Consciousness and Cognition 27 (2014) 109–120

Contents lists available at ScienceDirect

Consciousness and Cognition

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

Different subjective awareness measures demonstrate the influence of visual identification on perceptual awareness ratings

Michał Wierzchoń [a,*], Borysław Paulewicz [b], Dariusz Asanowicz [a], Bert Timmermans [c], Axel Cleeremans [d]

[a] Consciousness Lab, Institute of Psychology, Jagiellonian University, Krakow, Poland
[b] Warsaw School of Social Science and Humanities, Faculty in Katowice, Poland
[c] School of Psychology, University of Aberdeen, Scotland, United Kingdom
[d] Consciousness, Cognition and Computation Group, Center for Research in Cognition & Neurosciences, Universite Libre de Bruxelles, Belgium

A R T I C L E   I N F O

Article history:
Received 8 October 2013

Keywords:
Subjective measures of awareness
Confidence ratings
Perceptual awareness scale
Post-decision wagering
Feeling of warmth
Visual identification task
Consciousness, higher-order theories of consciousness

A B S T R A C T

We compare four subjective awareness measures in the context of a visual identification task and investigate quantitative differences in terms of scale use and correlation with task performance. We also analyse the effect of identification task decisions on subsequent subjective reports. Results show that awareness ratings strongly predict accuracy for all scale types, although the type of awareness measure may influence the reported level of perceptual awareness. Surprisingly, the overall relationship between awareness ratings and performance was weaker when participants rated their awareness before providing identification responses. Furthermore, the Perceptual Awareness Scale was most exhaustive only when used after the identification task, whereas confidence ratings were most exhaustive when used before the identification task. We conclude that the type of subjective measure applied may influence the reports on visual awareness. We also propose that identification task decisions may affect subsequent awareness ratings.

1. Introduction

Subjective measures of consciousness, such as confidence ratings (Dienes & Perner, 2004; Dienes & Seth, 2010) or the Perceptual Awareness Scale (Ramsøy & Overgaard, 2004) have now become very popular in studies dedicated to investigating conscious awareness. Such subjective scales are used to investigate awareness in subliminal perception (Sandberg, Timmermans, Overgaard, & Cleeremans, 2010), attentional blink (Sergent & Dehaene, 2004), blindsight (Overgaard, Fehl, Mouridsen, Bergholt, & Cleeremans, 2008) and implicit learning (Dienes & Seth, 2010), to list just a few examples. The main reason for this resurgent and now widespread use of subjective measures is that they afford easy quantification whilst retaining the very feature one is interested in when attempting to measure conscious experience, namely its subjective character (Overgaard, Jensen, & Sandberg, 2010). Indeed, just like objective measures such as identification or recognition tasks (Holender, 1986; Merikle, Smilek, & Eastwood, 2001), subjective scales of awareness quantitatively probe the accessibility of a percept or a mental representation on a numerical scale, allowing for the fine-tuned assessment of

* Corresponding author. Address: Institute of Psychology, Jagiellonian University, al. Mickiewicza 3, 31-120 Krakow, Poland.
E-mail address: michal.wierzchon@uj.edu.pl (M. Wierzchoń).

http://dx.doi.org/10.1016/j.concog.2014.04.009
1053-8100/

conscious awareness and its relationship with task performance, as well as facilitating comparisons between different conditions.

Congruently, many researchers now seem to agree that subjective measures are better tuned to the measurement of conscious awareness than objective tasks or free verbal reports (see: Dienes & Seth, 2010; Koch & Preuschoff, 2007; Sandberg et al., 2010; Wierzchoń, Asanowicz, Paulewicz, & Cleeremans, 2012). However, there is continuing disagreement on what is in fact measured with a particular subjective scale. This issue is related to the response taxonomies that are used to probe participants' awareness. For instance, consider the Perceptual Awareness Scale (PAS, Ramsøy & Overgaard, 2004; Sandberg et al., 2010) or the continuous scale (CS, Sergent & Dehaene, 2004). In both cases, participants are asked to rate the visibility of a stimulus using scale ratings that range between "no experience" and "maximal visibility" (in the case of PAS, judgements that fall between the scale's extremes are described with labelled intermediate steps, whereas CS is continuous). Researchers have claimed that both measures constitute direct and purely introspective measures of stimulus visibility (Ramsøy & Overgaard, 2004; Sandberg et al., 2010). In other words, PAS and CS are often considered as pure measures of awareness of the stimulus rather than as measures of awareness of knowing that one is aware of the stimuli (Dretske, 2006; so-called "judgement knowledge"). In contrast, confidence ratings (CR, Dienes & Seth, 2010) and post-decision wagering (PDW, Persaud, McLeod, & Cowey, 2007; Ruffman, Garnham, Import, & Connolly, 2001) are more explicitly directed towards assessing participants' certainty in their judgements (that is awareness of possessing experience-related knowledge relevant to the judgement rather than the direct experience itself). In confidence ratings, certainty is expressed through a scale describing the lowest ratings as "not confident at all" and the highest ratings as "very confident". Similarly, with post-decision wagering, participants express their certainty by betting on their judgements with real or artificial money. Here, the scale presupposes that low stakes express low certainty whereas high stakes are assumed to be used when participants are highly confident in their judgements (i.e. are aware of judgements). PDW is thought to be more intuitive for participants (e.g. it has been used in developmental, Ruffman et al., 2001) and used in animal studies (Middlebrooks & Sommer, 2012) but has also been criticised as being less sensitive. This is because the scale seems to involve more conservative response criteria due to risk aversion (Clifford, Arabzadeh, & Harris, 2008; Dienes & Seth, 2010; Fleming & Dolan, 2010; Schurger & Sher, 2008; Szczepanowski, 2010).

Further, we have recently proposed an alternative taxonomy the feeling of warmth scale (FOW, Wierzchoń et al., 2012). The scale has previously been used as a measure of intuition in problem solving (Metcalfe, 1986). We argued that FOW allows us to probe conscious awareness more directly than either CR or PDW, as the scale ratings are not influenced by participants' decision strategies, such as risk aversion (see: Schurger & Sher, 2008), or personal theories on certainty (e.g. the illusion of validity – see: Tobena, Marks, & Dar, 1999).

Given this (welcome) plethora of subjective scales of awareness, it seems essential to carry out experiments that aim to compare the scales directly so as to assess the sensitivity of each of them as well as the influence of other factors such as risk aversion. Different groups have recently carried out such systematic comparisons in the context of perceptual tasks (Sandberg et al., 2010; Szczepanowski, 2010; Szczepanowski, Traczyk, Wierzchoń, & Cleeremans, 2013; Zehetleitner & Rausch, 2013) and memory tasks (Dienes & Seth, 2010; Wierzchoń et al., 2012). However, not all available scales have been compared in the context of the perceptual task and there is no widespread agreement on the interpretation of the differences observed between the scales. Thus, our first aim is to investigate whether awareness ratings will show quantitative differences in terms of scale use and correlation with task performance for perceptual studies. Based on the previous studies we reason that the most pronounced differences should be observed between PAS, that is considered to be a direct measure of subjective visibility (Ramsøy & Overgaard, 2004), and other subjective scales, that rather seems to measure participants' metacognitive judgements on the perception ability. Furthermore, we expect that PDW results will be affected by risk aversion and thus PDW will assess the awareness with lower sensitivity.

In order to interpret the differences between subjective measures of awareness, we have recently proposed a model which aims to describe how awareness is examined with each of the scale taxonomies in context of an artificial grammar learning task (Wierzchoń et al., 2012). The model assumes that all subjective scales of awareness measure metacognitive awareness, i.e. participants' subjective judgement knowledge (Dienes & Scott, 2005). Following this assumption, we proposed that the different scales are distinguishable in two main aspects: the range of states of awareness that they are sensitive to, and the lowest point in the spectrum that they can reliably measure (Wierzchoń et al., 2012). It is, however, questionable whether the same assumptions apply to perceptual awareness scales, so we will analyse the results in this manner. It is essential to answer this question because it touches upon one of the most fundamental issues in consciousness research, namely the unity of consciousness (see e.g. Bayne, 2010). Perceptual awareness and introspective (metacognitive) awareness are often referred to as separate processes (Armstrong, 1981), but e.g. higher-order theories assume that conscious perception necessarily depends on awareness of knowing (Cleeremans, 2011; Lau & Rosenthal, 2011; Pasquali, Timmermans, & Cleeremans, 2010; Rosenthal, 2012). Thus, it seems important to see whether the model we proposed, which assumes that all subjective scales of awareness necessarily measure judgement knowledge, may be applied in context of perceptual awareness.

Furthermore, following other assumptions of higher-order theories of awareness, we hypothesised that the observed relation between task performance accuracy (the task upon which awareness is estimated) and awareness ratings may be modified not only by the type of subjective measure applied (and thus the rating criteria involved), but also by the availability of multiple different cues that may influence accuracy, participants' subjective experience, or both. One of the important cues that seems to be available to participants in all studies investigating subjective measures of awareness is their own reaction in the main task (the awareness ratings for CR and PDW necessarily refer to decision upon which certainty is rated; the similar procedure is usually applied in case of PAS). Thus, we hypothesised that the identification decision on what was perceived may shape participants' assessment regarding the extent to which they were aware of what was presented.

Finally, comparing different scales of awareness is also relevant to addressing another important theoretical issue, namely the question of whether consciousness has a graded or dichotomous character – an ongoing controversial issue in this field (Del Cul, Baillet, & Dehaene, 2007; Nieuwenhuis & de Kleijn, 2011; Overgaard, Rote, Mouridsen, & Ramsøy, 2006; Sergent & Dehaene, 2004). Whether awareness is graded or dichotomous is usually analysed by comparing the response distributions obtained through different scales. Thus, the distributions at each scale point are contrasted assuming that graded access results in equal distribution of the possible rating options (Ramsøy & Overgaard, 2004; Sergent & Dehaene, 2004). The results of previous studies (Sandberg et al., 2010; Wierzchoń et al., 2012) suggest that the conclusions on this theoretical problem may depend on the awareness measure applied, so the distribution of responses was also analysed.

To summarise, in this paper, we compare different measures of conscious awareness in the context of a perceptual task. We investigate the relationship between awareness ratings and visual identification task accuracy and look for quantitative differences in terms of scale use and correlation with task performance. We also analyse conscious experience mechanisms in more details, discussing whether the visual discrimination task response may influence awareness ratings. Here our aim is to investigate the effect of visual discrimination task/awareness rating response order, to see whether perceptual consciousness depends on awareness of knowing. Finally, we aim to analyse whether access observed using each of the scales is graded or dichotomous.

2. Method

2.1. Participants

One hundred and forty-six students of Jagiellonian University (77 females and 74 males) voluntarily participated in the study in exchange for course credits. All participants had normal or corrected-to-normal vision. Participants were randomly assigned to one of eight conditions (depending on applied measure of awareness and task order, i.e. 4 scale conditions (CR, PAS, PDW, FOW) × 2 order conditions (before, after) – see below).

2.2. Materials and procedure

The experiment was run on PC computers using DMDX software (Forster & Forster, 2003). The LCD monitors had a resolution of 1280 × 800 pixels and a frame refresh rate of 60 Hz. We presented participants with a simple visual identification task (VIT) following the psychophysical method of limits (Gescheider, 1997). Pictures from the NimStim set of facial expressions (Tottenham et al., 2009) were used as target stimuli and an image with a complex pattern (constructed with fragments of faces) was used as a mask. The stimuli and the mask were presented in greyscale on a light grey background. The size of the stimuli was 3.8 × 4.0 degrees of visual angle. The viewing distance was around 60 cm.

The procedure is depicted in Fig. 1. Each trial began with a fixation cross, presented at the centre of the screen for 1000 ms, and followed by the target stimulus. The target remained on the screen for a duration randomly chosen amongst

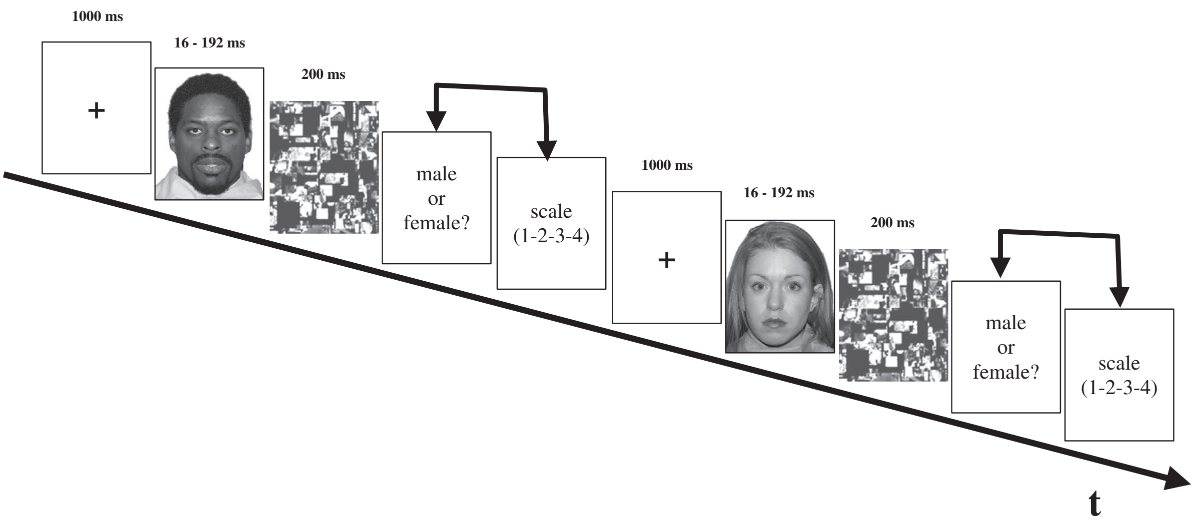

Fig. 1. Diagram of experimental procedure of gender identification task with subsequent trials; the subjective ratings were collected either after of before the forced-choice identification. CR, PAS, PDW and FOW scales were used to report awareness.

Table 1
Scale taxonomies applied in the study. CR – confidence ratings; PAS – perceptual awareness scale; PDW – post-decision wagering; FOW – feeling or warmth scale.

	Rating 1	Rating 2	Rating 3	Rating 4
CR	Guessing	Not confident	Quite confident	Very confident
PAS	No experience	A vague experience	An almost clear experience	Clear experience
PDW	20 PLN stake	40 PLN stake	60 PLN stake	80 PLN stake
FOW	Cold	Cool	Warm	Hot

twelve possibilities ranging from 16 to 192 ms in steps of 16 ms, and was immediately followed by the mask, which was displayed for 200 ms. On each trial, participants were asked to identify whether a target stimulus was a female or a male face as fast as possible and with maximal accuracy. Participants responded by pressing one of two keys (left or right 'ctrl' on a regular "qwerty" keyboard). In addition, participants were asked (either before or after their decision: see below) to rate their awareness of the target stimulus. To facilitate response, a graphical representation of one of four subjective scales of awareness (CR, PAS, PDW, or FOW) was presented. Each scale was displayed in a 2×4 table, with numerical values (1–4) in the first row, and descriptions of these scale points in the lower row. Participants were asked to indicate their response through pressing one of four keys ('1', '2', '3', or '4'). Table 1 presents the responses available to participants at each scale condition.

For CR, rating '1' represents 'guessing', '2' – 'not confident', '3' – 'quite confident', and '4' – 'very confident'. Accordingly, PAS participants rated stimulus visibility (i.e. whether they have a clear experience of an image) with 4-point scale. The descriptions were: 'no experience' (1), 'a vague experience' (2), 'an almost clear experience' (3), and 'clear experience' (4). For PDW, participants were asked to wager on each decision with one of four possible stakes: 20 PLN (1), 40 PLN (2), 60 PLN (3), or 80 PLN (4). When bet on the correct response, they gained the amount and it was added to their account. If not, they lost the amount, which was subtracted from their account. Participants started the task with 200 PLN on their account and were informed that the account would be visible at the end of the phase. Finally, for FOW, participants were asked to rate their awareness based on intuitive feeling of accuracy. The scale descriptions were cold (1), cool (2), warm (3), and hot (4). The scales were administered either before or after each identification task response, i.e. the objective, forced-choice identification either followed or preceded the second-order measure. Both the scales and the order of the responses were between-subject manipulations.

The task consisted of 12 practice trials and 384 experimental trials. The former were divided into two experimental blocks. Thus, each of the twelve target durations was repeated 32 times over the course of the experiment. The order of trials was random for each participant.

We also included the visual identification task/awareness rating order manipulation to control the effects of decision on scale ratings across all conditions.

3. Results

3.1. Identification task accuracy

We first analysed the data by comparing accuracy in the identification task between the scales. We tested whether the overall accuracy of the visual identification task remained unchanged regardless of the type of subjective measure involved. The data were analysed by means of mixed logistic regression with Scale (4 levels: CR, PDW, FOW, PAS), Presentation time (16–192 ms) and their interaction as fixed effects and subject specific random intercept. We observed significant effects of Presentation time that will be described later. Importantly, the main effects of Scale ($|z| \leqslant 0.8$, ns), Order ($|z| \leqslant -0.43$, ns) as well as the interactions of the Scale and Order ($|z| \leqslant 1.12$, ns), and the Scale and Presentation time ($|z| \leqslant 1.76$, ns) were not significant,[1] suggesting that accuracy on the VIT was comparable between groups. Thus, the effects of Awareness ratings described below reflect differences in scale sensitivity (accuracy at the lowest rating) and exhaustiveness (correlation with the objective measure) but not in target visibility.

3.2. Awareness ratings distribution

Next, we turned to awareness ratings distribution, analysing how often different scale ratings were used by the participants. First, we compared the raw response frequencies for each scale and each rating in each Presentation time condition both for the pre-decision and post-decision ratings conditions. As depicted in Fig. 2, the empirical distributions differ between the scales. The most important differences for the purpose of this analysis are those related to the frequency of the lowest ratings and the dichotomy vs. gradualness of the ratings. Preliminary data inspection suggested that the lowest ratings are more often used for PDW as compared to all other ratings, whereas the lowest FOW ratings seemed to be used

[1] A separate model with Scale, Order and Presentation Time as the only predictors was fitted to obtain these values.

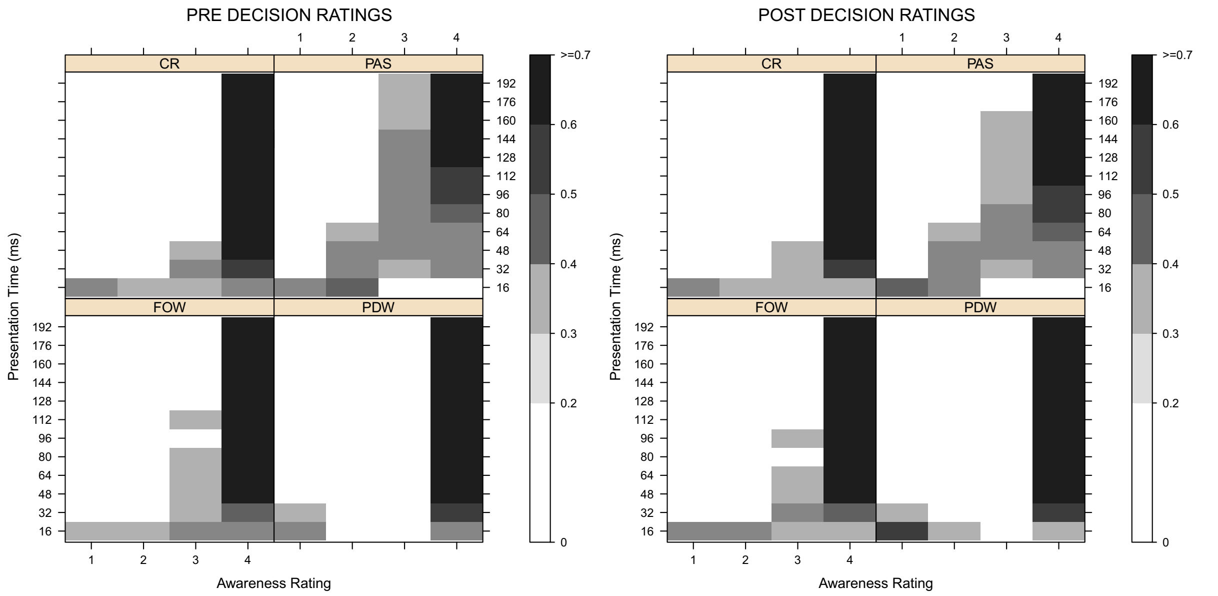

Fig. 2. Distribution of responses observed for a particular time condition (16–192 ms – see *Y* axis) in the pre-decision (left panel) and the post-decision (right panel) condition (colour represents the percentage of ratings for a given presentation time and scale condition). Note, that PDW distributions exhibit binominal character, whereas all other scales seem to be used in more gradual fashion. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

less frequently. As for the gradualness, the PAS scale seemed to be used in more graded fashion than CR and FOW, whereas PDW is used in the most dichotomous fashion. Distributions of responses for the pre-decision and post-decision conditions looked similar (see Fig. 2).

Preliminary data inspection also shows that observed distributions differed mainly for the shortest Presentation time conditions, while for the presentation times equal or longer than 48 ms, we did not observe any systematic differences. Thus, to simplify the statistical analyses we aggregated the distribution data for presentation times $\geqslant$48 ms and compared them with distributions observed for 16 ms and 32 ms. Then, we compared distributions by means of two analyses. First, we compared distribution of the lowest ratings and distribution of all other ratings. This allowed us to determine the response criterion applied by participants or other response strategies resulting in lowest rating overuse (e.g. elevated criterion to rate high, risk aversion strategy, etc.). Next, we analysed the gradualness of scale usage by comparing the distributions of scale extreme ratings with the distribution of middle ratings. We reasoned that these would allow us to test whether participants tend to use the scales in a dichotomous or graded fashion. In the former case, mainly extreme ratings should be used, in the latter, middle ratings should be used as well.

In order to examine the extent to which the lowest ratings were overused we have categorised the ratings into two disjoint categories, i.e., the lowest value vs. all the other values. This enabled us to model the probability of the lowest rating as a function of Scale, Presentation Time (fixed effects) and subject specific strategy (random intercept) with mixed logistic regression. Contrasts based on fitted model coefficients estimating differences between lowest rating probabilities are presented in Table 2.

The lowest PDW ratings were used more often than the lowest FOW ratings in both the 16 ms and 32 ms conditions for the pre-decision and post-decision ratings. Similarly, the lowest PDW ratings were used more frequently than the lowest CR ratings in the 16 ms condition (the effect was significant for post-decision only).

In order to examine the gradualness of the scale usage, we categorised the ratings into two disjoint categories, i.e., the extreme values (i.e. lowest and highest) vs. middle values. This enabled us to model the probability of the middle ratings as a function of Scale, Presentation time (fixed effects) and Subject specific strategy (random intercept) with mixed logistic regression. Contrasts based on fitted model coefficients estimating differences between middle ratings probabilities are presented in Table 3.

In line with the preliminary data inspection the extreme PDW ratings were used more often than the extreme FOW ratings in 16 ms conditions (both pre- and post-decision). We also again observed a significant difference between PDW extreme ratings and CR extreme ratings in 16 ms conditions (both pre- and post-decision). Based on the previous analyses we assumed those two results were an effect of the high frequency of the lowest ratings. However, other differences could be interpreted in terms of scale gradualness or dichotomy. The results clearly show that frequencies of the middle ratings are significantly higher for most of the PAS scale conditions (see Table 3, last column; the less-pronounced differences between scales at the longest presentation time condition seem to be an effect of the very high number of the highest ratings).

Table 2
Differences between the probabilities of the lowest ratings on each scale (ratings 1 on 4-point scales; row variables subtracted from the column variables).

	PDW	CR	PAS
Pre-decision ratings			
16 ms			
FOW	0.2 *	0.12	0.14
PDW		−0.08	−0.06
CR			0.02
32 ms			
FOW	0.04 *	0.01	0.01
PDW		−0.03	−0.03
CR			0
48+ ms			
FOW	0	0	0
PDW		0	0
CR			0
Post-decision ratings			
16 ms			
FOW	0.32 *	−0.01	0.23
PDW		−0.32 *	−0.09
CR			0.23
32 ms			
FOW	0.09 *	0.03	0.02
PDW		−0.06	−0.07
CR			−0.01
48+ ms			
FOW	0	−0.01	−0.01
PDW		−0.01	−0.02
CR			−0.01

* $p < .05$.
** $p < .01$.
*** $p < .001$.

Table 3
Differences between the probabilities of the middle ratings on each scale (ratings 2–3 on 4-point scales; row variables subtracted from the column variables).

	PDW	CR	PAS
Pre-decision ratings			
16 ms			
FOW	−0.37 ***	−0.07	0.06
PDW		0.3 **	0.43 ***
CR			0.13
32 ms			
FOW	−0.21 *	−0.03	0.38 ***
PDW		0.18	0.59 ***
CR			0.41 ***
48+ ms			
FOW	−0.07	−0.06	0.14
PDW		0.01	0.2 ***
CR			0.19 **
Post-decision ratings			
16 ms			
FOW	−0.32 **	−0.04	−0.11
PDW		0.28 **	0.21 *
CR			−0.06
32 ms			
FOW	−0.14	−0.04	0.37 ***
PDW		0.1	0.51 ***
CR			0.41 ***
48+ ms			
FOW	−0.03	0	0.14 *
PDW		0.03	0.17 **
CR			0.14

* $p < .05$.
** $p < .01$.
*** $p < .001$.

Importantly, all other taxonomies seem to be used in a more dichotomous fashion, but the effect is more pronounced in the PDW scale condition.

3.3. Awareness and accuracy

Preliminary data inspection of the VIT accuracy results showed that the effect of Presentation Time was nonlinear, i.e. the accuracy increased substantially between 16 ms, 32 ms, and 48 ms, but it did not increase any further for times longer than 48 ms. Furthermore, there were no systematic differences related to our research questions observed for these longer presentation times. Thus, we decided to use the aggregated model (with 3 Presentation Time conditions: 16 ms, 32 ms and =>48 ms), as it allowed us to simplify the analyses substantially.

The accuracy was analysed by means of a logistic mixed model with Scale (4 levels: FOW, PDW, CR, PAS), Order (2 levels: post- or pre-decision rating), Awareness Rating (4 levels) and Presentation time (3 levels: 16, 32 and 48+ ms), all the possible interactions, and random subject specific intercept using separate intercepts and slopes parameterization to improve the readability of regression coefficients. This model allows us to compare a lowest rating sensitivity and scale exhaustiveness for each subjective scale. Awareness Rating was centred on the lowest value (1), thus the first 12 coefficients in Table 4 (left panel) directly estimate the sensitivity of the lowest scale rating when used after the identification response. The following 12 coefficients estimate the effect of Awareness Rating on accuracy (scale exhaustiveness for post-decision ratings). The coefficients in the right panel of Table 4 represent the differences between the effects of awareness rating applied before and after the VIT response. These coefficients estimate the influence of order on the lowest scale rating sensitivity and scale exhaustiveness. The overall model fit was good ($\chi^2(47) = 8860$, $p < .0001$). As can be seen in Fig. 3 almost all of the aggregated data points fall within the 95% predictive intervals, showing that over-dispersion was not substantial.

3.4. Sensitivity of the lowest scale ratings

As mentioned before, we centred Awareness Rating on the lowest possible value to assess the sensitivity of the lowest rating. Since the logit of .5 equals 0, a significant positive regression intercept within a given scale condition indicates that accuracy was above chance when the lowest rating was used for that scale. As can be seen from Table 4 (lines 1–12, left column), accuracy at the lowest scale point differed significantly from the chance level for all the scales when the presentation

Table 4
Regression coefficients for the logistic regression mixed model for accuracy. The model was centred on the lowest post-decision ratings; left panel represent intercepts and slopes for the scales applied after identification task response; right panel depict differences between pre- and post-decision ratings. Note, that awareness ratings predict accuracy in case of all scales when used after the decision and that this effect is significantly smaller for most of the conditions, when scales are used before the decision.

$N = 146$ # observations 58607	Post-decision ratings				Pre–post-decision ratings			
	Estimate	SE	z	p	Estimate	SE	z	p
Scale FOW 16	0.001	0.197	0.007	.995	0.225	0.31	0.725	.468
Scale FOW 32	1.038	0.31	3.35	.001**	−0.078	0.466	−0.168	.867
Scale FOW 48+	1.083	0.216	5.01	<.001***	0.82	0.367	2.233	.026*
Scale PDW 16	0.11	0.193	0.571	.568	−0.016	0.296	−0.053	.958
Scale PDW 32	0.833	0.265	3.147	.002**	−0.265	0.387	−0.686	.493
Scale PDW 48+	1.079	0.214	5.047	<.001***	1.038	0.361	2.876	.004**
Scale CR 16	0.148	0.215	0.69	.49	0.198	0.315	0.629	.529
Scale CR 32	0.672	0.299	2.247	.025*	0.434	0.474	0.914	.36
Scale CR 48+	0.205	0.225	0.911	.362	0.295	0.403	0.733	.463
Scale PAS 16	0.225	0.199	1.132	.258	−0.005	0.281	−0.017	.986
Scale PAS 32	1.142	0.338	3.377	.001**	0.29	0.473	0.614	.539
Scale PAS 48+	0.414	0.243	1.705	.088	1.441	0.401	3.596	<.001***
Scale FOW 16 ∗ Awareness Rating	0.455	0.084	5.418	<.001***	−0.513	0.122	−4.206	<.001***
Scale FOW 32 ∗ Awareness Rating	0.821	0.133	6.152	<.001***	−0.036	0.202	−0.179	.858
Scale FOW 48+ ∗ Awareness Rating	1.274	0.077	16.589	<.001***	−0.423	0.127	−3.341	.001**
Scale PDW 16 ∗ Awareness Rating	0.679	0.094	7.191	<.001***	−0.542	0.124	−4.373	<.001***
Scale PDW 32 ∗ Awareness Rating	1.31	0.165	7.928	<.001***	−0.415	0.203	−2.043	.041*
Scale PDW 48+∗ Awareness Rating	1.474	0.082	17.875	<.001***	−0.522	0.131	−3.981	<.001***
Scale CR 16 ∗ Awareness Rating	0.333	0.087	3.844	<.001***	−0.358	0.122	−2.942	.003**
Scale CR 32 ∗ Awareness Rating	0.955	0.138	6.899	<.001***	−0.325	0.202	−1.61	.107
Scale CR 48+ ∗ Awareness Rating	1.474	0.072	20.347	<.001***	0.082	0.14	0.587	.557
Scale PAS 16 ∗ Awareness Rating	0.591	0.135	4.39	<.001***	−0.571	0.165	−3.47	.001**
Scale PAS 32 ∗ Awareness Rating	1.174	0.224	5.249	<.001***	−0.413	0.281	−1.471	.141
Scale PAS 48+ ∗ Awareness Rating	1.8	0.107	16.803	<.001***	−0.76	0.164	−4.64	<.001***

Likelihood ratio $\chi^2(47) = 8860$, $p < .0001$.
* $p < .05$.
** $p < .01$.
*** $p < .001$.

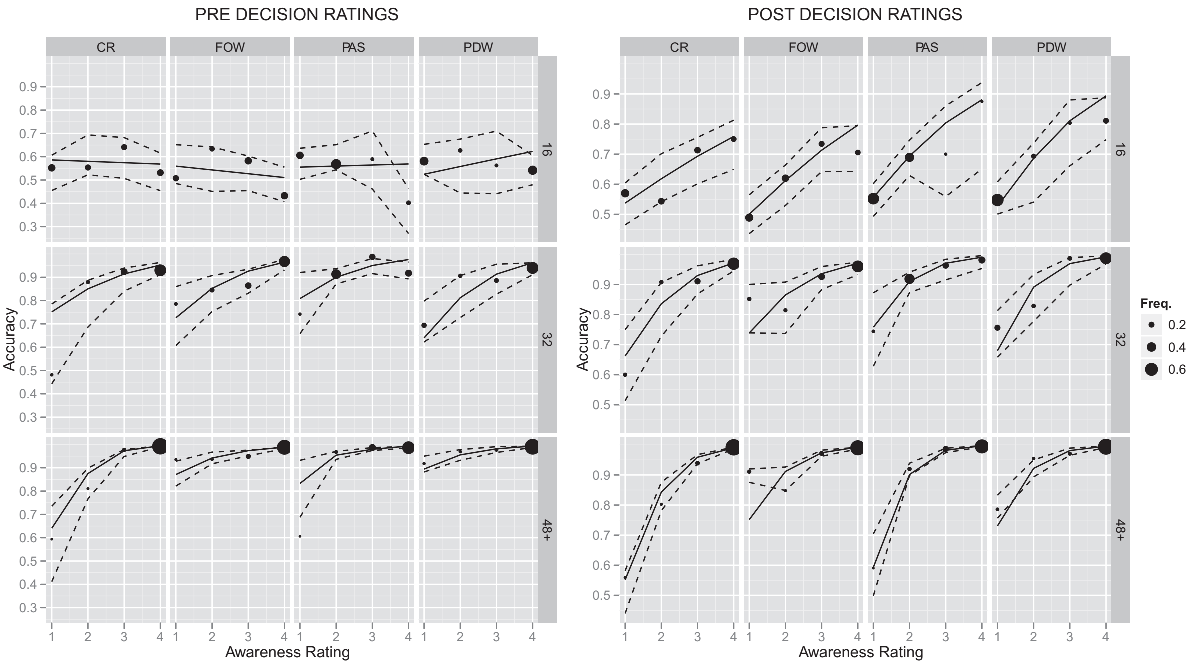

Fig. 3. Accuracy scores plotted separately for each scale (columns) and presentation time (rows) in the pre-decision (left panel) or the post-decision (right panel) condition. Solid lines represent fitted values based on fixed effects only. Dotted lines represents the 95% predictive intervals based on both the fixed and the random effects. Size of the dots represents the distribution of the responses (frequency within each condition). Note, that slopes are less pronounced on the left panel and disappear for the shortest presentation time condition.

time was longer than 16 ms. The exceptions were the CR and PAS scale when chance level accuracy was observed also at the longest presentation time (48+ ms), but this seems to be an effect of the response distribution, i.e. the much smaller number of observations in those conditions (see size of the dots at Fig. 3, left panel).

3.5. Scale exhaustiveness analysis

Effects of awareness ratings are significant for every scale (regression slopes in Table 4) meaning that the VIT accuracy increased with ratings for all the scales. To assess the exhaustiveness of the scales in more detail we compared the regression slopes by means of contrast analysis (see Table 5). The only systematic effects were observed at the longest presentation time conditions, in which the PAS slope was higher than the slopes of the other measures in the post-decision condition. This

Table 5
Differences between slopes (row variables subtracted from the column variables); post-decision ratings only. Contrasts were calculated based on the fit summary statistics for the full model of mixed regression parameterized for the analysis presented in Table 4.

	PDW	CR	PAS
16 ms			
FOW	0.22	−0.12	0.14
PDW		−0.35 **	−0.09
CR			0.26
32 ms			
FOW	0.49 *	0.13	0.35
PDW		−0.35	−0.14
CR			0.22
48+ ms			
FOW	0.2	0.2	0.53 ***
PDW		0	0.33 *
CR			0.33 *

* $p < .05$.
** $p < .01$.
*** $p < .001$.

Table 6
Differences between slopes (row variables subtracted from the column variables); pre-decision ratings only. Contrasts were calculated based on the fit summary statistics for the full model of mixed regression parameterized for the analysis presented in Table 4.

	PDW	CR	PAS
16 ms			
FOW	0.19	0.03	0.08
PDW		−0.16	−0.12
CR			0.05
32 ms			
FOW	0.11	−0.15	−0.02
PDW		−0.26	−0.13
CR			0.13
48+ ms			
FOW	0.1	0.7 ***	0.19
PDW		0.6 ***	0.09
CR			−0.52 **

* $p < .05$.
** $p < .01$.
*** $p < .001$.

suggests that PAS is more exhaustive than all the other scales when used after VIT. Note that the pattern does not generalise to the pre-decision rating Order (see below – Table 6).

3.6. Influence of the response order on awareness ratings

We now turn to the most interesting results of the present study, namely the effect of the response order. As can be seen in Table 4 (right panel, rows 1–12), Order influenced the sensitivity of the lowest scale values at the longest presentation times for all the scales with the exception of CR. The respective coefficients suggest that the sensitivity of the lowest ratings decreases for the =>48 ms presentation time when participants rate their awareness before identification task response. However, this effect does not seem very important, since the number of lowest ratings is very low in this condition. More interestingly, however, the exhaustiveness of the scales was also systematically lower in the pre-decision condition (see Table 4, right panel, rows 13–24). In most of the Scale conditions awareness ratings predicted the accuracy to a greater extent, when the scale was used after the identification task decision was overtly made. The exception is CR, which seems to be similarly exhaustive both in case of pre and post-decision conditions, when the presentation time was longer than 16 ms. We also observe similar slopes for the single presentation times in case of FOW and PAS (both in case of 32 ms condition). Finally, we observed significant differences between scales comparing regression slopes for the pre-decision condition (see Table 6). Interestingly, this time CR seems to be more exhaustive than any other scale, including PAS.

3.7. Response time effects

One way to explain the observed Order effects is by noting that awareness ratings made prior to the identification task responses are less indicative of accuracy simply because the participants had less time to judge their awareness, or their awareness had less time to develop. To test this hypothesis we compared the main model represented in Table 4 with an alternative model including the time of awareness rating response. In order to compare the awareness rating time factor for pre- and post-decision conditions, we prepared the standardised awareness rating time factor, that in the pre-decision group equals the awareness rating time, whereas for the post-decision condition, it corresponds to the VIT decision time plus the awareness rating time. Surprisingly, we observed a significant, but negative effect of awareness rating time within both pre- and post-decision ratings conditions ($ps < .01$). This is contradictory to the interpretation suggested above, so we modelled the awareness rating time effect with a spline curve and it turned out that the effect of awareness rating time was virtually flat in the <700 ms range within both decision phase conditions, but for the longer awareness rating times a much lower accuracy was observed. This is probably due to the fact that the longest awareness rating times represent random responses to a much greater degree. Nevertheless, the inclusion of the awareness rating time did not result in a significant improvement in model fit ($\chi^2(2) = 0$, $p = 1$).

4. General discussion

The aim of the reported study was to compare awareness ratings collected with four subjective measures of consciousness. We also aimed to estimate the influence of the visual identification task and awareness measure responses order on the relation between awareness ratings and identification task accuracy. We exclude the possibility that the observed results are an effect of stimulus visibility differences between the scales.

4.1. Scale comparison

There are at least three reasons why the response distribution analysis is important for our data analysis. First, differences in response distribution might have affected the relation between awareness and identification accuracy. In other words, some differences observed for the scale sensitivity and scale exhaustiveness analyses reported below might stem from differences in response distributions observed for the scales. This is especially important when a particular scale rating is used only occasionally and the reliability of the estimates for that rating is poor, or is used too often (as this may suggest the existence of some external variables that influence a rating – e.g. response strategy). Next, the frequency of the lowest ratings provides information about the strategy participants used when reporting their conscious awareness. Here, the more frequently the lowest rating is used, the more conservative the response strategy of one scale relative to another is (note that this analysis is informative only when at least two subjective measures are compared, as there is no objective estimate on how often the lowest rating should be used if unbiased by a strategy). This strategy may stem from risk aversion, or from other strategies induced by a given taxonomy (e.g. income maximisation strategy in PDW). Finally, the frequency of the intermediate ratings allows us to discuss the issue of conscious awareness gradualness.

Participants used all the ratings available for all the scales, however we mostly observed the highest awareness ratings for presentation times that exceeded 48 ms. This is probably the effect of the stimuli we applied, as it is well known that human faces are automatically processed and thus easily recognised (see e.g. Fox, Moon, Iaria, & Barton, 2009; Haxby, Hoffman, & Gobbini, 2002). Because of the high frequency of the highest ratings observed for longer presentation times, we focused the analysis on the three presentation times (i.e. 16 ms, 32 ms and 48+ ms).

The lowest PDW rating seemed to be used more frequently than in the other scales. We reason that in case of PDW scale the risk aversion induced by wagering effects participants' response strategy resulting in higher response criterion applied in this condition (Clifford et al., 2008; Schurger & Sher, 2008; Szczepanowski, 2010). We also observed higher frequency of the PAS intermediate ratings, as compared to all other scales. CR and FOW seem to be used in a slightly more all-or-nothing fashion, but the effect is most pronounced when PAS is compared with PDW, with the PAS middle rating being used more often. This clearly shows that the apparent gradualness of conscious awareness depends on the type of subjective scale that has been used. Importantly, the observed results patterns for FOW and CR are similar, suggesting that our new scale does allow for fine-tuned assessment of awareness. It also seems to apply a similar response criteria to CR. Finally, it seems that the more dichotomous result pattern may stem from a response strategy that is more rewarding for a particular scale. For example, it seems possible that the PDW all-or-nothing pattern is an effect of the income maximisation strategy applied by participants (Clifford et al., 2008). Thus, we conclude that conscious awareness seems to be more gradual than dichotomous.

We now turn to scale sensitivity analysis, comparing the level of awareness observed with the lowest ratings (i.e. criterion of the strength of sensory evidence applied by participants to rate higher than one) as well as the interactions between awareness ratings and accuracy of the identification task (exhaustiveness of a scale to probe different level of conscious awareness). The results revealed systematic differences between awareness taxonomy applied and identification task accuracy.

Identification accuracy for a presentation time of greater than 16 ms exceeds chance. This suggests that even though the response strategies differences observed, participants at all conditions seemed to be sensitive for a weak sensory evidences to the same extent. This may also suggest that the applied response criterion is distorted in all the scales to some extent, either as an effect of risk aversion (as in case of PDW – see: Clifford et al., 2008; Schurger & Sher, 2008), or the other assumptions people made on the scale requirements (e.g. illusion of validity – see: Tobena et al., 1999).

More interestingly, we observed significant differences in the scale exhaustiveness analysis. The results indicated that the awareness ratings predicted accuracy to a greater extent in the case of PAS compared to all other scales. This result is in line with (Sandberg et al., 2010) conclusions, which suggest that PAS is the most exhaustive measure of perceptual awareness. Interestingly, this effect was observed only in a post-decision ratings condition, i.e. when the awareness ratings were used after the identification task decision. When the scales were used before the VIT decision, CR appears to be more exhaustive than other scales. This effect required explanation, but before we comment on that, we turn to the most interesting part of the results, i.e. response order analyses.

4.2. Order factors

The observed differences between pre- and post-decision ratings as well as the negative effect of awareness rating time on accuracy were surprising, so we took a closer look at possible explanations. Two theories seem to be most probable here. The first would predict that awareness is inferred through the observation of actions (or even that the experience of seeing occurs when our actions result in sensorimotor contingencies – see: O'Regan and Noë, 2001). Here, we propose that the accurate reports on conscious awareness require both stimulus identification and the behavioural decision that results from this identification. This is in line with Higher-Order theories of consciousness (Lau & Rosenthal, 2011; Rosenthal, 2005, 2012). This assumes that first-order information processing provides content for the higher-order representations, and that those higher-order representations are necessary for consciousness. We reason that because the first-order processing often includes actions or action related processing (action planning, action controlling), the reports on awareness are more appropriate when a first-level process is completed (for an alternative explanation see also an analysis investigating the time requirements of judgement knowledge – see Mealor & Dienes, 2013). If this interpretation of our results is well-grounded, the accurate awareness ratings should be based on the decision-making (in fact all information processing resulting in decision making). Thus, the awareness ratings should be more sensitive when the identification task decision is made before awareness rating. In the context of the recent theoretical discussion on the characteristics of conscious processing (Block, 2011; Cohen & Dennett, 2011; Lamme, 2010; Lau & Rosenthal, 2011), our results suggest that perceptual consciousness should not be perceived as a passive, rich and detailed expression of the external stimulation formed at the low levels of the visual cortex, as the information processing involved in the decision making significantly increases the accuracy of consciousness reports.

An alternative explanation of the response order effect would assume that awareness requires time, because participants need to accumulate evidence allowing estimating awareness accurately (Cleeremans & Sarrazin, 2007). In the case of our study, this interpretation predicts that pre-decision ratings would be less accurate simply because participants had insufficient time to rate their awareness accurately. To discriminate between these two interpretations we ran an additional analysis comparing our basic model with the other model including rating time (time between target stimuli presentation and scale ratings key press). We reasoned that the second theory would predict that the interactions observed in the first model should disappear (or at least decrease), when the rating time is included in the model (because the rating time in fact estimates the time effects on consciousness ratings). The results clearly show that the inclusion of rating time did not result in model fit improvement, suggesting that this explanation is insufficient. This alone seems to suggest that the first interpretation we proposed is more probable and that identification task decision is necessary to rate awareness accurately. However, further analysis suggests that the time effects in the process of conscious experience formation remain ambiguous, as it is not clear whether the accumulation of information does not influence the accuracy or just whether the decision order factor has masked the time effects. Thus, further research investigating the effects of time and decision making separately should be conducted.

Finally, it seems interesting that we have not found any qualitative differences between the subjective scales in terms of how sensitive they are to task order. It has been suggested that PAS is a more direct measure of awareness than CR or PDW (Ramsøy & Overgaard, 2004; Sandberg et al., 2010; Zehetleitner & Rausch, 2013). If this position is well-founded, we should expect PAS to be less affected by the response order, as it should directly measure visibility but not higher-order reports on perceptual decision. Our results clearly show that the identification task decision influences PAS rating to the same extent as CR or PDW. Importantly, we even show that CR is more exhaustive than PAS when used before the visual identification (we reason this is because following the CR instruction participants need to decide on identification response before awareness rating even at the pre-decision condition, and only then do they make this decision). Thus, it seems that accurate PAS ratings require knowledge on the executed action, as do the other scales. In our opinion this suggest that PAS ratings should be interpreted as an awareness of the fact of seeing (Dretske, 2006), i.e. a metacognitive measure of the awareness, but not direct reports on visibility. Consequently, we also propose that differences in accuracy observed between scale ratings actually are the result of differences in scale sensitivity. In other words, in line with our theoretical model (Wierzchoń et al., 2012), we proposed that each scale can measure a different part of the consciousness spectrum, i.e. each rating of each scale estimates a slightly different level of awareness. Thus, we propose that all scales are capable of assessing conscious awareness, but one should avoid scales applying conservative rating criterion (e.g. PDW) when investigating low levels of consciousness. We also propose that a precise measure of awareness may require a combination of the subjective measures allowing for a more fine-tuned estimation of conscious awareness (as was also the case in this experiment with FOW and PAS scale: e.g. the former was more exhaustive at the 32 ms condition, whereas the latter was more exhaustive with longer presentation time. This was the case for post-decision condition only).

5. General conclusions

The most interesting result of our study is that the effects of awareness ratings seem to be based on the action associated with the identification response. In line with higher-order thought theory we conclude that action execution is necessary to rate awareness accurately, although the time effects should be investigated further. We reason that this is because an action is an outcome of the information processing on which the awareness ratings are based. This also suggests that ratings collected after the response are based on, or at least heavily influenced by judgement knowledge, regardless of the measure applied. Similar conclusions can be drawn for PAS, despite the fact that PAS does not explicitly require participants to reflect on their judgement. Our study also confirms that the PAS scale is more exhaustive in measuring conscious awareness than other scales. However, we also claim that PAS seems to measure metacognitive judgements on visibility, rather than direct subjective experience (as we have not observed any qualitative differences between order effects for this scale as compared to other scales). CR and FOW are also sufficiently sensitive (the former is even more exhaustive than PAS, when the taxonomy is applied before identification task response). Comparing results obtained with different subjective measures, we conclude that all measures probe participants' conscious awareness, although each scale seems to be differentially sensitive depending on task conditions. This suggests that scales should ideally be combined to probe conscious awareness with more details. Consequently, our new measure, FOW, is a reliable and sensitive measure of awareness that can be used especially when the level of awareness is low (i.e. when a conservative response criterion is problematic).

Acknowledgments

M.W. was supported with SONATA BIS Program granted by National Science Centre for the research described in this paper (Grant 2012/07/E/HS6/01037). A.C. is a Research Director with the F.R.S.-FNRS (Belgium). The authors would like to thank Marta Siedlecka for her help with data collection.

References

Armstrong, D. M. (1981). *The nature of mind and other essays*. Ithaca: Cornell University Press.

Bayne, T. (2010). *The unity of consciousness*. New York, USA: Oxford University Press.

Block, N. (2011). Perceptual consciousness overflows cognitive access. *Trends in Cognitive Sciences, 15*(12), 567–575. http://dx.doi.org/10.1016/j.tics.2011.11.001.

Cleeremans, A. (2011). Frontiers: The radical plasticity thesis: How the brain learns to be conscious. *Frontiers in Consciousness Research, 2*(86), 1–12.

Cleeremans, A., & Sarrazin, J. C. (2007). Time, action, and consciousness. *Human Movement Science, 26*(2), 180–202. http://dx.doi.org/10.1016/j.humov.2007.01.009.

Clifford, C. W., Arabzadeh, E., & Harris, J. A. (2008). Getting technical about awareness. *Trends in Cognitive Sciences, 12*(2), 54–58. http://dx.doi.org/10.1016/j.tics.2007.11.009.

Cohen, M. A., & Dennett, D. C. (2011). Consciousness cannot be separated from function. *Trends in Cognitive Sciences, 15*(8), 358–364. http://dx.doi.org/10.1016/j.tics.2011.06.008.

Del Cul, A., Baillet, S., & Dehaene, S. (2007). Brain dynamics underlying the nonlinear threshold for access to consciousness. *PLoS Biology, 5*(10), e260. http://dx.doi.org/10.1371/journal.pbio.0050260.

Dienes, Z., & Perner, J. (2004). Assumptions of a subjective measure of consciousness: Three mappings. In R. Gennaro (Ed.), *Higher order theories of consciousness* (pp. 173–199). Amsterdam: John Benjamins Publishers. Retrieved from Google Scholar.

Dienes, Z., & Scott, R. (2005). Measuring unconscious knowledge: Distinguishing structural knowledge and judgement knowledge. *Psychological Research Psychologische Forschung, 69*(5–6), 338–351.

Dienes, Z., & Seth, A. (2010). 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.

Dretske, F. (2006). Perception without awareness. *Perceptual Experience*, 147–180.

Fleming, S. M., & Dolan, R. J. (2010). Effects of loss aversion on post-decision wagering: Implications for measures of awareness. *Consciousness and Cognition, 19*(1), 352–363.

Forster, K. I., & Forster, J. C. (2003). DMDX: A windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers: A Journal of the Psychonomic Society Inc, 35*(1), 116–124.

Fox, C. J., Moon, S. Y., Iaria, G., & Barton, J. J. (2009). The correlates of subjective perception of identity and expression in the face network: An fMRI adaptation study. *NeuroImage, 44*(2), 569–580. http://dx.doi.org/10.1016/j.neuroimage.2008.09.011.

Gescheider, A. G. (1997). *Psychophysics: The fundamentals*. London: LEA.

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2002). Human neural systems for face recognition and social communication. *Biological Psychiatry, 51*(1), 59–67.

Holender, D. (1986). Semantic activation without conscious identification in dichotic listening, parafoveal vision, and visual masking: A survey and appraisal. *Behavioral and Brain Sciences, 9*(01), 1–23. Retrieved from Google Scholar.

Koch, C., & Preuschoff, K. (2007). Betting the house on consciousness. *Scientific American Mind, 18*(3), 16–17. Retrieved from Google Scholar.

Lamme, V. A. F. (2010). How neuroscience will change our view on consciousness. *Cognitive Neuroscience, 1*(3), 204–220. http://dx.doi.org/10.1080/17588921003731586.

Lau, H. C., & Rosenthal, D. (2011). Empirical support for higher-order theories of conscious awareness. *Trends in Cognitive Sciences, 15*(8), 365–373.

Mealor, A. D., & Dienes, Z. (2013). The speed of metacognition: Taking time to get to know one's structural knowledge. *Consciousness and Cognition, 22*(1), 123–136. http://dx.doi.org/10.1016/j.concog.2012.11.009.

Merikle, P. M., Smilek, D., & Eastwood, J. D. (2001). Perception without awareness: Perspectives from cognitive psychology. *Cognition, 79*(1–2), 115–134.

Metcalfe, J. (1986). Premonitions of insight predict impending error. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12*(4), 623–634.

Middlebrooks, P. G., & Sommer, M. A. (2012). Neuronal correlates of metacognition in primate frontal cortex. *Neuron, 75*(3), 517–530. http://dx.doi.org/10.1016/j.neuron.2012.05.028.

Nieuwenhuis, S., & de Kleijn, R. (2011). Consciousness of targets during the attentional blink: A gradual or all-or-none dimension? *Attention, Perception, & Psychophysics, 73*(2), 364–373. http://dx.doi.org/10.3758/s13414-010-0026-1.

O'Regan, J. K., & Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *The Behavioral and Brain Sciences, 24*(5), 939–973. discussion 973-1031.

Overgaard, M., Fehl, K., Mouridsen, K., Bergholt, B., & Cleeremans, A. (2008). Seeing without seeing? Degraded conscious vision in a blindsight patient. *PLoS One, 3*(8), e3028. http://dx.doi.org/10.1371/journal.pone.0003028.

Overgaard, M., Jensen, M., & Sandberg, K. (2010). Methodological pitfalls in the "objective" approach to consciousness: Comments on Busch et al. (2009). *Journal of Cognitive Neuroscience, 22*(9), 1901–1902. http://dx.doi.org/10.1162/jocn.2009.21403.

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.

Pasquali, A., Timmermans, B., & Cleeremans, A. (2010). Know thyself: Metacognitive networks and measures of consciousness. *Cognition, 117*(2), 182–190. http://dx.doi.org/10.1016/j.cognition.2010.08.010.

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

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

Rosenthal, D. M. (2005). *Consciousness and mind*. Alderley, UK: Clarendon Press.

Rosenthal, D. (2012). Higher-order awareness, misrepresentation and function. *Philosophical Transactions of the Royal Society of London, Series B: Biological sciences, 367*(1594), 1424–1438. http://dx.doi.org/10.1098/rstb.2011.0353.

Ruffman, T., Garnham, W., Import, A., & Connolly, D. (2001). Does eye gaze indicate implicit knowledge of false belief? Charting transitions in knowledge. *Journal of Experimental Child Psychology, 80*(3), 201–224. http://dx.doi.org/10.1006/jecp.2001.2633.

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.

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

Sergent, C., & Dehaene, S. (2004). Is consciousness a gradual phenomenon? *Psychological Science: A Journal of the American Psychological Society/APS, 15*(11), 720–728. Retrieved from Google Scholar.

Szczepanowski, R. (2010). Absence of advantageous wagering does not mean that awareness is fully abolished. *Consciousness and Cognition, 19*(1), 426–431.

Szczepanowski, R., Traczyk, J., Wierzchoń, M., & Cleeremans, A. (2013). The perception of visual emotion: Comparing different measures of awareness. *Consciousness and Cognition, 22*(1), 212–220. http://dx.doi.org/10.1016/j.concog.2012.12.003.

Tobena, A., Marks, I., & Dar, R. (1999). Advantages of bias and prejudice: An exploration of their neurocognitive templates. *Neuroscience and Biobehavioral Reviews, 23*(7), 1047–1058.

Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., et al (2009). The NimStim set of facial expressions: Judgements from untrained research participants. *Psychiatry Research, 168*(3), 242–249. http://dx.doi.org/10.1016/j.psychres.2008.05.006.

Wierzchoń, M., Asanowicz, D., Paulewicz, B., & Cleeremans, A. (2012). Subjective measures of consciousness in artificial grammar learning task. *Consciousness and Cognition, 21*(3), 1141–1153. http://dx.doi.org/10.1016/j.concog.2012.05.012.

Zehetleitner, M., & Rausch, M. (2013). Being confident without seeing: What subjective measures of visual consciousness are about. *Attention, Perception, & Psychophysics*. http://dx.doi.org/10.3758/s13414-013-0505.