1 ($z = 1.96$, $p = .0495$).

3.2.2. Gradual increase of consciousness

As in Experiment 1, consciousness was found to be gradual ($p < .0001$), and the hypothesis of identical slopes was rejected: We found that the d value for the awareness curve was 2.26 times higher (95% CI: 1.58; 3.23) than for the

Table 2
Experiment 2: Sigmoid parameters (with 95% confidence intervals).

Parameter		*a*	*b*	*c*	*d*
Scale					
PAS	Accuracy	0.24(0.18; 0.31)	0.96(0.94; 0.99)	64.7(59.9; 69.6)	9.78(7.53; 12.7)
	Awareness	1.20(1.05; 1.34)	2.89(2.82; 2.96)	83.2(78.0; 88.3)	22.1(17.3; 28.2)

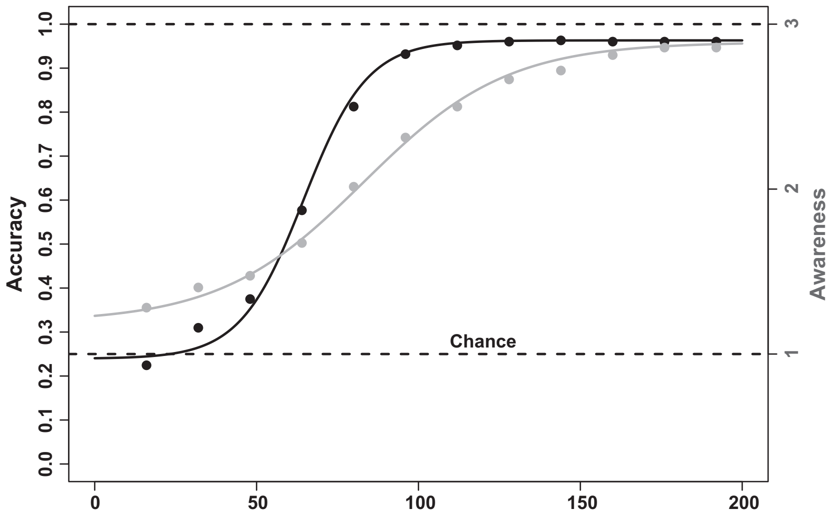

Fig. 5. Data from Experiment 2. Accuracy (black) and awareness ratings (gray) plotted as functions of stimulus duration for each group. 4-Parameter sigmoid functions are fitted to the results.

corresponding accuracy curve ($t = 4.44$, $p < .0001$). The differences in steepness between the accuracy curve and the awareness curve did not differ between experiments 1 and 2 ($z = 0.73$, $p = .46$).

3.3. Discussion

Using a scale that was designed to meet the preconditions of maximum exhaustiveness and exclusiveness at both ends of the scale, we were able to decrease the amount of unconscious processing indicated by the analysis. The amount of unconscious processing indicated by the analysis was significantly lower in the second experiment than in the first experiment.

Interestingly, the awareness function did not simply gain a steeper slope, the entire curve shifted to the left, and a relatively higher top plateau was found. The relatively higher top plateau of the 3-step PAS cannot explain the smaller lag as the d parameter did not differ from the one found for the 4-step PAS. Given a stable d, a change in either of the plateaus does not change the c value. However, it is noteworthy that the new PAS awareness curve did not simply become identical to the CR awareness curve observed in Experiment 1, indicating that even when PAS was used in a more exclusive manner, it was still not used as a confidence scale, and the two types of reports still seem to differ as indicated by previous research (Sandberg et al., 2010).

4. General discussion

Over two experiments, we have introduced a novel method to be used in consciousness research. Using nonlinear regression models (4-parameter sigmoid functions), we were able to describe data obtained with four different awareness scales well, also on an individual subject level.

Using curve parameters, we estimated the amount of unconscious processing involved in a specific task over a wide range of stimulus intensities while making use of all the data and not just the cases in which awareness is reported as completely absent; we estimated the average lag between accuracy and awareness, and we examined how gradually awareness builds up as a function of stimulus intensity (note that the method in itself remains neutral on the discussion of whether each individual experience is an all-or-none phenomenon or whether partial awareness (in some form) exists – see below for a discussion of partial awareness). A shallow slope of the awareness function means that awareness builds up slowly. These two results were combined in a difference function estimating how much accuracy had increased without a corresponding increase of awareness as a function of stimulus duration.

Our results are interesting as they emphasize the importance of a method that is able to estimate unconscious processing at various stimulus intensities and levels of experience. The relative frequency of the subjective reports used in subjective threshold/guessing criterion analysis compared to those not used (reports of "no experience" compared to any other report) decrease as a function of stimulus duration, meaning that the stimulus intensities at which unconscious processing would be most likely to be found, the amount of data is already reduced. Using the estimated sigmoid functions a relative difference curve (between accuracy and awareness) can be created that does not share this problem (see Fig. 6 for examples of such curves). The curves in Fig. 6 are plots of relative changes in accuracy and awareness – i.e. $f(y) = ((y_{\text{accuracy}} - a_{\text{accuracy}})/(b_{\text{accuracy}} - a_{\text{accuracy}})) - ((y_{\text{awareness}} - a_{\text{awareness}})/(b_{\text{awareness}} - a_{\text{awareness}}))$, where y is an y axis value between a and b. In the generation of the curves, we exploit the fact that our curve parameters are estimated with a confidence interval; we have thus simulated

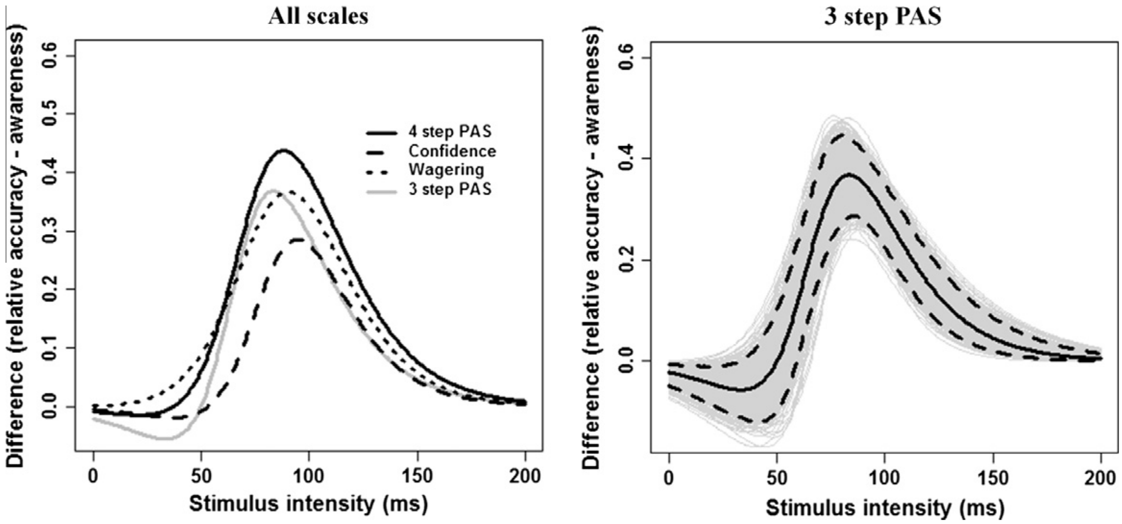

Fig. 6. Relative difference curves. Awareness was subtracted from accuracy in a relative manner, i.e. $((y_{accuracy} - a_{accuracy})/(b_{accuracy} - a_{accuracy})) - ((y_{awareness} - a_{awareness})/(b_{awareness} - a_{awareness}))$. Accuracy and awareness curves are simulated based on the asymptotic normal distribution of the parameter estimates defined by the estimated mean and variance–covariance matrix. The graphs are based on differences between 1000 such simulated curves and empirical point wise confidence intervals. Left: Relative accuracy–awareness for all scales. Right: Relative accuracy–awareness for the 3-step PAS (note the narrow confidence intervals for low and high stimulus durations).

accuracy and awareness curves based on the asymptotic normal distribution of the parameter estimates defined by the estimated mean and variance–covariance matrix. The graphs in Fig. 6 are based on differences between 1000 such curves and empirical point wise confidence intervals. The graphs rise from baseline from around 50–60 ms, they peak at around 80–95 ms, and they return to baseline at around 150–160 ms. Interestingly, these time points are quite similar to the results indicated by the subjective threshold/guessing criterion analyses displayed in Fig. 1, thus indicating that the two quite different methods indicate the presence of unconscious processing at roughly the same time.

4.1. Traditional methods

The novel method of curve comparisons is not presented in order to make the traditional analyses obsolete. In some types of experiments, it is practically impossible (or at least very inconvenient) to gather the amount of data needed for constructing the accuracy and awareness curves. In our experiments, this was not an issue because various stimulus conditions were created simply by varying stimulus duration. In other types of consciousness studies, such as artificial grammar tasks, creating and testing 10 (or more) different stimulus conditions might be practically impossible.

Similarly, much work still needs to be done comparing the various behavioral methods for examining conscious perception. Traditional methods will be better suited for at least parts of this work. An example could be the examination of the criticism that exclusion tasks are not exhaustive using, for instance, PAS and traditional methods of analysis. Here, it could be examined whether any residual awareness is present when participants use primed words at a rate above chance.

Inspecting the literature on exclusion tasks, we notice that when multiple stimulus durations are used in exclusion tasks, different amounts of unconscious processing are found for these stimulus durations (e.g. (Merikle & Joordens, 1997). More unconscious processing is found for stimulus durations of 43–71 ms than for 0–29 ms and 214 ms. The peak in unconscious processing seems to be somewhat earlier than subjective threshold approaches (and curve analyses) indicate. Some (or possibly all) of this difference could be explained by the fact that any trials in which the stimulus is seen fairly clearly, it will be excluded and thus push average "performance" (i.e. reports of the target) towards zero. Logically, the chance of this occurring is larger as stimulus intensity increases. In subjective threshold approaches, on the other hand, calculation of classification accuracy for images that are reported as "unseen" is not directly influenced by performance on other trials. Interestingly, the probability of participants using PAS ratings of 2 (i.e. "weak glimpse" or "brief experience") as a function of stimulus duration follows a positively skewed distribution with a median of around 60–80 ms. Reports of "weak glimpses" are thus much more likely to occur at the times when unconscious processing is indicated by exclusion tasks. This would indeed seem to indicate that weak glimpses account for at least some of the unconscious processing in exclusions tasks, and that trial rejections based on ratings on a subjective scale might improve the validity of results obtained with exclusion tasks. In subjective threshold approaches, on the other hand, the drop in unconscious processing is presumably caused the fact that at high stimulus intensities, the lowest awareness rating is only used when participants are aware that they were not attending the stimuli (and therefore are likely to be at chance). However, further experiments are needed before any definite conclusions are made.

4.2. Illusory perception?

It is interesting to notice that the plateaus of the awareness functions seem to differ somewhat (relatively) to those of the accuracy functions. It seems that participants do not consistently report awareness as fully absent, and neither do they consistently report awareness as completely clear (similar observations are made for confidence and wagering behavior). The question is why this is the case, and whether it impacts on the validity of the results.

Naturally, the accuracy function must start at chance under normal circumstances. It is hard to imagine (in a task that does not examine illusions of some kind) that behavior falls below chance, and by increasing the difficulty, chance performance can usually be achieved. Similarly, near-100% accuracy can be achieved by making the task easier. In our experiment, accuracy peaked at 96%, the few misses probably caused by the participants not attending the screen or accidentally pressing the wrong response keys on these trials. Interestingly, at least the bottom plateau of the accuracy is independent of whatever response criterion the participant adopts, and the upper plateau can be reached easily if the participant shows no strong bias for one response option. The awareness function behaves somewhat differently. At the lower plateau, the type 1 d' is 0 (participants cannot distinguish the stimuli behaviorally), and naturally this will usually cause the type 2 d' to be 0 as well (participants cannot tell which of their responses were correct). However, unless participants adopt an ultraconservative type 2 criterion, some reports of weak confidence or experience will be given. They will not meaningfully distinguish the stimuli, but they will nevertheless cause the plateau of the awareness function to be higher than the absolute minimum value. In other words, the bottom plateau of the accuracy function is independent of whatever response criterion the participants might have, but the bottom plateau of the awareness function is dependent on how likely the participants are to report vague experiences in general.

The nature of such reports of vague experiences or small amounts of confidence is difficult to be sure of. One possibility draws on SDT: Even when at chance, the neural activity generated on some trials appear more like those generated by a particular stimulus than by another, even if this occurs simply by coincidence. The awareness ratings above 1 simply reflect these trials, and it might even be possible that the participant experienced a vague illusion of perceiving a stimulus that is not the one presented. It is also possible that top-down influences on perception, rather than simply coincidence, bias the processing of a sensory stimulus in one direction (Bachmann, 2004). In this sense, the fact that awareness ratings are not all ones (i.e. no experience) when the participants are at chance seems logical if one assumes that the same sensory stimulus is not always processed in a same way, and part of the reason for such changes in processing could in principle be changes in either response bias or top-down bias in the perceptual system (causing illusory perception).

Importantly, this relative difference between the plateaus of the accuracy and awareness function should not be problematic for the analyses. As mentioned in "Materials and methods", both the parameters *c* and *d* are defined independently of *a* and *b*, and any change in *a* and/or *b* will not impact on *c* or *d*. This means, for instance, that the relatively higher upper plateau of the awareness function observed Experiment 2 compared to Experiment 1, could not have caused a shift of the curve and thus a smaller estimate of unconscious processing.

4.3. Awareness as graded or dichotomous

It has been a long-standing debate in consciousness research whether awareness is gradual or dichotomous. Based on Global workspace theory and data from attentional blink experiments, it has been proposed that awareness is dichotomous, at least in some instances (Sergent & Dehaene, 2004), yet other evidence from masking (Overgaard et al., 2006) and attentional blink (Nieuwenhuis & Kleijn, 2010) experiments provide evidence for a continuous transition between non-conscious and conscious perception. Recently, it has been suggested that non-dichotomous awareness reports (and the fact that they correlate well with accuracy) reflect dichotomous access to processing at different stages of the visual system (Kouider & Dupoux, 2004; Kouider et al., 2010). This fits well with experimental results showing that when a face stimulus is masked after a short SOA (stimulus onset asynchrony, i.e. the number of milliseconds between target and mask), blurred Gaussian noise displays (of similar power spectra) mask just as well as a blurred face of a different identity whereas after a long SOA, this is no longer the case (Bachmann, Luiga, & Põder, 2005). In the first case, the stimulus is masked already at early processing stages where only a very coarse representation (of spectral power) is present, and in the second case the stimulus is masked at a later processing stage where some configurational information is present (Bachmann, 2006).

Experiments such as the above-mentioned are, however, also compatible with the view that consciousness is a gradual/graded phenomenon. In this view, the clarity of the experience could be hypothesized to reflect how unambiguous the stimulus related neural activity is. There is no *a priori* reason why this evidence should be dichotomous as firing patterns and frequencies of neurons are by no means dichotomous. Even if one "strength of evidence" evaluation is somehow present at the final processing stage, there is little reason that this should be dichotomous. For these reasons, the discussion of whether conscious is gradual or dichotomous is still unresolved, and it is still unknown whether the partial awareness hypothesis is the best solution.

It is not a trivial matter designing an experiment to distinguish the two ways in which awareness can be partial (all-or-none access to each of a number of stages vs. incomplete/gradual access to one or more stages). The results of the present experiments do not support one version specifically, yet the results confirm the results of other statistical methods that there does seem to be some unconscious processing, and that the amount of unconscious processing varies across stimulus duration, and possibly that a large cause of effective unconscious processing is that awareness on average increases at a slower rate than accuracy. We did indeed find that awareness increases slower than accuracy (i.e. less dichotomously), causing the largest vertical gap between the functions (and the largest difference in our relative difference curves) to occur at around the stimulus durations for which the most unconscious processing is also found with subjective threshold approaches (see Fig. 1). This finding that two methods now indicate what might be termed a "window of unconscious perception" located at roughly the same stimulus intensities warrants further investigation in the field.

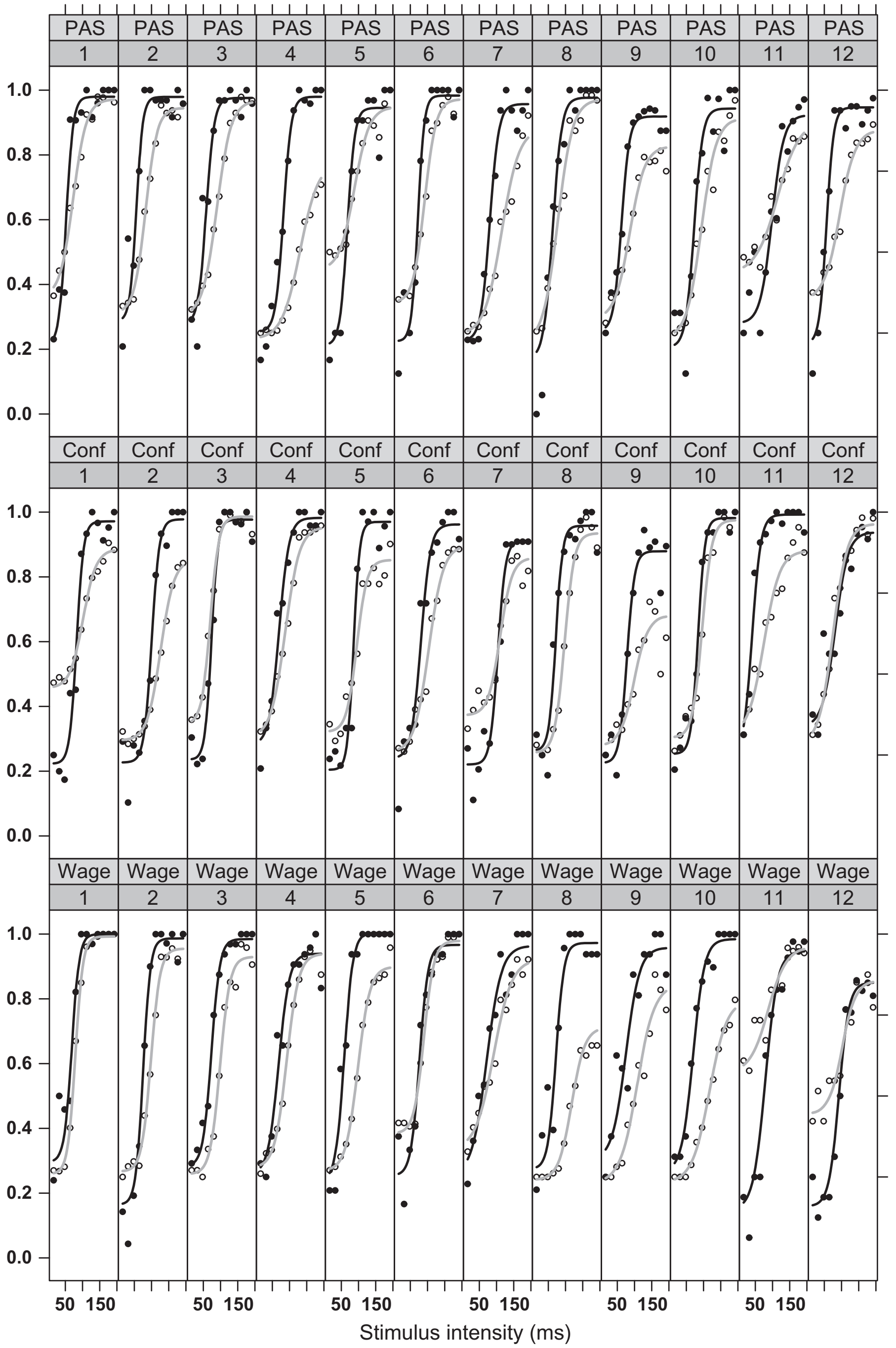

Fig. A1. Individual subject data from Experiment 1. Accuracy (black) and awareness ratings (gray) plotted as functions of stimulus duration for each subject. 4-Parameter sigmoid functions are fitted to the results.

4.4. 3-step versus 4-step PAS

The fact that the 3-step PAS performed very well, i.e. better than the 4-step PAS, when analysing by curve comparisons is interesting as the 4-step PAS was created by subjects who were asked to categorize their experience, and they all ended up using a 4-step scale. It was, however, suggested to them that they could start the scale with "no experience at all" and end it with "a clear image". This is likely to be a suitable instruction in most cases (when only analysing by the traditional criteria or when conducting neuroimaging experiments in which a difference between ratings 3 and 4 might be identified), but when analysing curve differences it would not ensure that the scale is also fully exhaustive at the high end of the scale. Based on the present study alone, we cannot claim that a 3-step PAS is always the appropriate choice of awareness scale. It is important to note, however, that a completely dichotomous scale would not be able to meet the preconditions of exhaustiveness and exclusiveness in both ends of the scale as long as awareness in itself is not dichotomous; a third step is necessary for instances in which conscious experience is neither completely absent nor clear enough to identify the stimulus.

5. Conclusion

In the reported masked visual identification experiments, we successfully employed nonlinear regression models in order to examine how task accuracy and reported awareness increase as sigmoid functions of stimulus duration. Unlike previously employed methods, the current method allowed us to base our estimates on a large range of stimuli and corresponding experiences, and thus not only very vague experiences. The curve estimations can further be used to predict an appropriate awareness level (with a confidence interval) for an obtained accuracy level following experimental manipulations. The constructed models described the data very well, also on an individual subject level. We argue that preconditions of maximum exhaustiveness and exclusiveness must be met in both ends of an awareness scale in order to obtain the most valid results. Although the analyses were made using data from visual consciousness experiments, they should in principle be applicable within other lines of research in which objective and subjective measures are compared.

Acknowledgments

This work was supported by a Starting Grant, European Research Council (K.S. and M.O.) and Marie Curie Action IEF #237502 "Social Brain" (B.T.). A.C. is a Research Director with the National Fund for Scientific Research (FRS – FNRS, Belgium).

Appendix A.

4-Parameter sigmoid functions were fitted to the data of the individual subjects. The curves for each subject are presented here (see Figs. A1 and A3). Also, the accuracies corresponding to each awareness rating are plotted for the 3-step PAS compared to the 4-step PAS in Fig. A2.

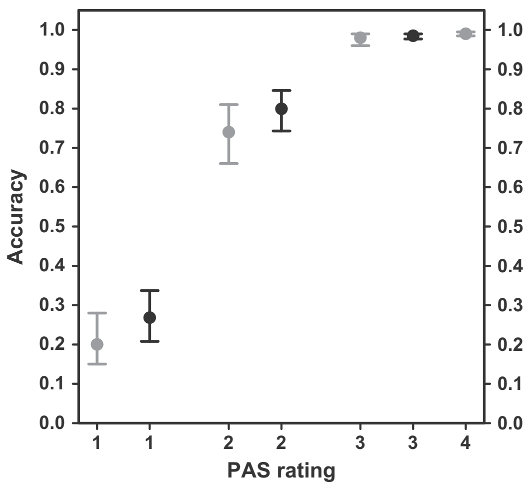

Fig. A2. Accuracy for each awareness rating. The plot shows the accuracies corresponding to each awareness rating for the 3-step PAS (black dots) compared to the 4-step PAS (gray dots). Error bars indicate 95% confidence intervals.

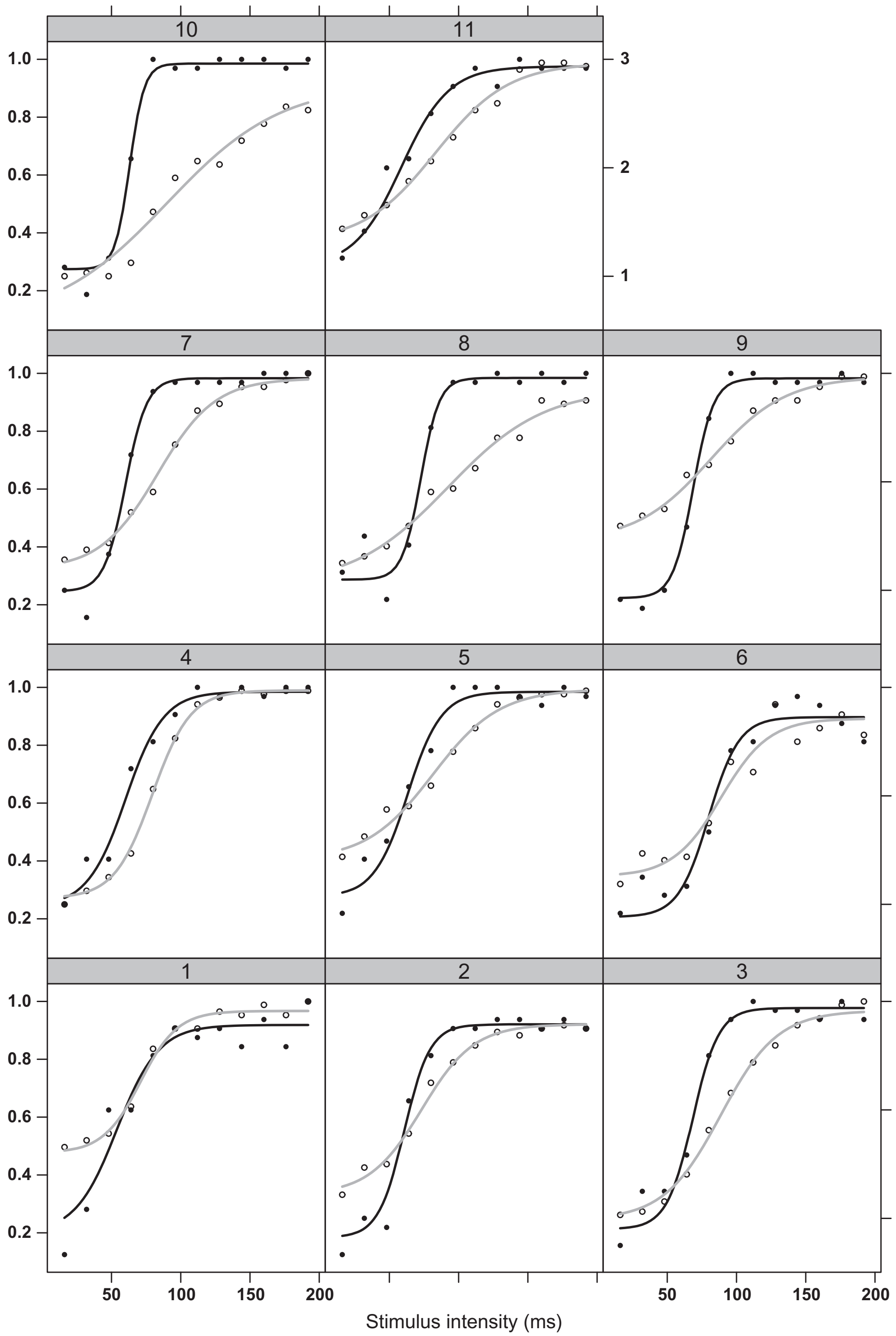

Fig. A3. Individual subject data from Experiment 2. Accuracy (black) and awareness ratings (gray) plotted as functions of stimulus duration for each subject. 4-Parameter sigmoid functions are fitted to the results.

References

Bachmann, T. (2004). Inaptitude of the signal detection theory, useful vexation from the microgenetic view, and inevitability of neurobiological signatures in understanding perceptual (un)awareness. *Consciousness and Cognition, 13*(1), 101–106. doi:10.1016/j.concog.2003.10.005.

Bachmann, T. (2006). Microgenesis of perception: Conceptual, Psychophysical, and Neurobiological Aspects. In *The first half second: The microgenesis and temporal dynamics of unconscious and conscious visual processes* (pp. 11–34). Cambridge, MA: MIT Press.

Bachmann, T., Luiga, I., & Põder, E. (2005). Variations in backward masking with different masking stimuli: II. The effects of spatially quantised masks in the light of local contour interaction, interchannel inhibition, perceptual retouch, and substitution theories. *Perception, 34*(2), 139–154. doi:10.1068/p5344b.

Boyer, J. L., Harrison, S., & Ro, T. (2005). Unconscious processing of orientation and color without primary visual cortex. *Proceedings of the National Academy of Sciences of the United States of America, 102*(46), 16875–16879. doi:10.1073/pnas.0505332102.

Chan, C. (1992). *Implicit cognitive processes: Theoretical issues and applications in computer systems design*. Unpublished doctoral dissertation, University of Oxford, Oxford, England.

Cheesman, J., & Merikle, P. M. (1984). Priming with and without awareness. *Perception & Psychophysics, 36*(4), 387–395.

Cheesman, J., & Merikle, P. M. (1986). Distinguishing conscious from unconscious perceptual processes. *Canadian Journal of Psychology Revue Canadienne de Psychologie, 40*(4), 343–367. doi:10.1037/h0080103.

Curran, T., & Hintzman, D. L. (1995). Violations of the independence assumption in process dissociation. *Journal of Experimental Psychology: Learning, Memory and Cognition, 21*(3), 531–547.

Debner, J. A., & Jacoby, L. L. (1994). Unconscious perception: Attention, awareness, and control. *Journal of Experimental Psychology: Learning, Memory and Cognition, 20*(2), 304–317.

Destrebecqz, A., & Cleeremans, A. (2001). Can sequence learning be implicit? New evidence with the process dissociation procedure. *Psychonomic Bulletin & Review, 8*(2), 343–350.

Dienes, Z. (2008). Subjective measures of unconscious knowledge. *Progress in Brain Research, 168*, 49–64. doi:10.1016/S0079-6123(07)68005-4.

Dienes, Z., Altmann, G. T. M., Kwan, L., & Goode, A. (1995). Unconscious knowledge of artificial grammars is applied strategically. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(5), 1322–1338. doi:10.1037/0278-7393.21.5.1322.

Dienes, Z., & Seth, A. (2009). Gambling on the unconscious: A comparison of wagering and confidence ratings as measures of awareness in an artificial grammar task. *Consciousness and Cognition, 19*(2), 674–681. doi:10.1016/j.concog.2009.09.009.

Dienes, Z., & Seth, A. K. (2010). Measuring any conscious content versus measuring the relevant conscious content: Comment on Sandberg et al.. *Consciousness and Cognition, 19*(4), 1079–1080. doi:10.1016/j.concog.2010.03.009.

Dodson, C. S., & Johnson, M. K. (1996). Some problems with the process-dissociation approach to memory. *Journal of Experimental Psychology: General, 125*(2), 181–194.

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

Graf, P., & Komatsu, S.-I. (1994). Process dissociation procedure: Handle with caution! *European Journal of Cognitive Psychology, 6*(2), 113–129. doi:10.1080/09541449408520139.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language, 30*(5), 513–541. doi:10.1016/0749-596X(91)90025-F.

Jacoby, L. L. (1998). Invariance in automatic influences of memory: Toward a user's guide for the process-dissociation procedure. *Journal of Experimental Psychology: Learning, Memory and Cognition, 24*(1), 3–26.

Jacoby, L. L., Toth, J., & Yonelinas, A. P. (1993). Separating conscious and unconscious influences of memory: Measuring recollection. *Journal of Experimental Psychology: General, 122*(2), 139–154. doi:10.1037/0096-3445.122.2.139.

Jiménez, L., Méndez, C., & Cleeremans, Axel. (1996). Comparing direct and indirect measures of sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*(4), 948–969. doi:10.1037/0278-7393.22.4.948.

Kanai, R., Walsh, V., & Tseng, C.-H. (2010). Subjective discriminability of invisibility: A framework for distinguishing perceptual and attentional failures of awareness. *Consciousness and Cognition, 19*(4), 1045–1057. doi:10.1016/j.concog.2010.06.003.

Koch, C., & Preuschoff, K. (2007). Betting the house on consciousness. *Nature Neuroscience, 10*(2), 140–141. doi:10.1038/nn0207-140.

Koivisto, M., Mäntylä, T., & Silvanto, J. (2010). The role of early visual cortex (V1/V2) in conscious and unconscious visual perception. *NeuroImage, 51*(2), 828–834. doi:10.1016/j.neuroimage.2010.02.042.

Kouider, S., de Gardelle, V., Sackur, J., & Dupoux, E. (2010). How rich is consciousness? The partial awareness hypothesis. *Trends in Cognitive Sciences, 14*(7), 301–307. doi:10.1016/j.tics.2010.04.006.

Kouider, S., & Dupoux, E. (2004). Partial awareness creates the "illusion" of subliminal semantic priming. *Psychological Science: A Journal of the American Psychological Society/APS, 15*(2), 75–81.

Kuhn, G., & Dienes, Z. (2006). Differences in the types of musical regularity learnt in incidental- and intentional-learning conditions. *Quarterly Journal of Experimental Psychology, 59*(10), 1725–1744. doi:10.1080/17470210500438361.

Merikle, P. M. (1982). Unconscious perception revisited. *Perception & Psychophysics, 31*(3), 298–301.

Merikle, P. M., & Daneman, M. (2000). Conscious vs. unconscious perception. In M. Gazzaniga (Ed.), *The new cognitive neurosciences* (2nd ed., pp. 1295–1303). Cambridge, MA: MIT Press.

Merikle, P. M., & Joordens, S. (1997). Measuring unconscious influences. In J. Cohen & J. Schooler (Eds.), *Scientific approaches to consciousness* (pp. 109–123). Mahwah, NJ: Erlbaum.

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

Nieuwenhuis, S., & Kleijn, R. (2010). Consciousness of targets during the attentional blink: A gradual or all-or-none dimension? *Attention, Perception, & Psychophysics, 73*(2), 364–373. doi:10.3758/s13414-010-0026-1.

Overgaard, M. (2006). Introspection in science. *Consciousness and Cognition, 15*(4), 629–633. doi:10.1016/j.concog.2006.10.004.

Overgaard, M., Gallagher, S., & Ramsøy, T. Z. (2008). An integration of first-person methodologies in cognitive science. *Journal of Consciousness Studies, 15*(5), 100–120.

Overgaard, M., Rote, J., Mouridsen, K., & Ramsøy, 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. doi:10.1016/j.concog.2006.04.002.

Persaud, N., & McLeod, P. (2008). Wagering demonstrates subconscious processing in a binary exclusion task. *Consciousness and Cognition, 17*(3), 565–575. doi:10.1016/j.concog.2007.05.003.

Persaud, N., McLeod, P., & Cowey, A. (2007). Post-decision wagering objectively measures awareness. *Nature Neuroscience, 10*(2), 257–261. doi:10.1038/nn1840.

Peterson, W., Birdsall, T., & Fox, W. (1954). The theory of signal detectability. *IEEE Transactions on Information Theory, 4*(4), 171–212. doi:10.1109/TIT.1954.1057460.

Ramsøy, T. Z., & Overgaard, M. (2004). Introspection and subliminal perception. *Phenomenology and the Cognitive Sciences, 3*(1), 1–23. doi:10.1023/B:PHEN.0000041900.30172.e8.

Reingold, E. M., & Merikle, P. M. (1988). Using direct and indirect measures to study perception without awareness. *Perception & Psychophysics, 44*(6), 563–575.

Reingold, E. M., & Merikle, P. M. (1990). On the inter-relatedness of theory and measurement in the study of unconscious processes. *Mind & Language, 5*(1), 9–28. doi:10.1111/j.1468-0017.1990.tb00150.x.

Sandberg, K., Timmermans, B., Overgaard, M., & Cleeremans, A. (2010). Measuring consciousness: Is one measure better than the other? *Consciousness and Cognition, 19*(4), 1069–1078. doi:10.1016/j.concog.2009.12.013.

Schurger, A., & Sher, S. (2008). Awareness, loss aversion, and post-decision wagering. *Trends in Cognitive Sciences, 12*(6), 209–210 [author reply 210. doi:10.1016/j.tics.2008.02.012].

Scott, R. B., & Dienes, Z. (2008). The conscious, the unconscious, and familiarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(5), 1264–1288. doi:10.1037/a0012943.

Sergent, C., & Dehaene, S. (2004). Is consciousness a gradual phenomenon? Evidence for an all-or-none bifurcation during the attentional blink. *Psychological Science: A Journal of the American Psychological Society/APS, 15*(11), 720–728. doi:10.1111/j.0956-7976.2004.00748.x.

Sidis, B. (1898). *The psychology of suggestion*. New York: Appleton.

Snodgrass, M. (2002). Disambiguating conscious and unconscious influences: Do exclusion paradigms demonstrate unconscious perception? *The American Journal of Psychology, 115*(4), 545–579.

Snodgrass, M., & Shevrin, H. (2006). Unconscious inhibition and facilitation at the objective detection threshold: Replicable and qualitatively different unconscious perceptual effects. *Cognition, 101*(1), 43–79. doi:10.1016/j.cognition.2005.06.006.

Tanner, W. P., Jr., & Swets, J. A. (1954). A decision-making theory of visual detection. *Psychological Review, 61*(6), 401–409. doi:10.1037/h0058700.

Timmermans, B., Sandberg, K., Cleeremans, A., & Overgaard, M. (2010). Partial awareness distinguishes between measuring conscious perception and conscious content: Reply to Dienes and Seth. *Consciousness and Cognition, 19*(4), 1081–1083. doi:10.1016/j.concog.2010.05.006.

Toth, J., Reingold, E. M., & Jacoby, L. L. (1995). A response to graf and komatsu's critique of the process dissociation procedure: When is caution necessary? *European Journal of Cognitive Psychology, 7*(2), 113–130. doi:10.1080/09541449508403095.

Visser, T. A., & Merikle, P. M. (1999). Conscious and unconscious processes: The effects of motivation. *Consciousness and Cognition, 8*(1), 94–113. doi:10.1006/ccog.1998.0378.