Subjective discriminability of invisibility: A framework for distinguishing perceptual and attentional failures of awareness

Ryota Kanai a,*, Vincent Walsh a, Chia-huei Tseng b,c

a Institute of Cognitive Neuroscience & Dept. of Psychology, University College London, 17 Queen Square, WC1N 3AR London, United Kingdom
b Department of Psychology, National Taiwan University, Taipei, Taiwan
c Department of Psychology, The University of Hong Kong, Hong Kong

Abstract

Conscious visual perception can fail in many circumstances. However, little is known about the causes and processes leading to failures of visual awareness. In this study, we introduce a new signal detection measure termed subjective discriminability of invisibility (SDI) that allows one to distinguish between subjective blindness due to reduction of sensory signals or to lack of attentional access to sensory signals. The SDI is computed based upon subjective confidence in reporting the absence of a target (i.e., miss and correct rejection trials). Using this new measure, we found that target misses were subjectively indistinguishable from physical absence when contrast reduction, backward masking and flash suppression were used, whereas confidence was appropriately modulated when dual task, attentional blink and spatial uncertainty methods were employed. These results show that failure of visual perception can be identified as either a result of perceptual or attentional blindness depending on the circumstances under which visual awareness was impaired.

1. Introduction

Conscious visual perception can be impaired in many circumstances such as when a visual stimulus is degraded or when we are distracted by other stimuli (Kim & Blake, 2005). Such instances of stimulus blindness have revealed that visual stimuli are processed in many brain regions even when they do not reach visual awareness (Lin & He, 2009). However, the causes and processes leading to failure of visual awareness are likely to be different depending on the circumstances under which visual awareness is impaired. Disruption of visual awareness could occur when stimulus signals are degraded to the degree that they are hardly indistinguishable from any signal at all (Dehaene, Changeux, Nacache, Sackur, & Sergant, 2006; Lamme, 2003, 2004). Even if sensory signals remain strong, they could go unnoticed when attention was distracted by other stimuli (Chun & Potter, 1995; Dehaene et al., 2006; Norman & Bobrow, 1975; Pashler, 1994; Sigman & Dehaene, 2006) (Fig. 1). In other words, psychophysically induced unawareness seems to be divided into perceptual blindness, in which subjective invisibility is caused by suppression of low-level sensory signals, and attentional blindness, in which observers fail to notice the presence of a target due to failure to access low-level sensory signals despite their presence. However, it has been difficult to operationally distinguish these causes of blindness. Yet, the distinction of the two is critical for understanding the neuronal mechanisms underlying conscious visual perception (Block, 2007; Lamme, 2006).

The two levels of awareness failure have also been proposed on the basis of neuroimaging studies of conscious face perception (Block, 2007). In a study with binocular rivalry with faces and houses as stimuli, the activity in the fusiform face area (FFA) was suppressed when conscious perception of a face was impaired by a competing stimulus presented to the other eye.
On the other hand, a visual extinction patient exhibited FFA activity even when he failed to consciously register the face presented in his left visual field due to a competing stimulus on the right (Rees et al., 2000). The former case of binocular rivalry can be interpreted as suppression of visual signals in FFA or even at earlier stages, whereas the latter case appears to be due to lack of attentional (or cognitive) access to the information of the face present in the FFA.

Despite the mounting evidence, it has been difficult to operationally distinguish at a behavioural level the suppression of low-level sensory signals and failure to access sensory signals, because they both result in the same responses, i.e., reports of absence of stimuli. Moreover, the fact that attention often boosts sensory signals independent of awareness (Koch & Tsuchiya, 2007) makes it difficult to separate the effect of attentional access from modulation of sensory signals.

The goal of the present study was to introduce a framework to distinguish perceptual and attentional blindness and examine the framework in light of empirical data from a series of psychophysical experiments. To illustrate the idea behind the development of such a framework, we first provide an overview of signal detection theory (SDT) (Green & Swets, 1966; Macmillan & Creelman, 2005) as applied in awareness studies and discuss their limitations.

### 1.1. Objective sensitivity: Type I signal detection theory

The most common use of SDT in experimental psychology is a Type I task in which observers make decisions in a presence–absence judgment task or in a stimulus discrimination task. In Type I SDT, objective sensitivity independent of decision criterion is computed. In a presence–absence judgment task, for example, observers make judgments as to whether a stimulus is present or absent. Trials in a presence–absence judgment task are categorized into the following four types:

1. **Type I hits**: stimulus present & response present.
2. **Type I correct rejections**: stimulus absent & response absent.
3. **Type I misses**: stimulus present & response absent.
4. **Type I false alarms**: stimulus absent & response present.

From the z-scores of hit and false alarm rates (i.e., hits as a proportion of present trials and false alarms as a proportion of absent trials), sensitivity measures such as $d'$ or the AUC (the Area Under the ROC curve) are computed and they are used as criterion free measures of sensitivity (see a standard textbook on SDT for the formulas of these measures; e.g., Green & Swets, 1966; Macmillan & Creelman, 2005).

Null sensitivity to a stimulus (i.e., $d' \sim 0$, misses and false alarms make up 50% of responses) has been used as evidence of no awareness under experimental situations (e.g., Dagenbach, Carr, & Wilhelmsen, 1989; Eriksen, 1960; Fowler, Wolford, Slade, & Tassinary, 1981; Kemp-Wheeler & Hill, 1988; Marcel, 1983; Nolan & Caramazza, 1982; Schurger, Pereira, Tresiman, ...
& Cohen, 2009; Schwiedrzik, Singer, & Melloni, 2009). The advantage of using objective Type I sensitivity as a marker for absence of awareness is that it does not involve subjective estimates of awareness and thus provides a conservative criterion when experimenters wish to ensure that stimuli evoked no awareness (Cheesman & Merikle, 1984).

However, this criterion may be too stringent to capture the nature of awareness in some cases. For example, blindsight patients with a lesion to the visual cortex can detect visual stimuli above a chance level while they claim to have no subjective experiences of the stimuli (Stoerig & Cowey, 1997; Weiskrantz, 1998). Similar situations have been reported in normal healthy observers under various experimental conditions (Kolb & Braun, 1995; Lau & Passingham, 2006). Objective sensitivity as measured by Type I SDT would indicate that observers are aware of the stimuli in these situations, because they could detect the presence of a stimulus. However, this is in contradiction with the subjective reports of absence of awareness. Thus, the objective sensitivity (e.g., \( d' = 0 \)) is insufficient to capture possible dissociations between objective performance and subjective report of conscious percepts.

1.2. Subjective sensitivity: Type II signal detection analysis

One way to overcome the limitation of the objective criterion of awareness is to collect introspective reports of observers’ confidence while performing an objective Type I task. In the classical Type II task, observers are asked to report their confidence of their correctness in an objective Type I task and the relationship between objective correctness and subjective confidence is analyzed (Fig. 1B). The idea is that if observers were aware of a stimulus, their correct responses would be accompanied by high confidence, while low confidence would be linked with incorrect trials. Thus, if observers were aware of stimuli, they would be able to adjust their confidence according to correctness of their responses.

Performance in a Type II task – observers’ ability to distinguish their objective correctness based on confidence ratings – is calculated from the following four types of second-level responses:

1. Type II hits: correct response & high confidence.
2. Type II correct rejections: incorrect response & low confidence.
3. Type II misses: correct response & low confidence.
4. Type II false alarms: incorrect response & high confidence.

From the proportions of Type II Hits and Type II False Alarms, criterion-free sensitivity measures such as \( d' \) or AUC can be computed (Galvin, Podd, Drga, & Whitmore, 2003; Kunimoto, Miller, & Pashler, 2001; Pollack & Decker, 1958), and Type II performance has been used as a criterion for conscious awareness (Kolb & Braun, 1995; Kunimoto et al., 2001; Persaud, McLeod, & Cowey, 2007; Szczepanowski & Pessoa, 2007). The idea is that if Type I performance is above chance, but Type II performance is at chance, the observers cannot access the sensory signals used for the Type I task. The recently proposed post-decision wagering (PDW) is a variant of Type II task in which observers are asked to wager on their Type I decision (Persaud et al., 2007). The PDW has been successfully applied to capture a variety of dissociations between objective performance and subjective awareness including blindsight (Persaud et al., 2007; but see Clifford, Arabzadeh, & Harris, 2008; Dienes & Seth, 2010).

While this approach has been successful in capturing the subjective nature of conscious awareness, the classical Type II task is not suitable for probing the different mechanisms underlying miss trials. As we argue in the next section, it is essential to keep separate the four trial types (i.e., hits, correct rejections, misses and false alarms) without aggregating them as correct (hits and correct rejections) and incorrect trials (misses and false alarms). However, the classical Type II tasks reported to date always combined hits and correct rejections as correct trials and misses and false alarms as incorrect trials. This aggregation does not allow separate analyses on the confidences of misses from those of false alarms, despite their potentially very different causes of errors. Moreover, most studies in the past employed a discrimination task between orientations (Wilming, Tsuchiya, Fahle, Einhäuser, & Koch, 2008), shapes (Lau & Passingham, 2006) or locations (Szczepanowski & Pessoa, 2007). In such cases, hits and correct rejections were interchangeable and were thus arbitrarily assigned to one response or to the other.

1.3. Dissociation of perceptual blindness and attentional blindness

To dissociate perceptual blindness and attentional blindness, the four trial types need to be kept separate instead of aggregating them as correct and incorrect trials. This is because perceptual and attentional blindness make distinct predictions as to the confidence rating behaviors for trials when they reported absence of a stimulus (i.e., misses and correct rejections). In perceptual blindness, sensory signals are already weak before they reach the decision stage and are hardly distinguishable from no signals at all. Thus, if observers tried to access the sensory signals, they would feel confident that there was no stimulus. This predicts that for perceptual blindness, observers cannot discriminate misses and correct rejections when they reported absence of a stimulus (Fig. 1c). In attentional blindness, on the other hand, the process of accessing sensory signals is disrupted due to psychophysical manipulations such as attentional distraction by a concurrent task. In such cases, observers would be able to adjust their confidence depending upon the availability of attentional resources at the time of stimulus detection. This predicts that for attentional blindness, observers will be able to discriminate misses and correct rejections by their confidence ratings.

As an illustration, consider disappearances of salient stimuli in motion-induced blindness or interocular suppression as in binocular rivalry. In these psychophysical phenomena, we feel that stimuli are invisible. This might result in high confidence...
in reporting absence of a stimulus if one was not informed of the fact that illusory disappearances could be induced by stimulus manipulations. On the other hand, when observers missed a target due to attentional distraction, they might have a feeling that they have missed a target because they could not pay full attention to the task. In the former case, observers would not be able to lower their confidence in reporting absence of stimuli, whereas in the latter case, they would be able to lower their confidence in accordance with their performance. In this study, we examine whether the differences in subjective impression of this sort could be captured by comparing confidence ratings for miss trials.

To formally quantify observers’ ability to adjust their confidence in psychophysically induced failures of awareness, we introduce an index termed subjective discriminability of invisibility (SDI). The SDI was developed based on Type II analysis (Galvin et al., 2003; Pollack & Decker, 1958), but computes Type II performance only for trials in which observers reported absence of a stimulus, namely, misses and correct rejections (Fig. 1c). This measure quantifies how confidence ratings for trials with impaired awareness (miss trials) differ from those with no target (correct rejections). Based on the rationale described earlier, we predict that SDI performance would fall to a chance level for perceptual blindness but remain above chance for attentional blindness.

The goal of the present study is to examine the new measure, SDI, in various circumstances under which visual awareness is impaired (Fig. 2). To do so, we applied the measure of the SDI while suppressing subjective visibility of a stimulus by contrast reduction, backward masking (Breitmeyer & Ogmen, 2006), flash suppression (Wolfe, 1984), dual task (Lee, Itti, Koch, & Braun, 1999), attentional blink (Raymond, Shapiro, & Arnell, 1992) and increase in spatial uncertainty (Foley & Schwarz, 1998) (see Section 2 for full details of the experiments). In all the six tasks, the target was presented on only 50% of the trials and observers were asked to report the presence or absence of a target stimulus together with their confidence rating (high, mid or low).

2. Materials and methods

2.1. Stimuli and procedure

Visual stimuli were presented on a 19-inch CRT monitor using Psychtoolbox 3 running on Matlab (Brainard, 1997; Pelli, 1997). In all the experiments, the task was to report the presence or absence of the target probe together with their
confidence rating in three levels; low (completely guessing), medium and high (certain). The stimulus parameters for each experiment are summarized below. In all the experiments, the target was presented on half of the trials and was absent in the remaining half.

2.1.1. Contrast
Six naive observers participated in the luminance probe detection task. The contrast of the target was varied between 5.0%, 6.5%, 8.0%, 9.5% and 11.0%. The sigma of the Gaussian probe was 0.22°. The target was presented for 24 ms below the fixation. The distance between the fixation and the centre of the target was 4.38°. One block of trials consisted of 220 trials. The timing of the target onset was varied between 1 and 3 s from the beginning of a trial. The moment in which the target was potentially presented at a probability of 50% was indicated by a brief colour change of the fixation cross from black to red. Within a block, there were 20 target-present trials for each contrast level and an equal number of target-absent trials. We included fixation control trials in which the subjects had to respond to the colour change of the fixation cross (9% of trials). Each participant completed a minimum of three blocks (i.e., 60 target-present trials per contrast) and on average 3.67 blocks.

2.1.2. Backward masking
Six naive observers participated in this experiment. The target probe was identical to the contrast experiment above, but the contrast was fixed at 40%, which is well above the detection threshold when isolated alone without a mask. The mask consisted of black and white hollow square frames each subtending 0.57° × 0.57° (see Fig. 2B). The detection performance was manipulated by adjusting the inter-stimulus interval between the probe and the mask between 12 ms, 24 ms, 35 ms, 47 ms, 94 ms and 188 ms. The mask was presented for 253 ms. As in the contrast experiment, we included fixation control trials in which the subjects had to respond to the colour change of the fixation cross. Within a block, there were a total of 132 trials. There were 10 target-present trials per ISI and a total of 60 target-absent trials and 12 fixation control trials.

2.1.3. Flash suppression
Six observers participated in this experiment. Stimuli were orthogonal gratings (45° and 135° in orientation), which were presented to each eye. They were presented with a temporal offset, i.e., the second stimulus was presented 1470 ms after the onset of the first stimulus. This temporal offset is known to be effective at flipping the percept to a new stimulus (flash suppression) (Wolfe, 1984). A probe was presented 365 ms after the onset of the flash suppression either to the suppressed eye or to the dominant eye. The probe was a Gaussian-enveloped (sigma = 106 ms) luminance blob superimposed on one of the gratings. The sigma of the spatial Gaussian envelope was 0.22° and its peak was 1.32° below fixation. Within a block, the combinations of orientation and eye (left-eye 45° or right-eye 45°), and the stimulus presentation order (left eye first or right-eye first) were counterbalanced across trials. The target-present trials across the four counterbalanced conditions were collapsed into suppressed eye trials and dominant eye trials, depending on whether the probe was presented either to the dominant eye (i.e., the eye to which the second stimulus was presented) or to the suppressed eye (i.e., the eye to which the first stimulus was presented). On half of trials, the probe was absent. Within a block, there were 12 dominant eye trials, 12 suppressed eye trials and 24 target-absent trials. All subjects completed a minimum of three blocks and on average 3.5 blocks.

2.1.4. Dual task
Six observers participated in this experiment. The target probe was the luminance Gaussian blob used in the contrast experiment and its contrast was fixed at 30%, well above all observers’ detection thresholds. In the concurrent visual search task, subjects were required to report the presence or absence of the target letter ‘T’ among distractor ‘L’s. The letters were randomly oriented and were presented at an eccentricity of 2° for 58 ms. After a blank period of 177 ms, F-shaped masks, which overlapped with both ‘T’s and ‘L’s, were presented for 177 ms. The Gaussian probe was presented during the last 24 ms of the letter array presentation. The number of the items in the search array was varied between 1, 2, 4 and 8 across trials. Within a block, there was a total of 160 trials. There were 20 target-present trials per number of items in the search array and a total of 80 target-absent trials were randomly mixed across trials. All subjects completed three blocks of trials (i.e., 480 trials in total).

2.1.5. Attentional blink
Six observers participated in the attentional blink experiment. For the rapid serial visual presentation (RSVP), all the alphabet except for the letters ‘I’, ‘O’, and ‘W’ were used and the same letter was not repeated in an RSVP. The letters were presented in black with the 80 point bold Helvetica font, and subtended 2.8° vertically. On one trial, 16 letters were presented in the RSVP. The task was to report the identity of the first target (T1) indicated by the red colour and to report the presence or absence of the second target, a black letter ‘X’ (T2) in the RSVP stream together with confidence rating for the T2 report. T1 was presented at either the 4th, 6th or 8th temporal position in the RSVP. For each position, T1–T2 lags of 2, 3, 4, 5, 6, and 7 were tested for equal numbers of trials. Data from different T1 positions were collapsed. 120 trials per lag were tested per observer.
2.1.6. Spatial uncertainty

Six observers participated in the spatial uncertainty experiment. The probe was the luminance Gaussian blob as in the contrast and backward masking experiments and was presented at one of the eight positions at an eccentricity of 6.58°. The peak luminance of the probe was 2.5 times the 79.4% correct detection threshold for each target location, which was determined for an individual observer by a 1-up-3-down staircase with 2-interval forced choice before the main experiment. At the beginning of a trial, either a single predictive cue (a white line subtending 4.38° from the fixation point) or non-predictive cues pointing all the eight possible target positions were presented for 706 ms. After 353 ms, the probe was presented at one of the eight locations on 50% of trials. When only one cue was presented, the probe was always presented at the location indicated by the cue. The probe was presented for an equal number of trials in each of the eight positions. Observers completed a minimum of 128 trials (8 [probe positions] × 2 [levels of cue number] × 2 [present and absent trials]). The data from different probe positions were pooled, resulting in a minimum of 32 trials per condition.

2.2. Analysis

2.2.1. Objective detection sensitivity

In order to quantify objective detection sensitivity, we constructed an ROC curve by plotting hit rates against false alarm rates. Two inflection points were included in the ROC curve using the confidence data (Green & Swets, 1966; Macmillan & Creelman, 2005). The objective detection performance was quantified as the area under the ROC curve.

2.2.2. Subjective discriminability of invisibility

We constructed a second-level ROC curve for trials in which subjects reported absence of a target (i.e., miss trials and correct rejection trials) with two inflection points. The rate of high-confidence correct rejections (analogous to hits in the first level ROC) was plotted against the rate of high-confidence miss trials (analogous to false alarm in the first level ROC). The SDI was quantified as the AUC of the second-level ROC curve (Szczepanowski & Pessoa, 2007; Wilmzig et al., 2008).

3. Results

3.1. Objective performance

Objective performances for target detection were impaired by the relevant manipulation for each task (Fig. 3). In the contrast reduction experiment (Fig. 3A), the detection performance decreased with decreasing the contrast (Spearman’s $\rho = 0.57$, $p < .01$). In the backward masking experiment (Fig. 3B), the objective performance monotonically increased as the inter-stimulus interval (ISI) between the target and the mask increased ($\rho = 0.72$, $p < .001$). In the flash suppression experiment (Fig. 3C), the detection performance for a probe presented to a suppressed eye was significantly lower than that for a probe presented to a dominant eye ($t(5) = 5.32$, $p < .01$). In the dual task experiment (Fig. 3D), the detection performance decreased as the number of items in the search array increased ($\rho = -0.57$, $p < .01$). In the attentional blink experiment (Fig. 3E), the detection performance was lower when the lag between T1 and T2 was short ($\rho = 0.57$, $p < .01$). Finally, in the spatial uncertainty experiment, the objective performance decreased when potential target positions was uncertain (Fig. 3F; $t(5) = 3.47$, $p < .05$). These results confirm that we obtained expected performance deteriorations for all the six task types.

3.2. Subjective discriminability of invisibility

To compare the SDI performance across different experimental manipulations, we selected the condition for each of the six tasks that yielded a comparable level of sensitivity reduction in objective performance (AUC = 0.7) (indicated by the gray stripes in Fig. 3). To illustrate the process of computing the SDI values, the histograms of confidence ratings for miss and correct rejection trials are shown for the visibility suppression conditions for the six tasks (Fig. 4). ROC curves were constructed based on the confidence histograms by plotting the cumulative probability function of the confidence in correct rejection trials against the cumulative probability function of the confidence in miss trials (Fig. 5).

We tested whether SDI was statistically greater than chance (one-tailed $t$-test against 0.5) for the six tasks (Fig. 6). The SDI was not significantly greater when the target visibility was suppressed in the contrast ($t(5) = -0.393$, $p = .64$), backward masking ($t(5) = 0.13$, $p = .45$) and flash suppression ($t(5) = -0.04$, $p = .51$) conditions. By contrast, it was significantly greater in the dual task ($t(5) = 2.37$, $p < .05$), attentional blink ($t(5) = 2.21$, $p < .05$) and spatial uncertainty ($t(5) = 2.60$, $p < .05$) experiments. These statistical analyses support the notion that the former three types of visual phenomena lead to subjective blindness indistinguishable from physical absence, whereas observers can distinguish the invisibility induced by the latter three manipulations from physical absence.

3.3. Standard Type II performance

To compare SDI measures with a more traditional approach, we computed AUC for the standard Type II performance (see, Fig. 1b). For all the six manipulations, the Type II AUC performance was significantly higher than chance including the three manipulations that were classified as perceptual blindness by SDI. This is due to the trials in which observers reported
presence of a target stimulus (i.e., hits and false alarms). Given that observers would be more confident when they detect a target (i.e., hits) than when they simply guessed (false alarms), the above-chance Type II performance is not surprising. In other words, the aggregation of different types of correct trials (hits and correct rejections) and incorrect trials (misses and false alarms) obscures the nature of the failures specifically involved in a target miss, because observers performed a Type II task well for the response-present trials. This comparison demonstrates the point that SDI is preferable to a more conventional Type II analysis for dissociating perceptual and attentional blindness.

3.4. Subjective discriminability of invisibility (full version)

Finally, we computed the SDI for all the conditions in the six tasks. We observed decreases in the SDI performance as objective visibility was reduced by reducing the contrast (Fig. 7A), backward masking (Fig. 7B) and by flash suppression (Fig. 7C). On the other hand, the SDI remained above chance even when the objective performances were decreased by dual task (Fig. 7D), attentional blink (Fig. 7E) or spatial uncertainty (Fig. 7F). These patterns of results corroborate the finding that unawareness due to contrast reduction, backward masking and flash suppression are categorized as perceptual blindness, whereas unawareness due to dual task, attentional blink and spatial uncertainty are categorized as attentional blindness. Interestingly, the SDI performances for the tasks categorized as perceptual blindness (i.e., contrast reduction, backward masking and flash suppression) were above chance when the objective sensitivity was high. For example, the SDI for high contrast stimuli (Fig. 7A) is classified as attentional blindness in our definition, as the value remained well above chance. This implies that when observers missed a high contrast target, they are aware that somehow they were not paying enough attention on those trials and reduced their confidence accordingly. Thus, observers could attribute the reason for missing a target of high contrast stimuli to some internal (attentional) state, whereas they could not subjectively distinguish a missed low contrast target from no stimulus. One important implication of this finding is that the distinction between perceptual and attentional blindness is not simply attributable to the design of a psychophysical experiment, but depends on the level of reduction in objective visibility.

4. Discussion

Our present results show that impairments of objective detection performance due to contrast reduction, backward masking and flash suppression were accompanied by a decline of the subjective discriminability of invisibility (SDI) to
chance performance, whereas the SDI for the dual task, attentional blink and spatial uncertainty experiments remained above chance even when their objective performances declined by a comparable degree.

We refer to the former cases as *perceptual blindness* and the latter cases as *attentional blindness*. The distinction between these two types of blindness supports the idea that suppression of visual awareness takes place as a consequence of reduction in the sensory signal level and/or failure to exert attentional access to the sensory signals. Our findings suggest that different types of visibility manipulation differentially interfere with those processes required for conscious report.

4.1. Comparison with Type II performance

To compare our results with conventional Type II performance, we computed subjective AUC for the Type II task by combining hits and correct rejections as correct trials and misses and false alarm rates as incorrect trials. In all the six tasks, the Type II performance was significantly above chance (Fig. 6B). Thus, the classical Type II performance did not differentiate perceptual and attentional blindness. The above-chance Type II performance was due to the general tendency that confidence ratings are low for false alarm trials and high in many hit trials. Thus, the contamination of the trials in which observers reported the presence of a target obscured the qualitative difference between perceptual and attentional blindness. These results validate the point that only report-absent trials should be included in the SDI analysis to quantify observers' ability to distinguish induced invisibility from physical absence of a stimulus.

4.2. Comparison with criterion shifts

One SDT-based method to capture the mechanisms underlying subjective blindness is to assess the decision criterion under experimental manipulations. The decision criterion (e.g. \( c_0 \)) can be computed from the \( z \)-scores of hit and false alarm rates. If the decision criterion were set too high for the signal to be reported as ‘seen’, this would result in strong stimuli being reported as absent (e.g., Super, Spekreijse, & Lamme, 2001). Previous studies suggested that shifts in decision criteria might underlie various pathological impairments of visual awareness such as neglect (Ricci & Chatterjee, 2004) and blindsight (Azzopardi & Cowey, 1997, 1998; Lau, 2008). Moreover, it has been shown that decision criteria become more conservative in attentional blindness, repetition blindness (Caetta & Gorea, 2010), motion-induced blindness (Caetta, Gorea, & Bonneh, 2007) and visual extinction in normal observers (Gorea & Sagi, 2000, 2001, 2002). The notion of criterion shifts provides a useful theoretical framework to account for how strong sensory signals could go unnoticed by observers without invoking subjective judgments about awareness of stimuli.

Fig. 4. Histograms of confidence ratings for miss and correct rejection (CR) trials under visibility suppression. The probability of confidence ratings is graphed for low, medium and high confidences for (A) contrast reduction, (B) backward masking, (C) flash suppression, (D) dual task, (E) attentional blink and (F) spatial uncertainty experiments. The total number of miss and CR trials for computing the probability distributions is shown in the insets. In all the psychophysical manipulations, conditions that produced suppression of visibility (indicated by the shades in Fig. 3; objective \( \text{AUC} \sim 0.7 \)) were used to produce these graphs.
To compare our SDI analyses with decision criterion, we computed normalized decision criteria $c_0$ ($d_0 = \frac{1}{\sqrt{2}} (z(\text{Hit}) + z(\text{FA})) / \sigma$) for the six tasks under the conditions in which the target was clearly visible in most trials (Fig. 8A) and under the conditions in which visibility was suppressed (Fig. 8B). The results show that the criteria were neutral when the suppression was weak (Fig. 8A) except for the flash suppression task for which the criterion was significantly negative, indicating a liberal criterion in this task. The neutral criterion is optimal in the sense that responses yield the greatest number of correct trials. On the other hand, the criteria were significantly positive in the reduced visibility conditions we used for the SDI analysis. All the task manipulations shifted the criterion positively, indicating a more conservative decision criterion under these conditions (Fig. 8C). The shift in criterion was significant in all the task manipulations except for the dual task condition. The observed positive shifts in decision criteria are consistent with a number of studies reported by Gorea and colleagues (Caetta & Gorea, 2010; Caetta et al., 2007; Gorea & Sagi, 2002).

In our experimental results, a commonly used criterion measure did not show clear differences between the tasks, revealing general upshifts in criterion whenever the visibility was reduced by any manipulation. However, it should be noted that our criterion measures were slightly different from those developed by Gorea and colleagues. The criterion measure we used is $c'$ as referred to in Macmillan and Creelman (2005), but this is not identical to $c'$ ($z(\text{FA})$) used in Gorea and Sagi (2000) or in Caetta et al’s study (2010). Our current experimental design does not allow direct comparison with their measure, since their criterion measure also requires blocked trials for each condition to estimate the false alarm rate separately. Thus, although both Type I criteria and SDI aim to capture the causes of reporting absence in the presence of strong sensory signals, further studies are needed to establish their exact relationships.

While criterion shift explains misses induced by subjective visibility manipulations, a potential caveat is that criterion-based analysis does not directly test whether criterion shifts affect perceptual state or decision making. Specifically, it is possible that even though observers responded that a stimulus was absent, they might have some awareness, but simply reported absence due to the conservative criterion. This is one of the rationales behind adopting Type II performance as a marker for awareness. Further studies that manipulate decision criterion (e.g., Summerfield & Koechlin, 2008) combined with a Type II task may be informative for resolving this issue.

### 4.3. Putative origins of perceptual and attentional blindness

The distinction between perceptual and attentional blindness can be understood in terms of accessibility of the critical cognitive process for each task (Corallo, Sackur, Dehaene, & Sigman, 2008). In perceptual blindness, the lack of discrimi-
between correct rejections and misses implies that a trial-by-trial fluctuation of the critical factor leading to invisibility is not accessible to the mechanisms for estimating confidence. In other words, the cause of a failure to detect a target occurs independent of internal factors that can be monitored by observers. One such example of fluctuation is that of ongoing neural activity. Recent studies showed that the amplitude (Romei et al., 2007) and phase (Busch, Dubois, & VanRullen, 2009) of ongoing alpha activity determines visibility of a stimulus presented at the threshold strength (i.e., perceptual blindness conditions). However, the information contained in such fluctuations does not seem to be available for confidence decisions, thus resulting in unawareness of the target, which is subjectively indistinguishable from physical absence.

On the other hand, when objective visibility is impaired by dual task, attentional blink or spatial uncertainty, the state of the fluctuation relevant for the task performance seems to be available to observers and thus incorporated into their decision on confidence rating. In other words, observers were able to adjust their confidence depending on their internal state. For example, observers may monitor how much attention they were able to allocate to the target detection task in the dual task experiment (Yeh & Wickens, 1988). If failure of attentional allocation is the critical factor for a target miss in a given task, such subjective trial-by-trial estimates of attentional availability for the task can predict their correctness. In such a case, observers can correctly attribute their potential target miss to their inattention to the target.

A previous study found a fundamental difference between unawareness induced by backward masking and attentional blink (Sergent & Dehaene, 2004). Observers in that study gave confidence ratings for target visibility in an all-or-none fashion during the attentional blink, whereas they gave confidence ratings in a more continuous scale for a masked target. This difference was interpreted as follows: attentional blink is caused by a failure of bottom-up signals accessing the global workspace (Baars, 1997; Dehaene, Sergent, & Changeux, 2003), whereas perception of a masked stimulus depends on the gradual activation of sensory regions. The conclusion drawn from our present study is consistent with this view: attentional blink is a failure of attentional access, whereas backward masking reduces visibility by curtailing activations of sensory signals.

4.4. Limitations of meta-cognitive approach to awareness

In the present study, we took the position that meta-cognition of perception reflects awareness (e.g., Kunimoto et al., 2001). A common criticism of this approach, including SDI as well as to any attempt to define a marker for awareness, is that operational markers are not a direct measurement of awareness. Instead, their interpretations should be limited to what the operational marker directly measures (Hulme, Friston, & Zeki, 2009). In this sense, our SDI is also limited and what it directly measures is the ability to distinguish psychophysically induced suppression of visibility from physical absence of stimuli.
namely, meta-cognition of subjective blindness to stimuli. Nevertheless, our experimental results indicate that SDI can be used to demonstrate that there are qualitative differences among commonly used psychophysical manipulations for reducing visibility, presumably due to distinct underlying mechanisms.

5. Concluding remarks

In summary, our results support the idea that failure of visual awareness could arise from suppression of early perceptual signals or failure of attentional access. This distinction offers a natural account for the intuition that in some types of subjective invisibility such as binocular rivalry, we have the impression that a target is phenomenally invisible (i.e., awareness of absence), whereas in other types of manipulations such as attention demanding situations, we do have a sense that we missed a target (i.e., absence of awareness). Finally, the SDI analyses reported here can be applied to probe the nature of other circumstances of impaired awareness such as crowding (Toet & Levi, 1992) or motion-induced blindness (Bonneh, Cooperman, & Sagi, 2001) as well as deficits of visual awareness caused by brain lesions (Humphreys, Romani, Olson, Riddoch, & Duncan, 1994; Weiskrantz, 1997).
Acknowledgments

We would like to thank Victor Lamme, Ned Block, Naotsugu Tsuchiya, Neil Muggleton and Daw-An Wu for comments on an earlier version of the manuscript. This work was supported by Human Frontier Science Foundation for R.K. and the U.K. Medical Research Council (G0700929) for V.W.

References


