The Level of Processing Modulates Visual Awareness: Evidence from Behavioral and Electrophysiological Measures

Mikel Jimenez1,2, Claudia Poch3,4, Cristina Villalba-García4,5, Lucía Sabater4, José Antonio Hinojosa3,4, Pedro R. Montoro1, and Mika Koivisto6

Abstract

The level of processing hypothesis (LoP) proposes that the transition from unaware to aware visual perception is graded for low-level (i.e., energy, features) stimulus whereas dichotomous for high-level (i.e., letters, words, meaning) stimulus. In this study, we explore the behavioral patterns and neural correlates associated with different depths (i.e., low vs. high) of stimulus processing. The low-level stimulus condition consisted of identifying the color (i.e., blue/blueish vs. red/reddish) of the target, and the high-level stimulus condition consisted of identifying stimulus category (animal vs. object). Behavioral results showed that the levels of processing manipulation produced significant differences in both the awareness rating distributions and accuracy performances between tasks, the low-level task producing more intermediate subjective ratings and linearly increasing accuracy performances and the high-level task producing less intermediate ratings and a more nonlinear pattern for accuracies. The electrophysiological recordings revealed two correlates of visual awareness, an enhanced posterior negativity in the N200 time window (visual awareness negativity [VAN]), and an enhanced positivity in the P3 time window (late positivity [LP]). The analyses showed a double dissociation between awareness and the level of processing hypothesis manipulation: Awareness modulated VAN amplitudes only in the low-level color task, whereas LP amplitude modulations were observed only in the higher level category task. These findings are compatible with a two-stage microgenesis model of conscious perception, where an early elementary phenomenal sensation of the stimulus (i.e., the subjective perception of color) would be indexed by VAN, whereas stimulus’ higher level properties (i.e., the category of the target) would be reflected in the LP in a later latency range.

INTRODUCTION

Understanding how we become aware of a stimulus, how the subjective experience of seeing the stimulus and its contents emerges, is a crucial problem within consciousness research. A percept may emerge to consciousness evolving through increasing degrees of clarity (i.e., gradually) or abruptly transitioning from unawareness to awareness (i.e., following a dichotomous, “all-or-none” pattern). Research on this issue has produced results supporting both accounts, some studies suggesting an all-or-none (i.e., abrupt) transition from unaware to aware visual perception (Asplund, Fougnie, Zughni, Martin, & Marois, 2014; Sekar, Findley, Poeppel, & Llinás, 2013; Del Cul, Baillet, & Dehaene, 2007; Sergent & Dehaene, 2004), whereas others showing that awareness can evolve through different intermediate perceptions (Pretorius, Tredoux, & Malcolm-Smith, 2016; Sandberg, Timmermans, Overgaard, & Cleeremans, 2010; Overgaard, Rote, Mouridsen, & Ramsoy, 2006; Ramsoy & Overgaard, 2004). In an attempt to reconcile both accounts, Windey and Cleeremans (2015) recently proposed a theoretical framework—the level of processing hypothesis (LoP)—by which the transition from unaware to aware visual experience is graded for low-level stimulus representations (i.e., stimulus “energy” or “feature” levels), whereas it is abrupt or dichotomous for high-level (i.e., the perception of “letters,” “words,” or “meaning”) stimulus perception. The hierarchy of representational levels was based on the proposal by Kouider, de Gardelle, Sackur, and Dupoux (2010), where stimulus information would be accessed independently at different (“energy,” “feature,” “letters,” etc.) stimulus levels or representations. According to the LoP, differences in transitions to awareness between lower and higher levels of stimulus processing would be explained by the different nature of their representations: Whereas high-level stimulus perception (such as letters, numbers, or words) usually refers to a precise, qualitatively distinct concept, low-level stimulus characteristics (such as contrast, color, or basic features) quantitatively vary on a physical continuum and are not easily associated to a single label. Therefore, a high-level stimulus perception would “pop-up” into awareness in an all-or-none
manner, whereas a low-level stimulus perception would emerge into awareness following a more gradual pattern.

Behavioral evidence has provided some support to the postulates of the LoP hypothesis (Derda et al., 2019; Jimenez, Villalba-García, Luna, Hinojosa, & Montoro, 2019; Binder et al., 2017; Anzulewicz et al., 2015; Windey, Vermeiren, Atas, & Cleeremans, 2014; Windey, Gevers, & Cleeremans, 2013). LoP effects are usually visible at threshold stimulus durations (commonly established at around 50 msec target–mask interval; see Del Cul et al., 2007) when comparing a low-level color task and a more complex high-level task usually involving postperceptual numerical judgments (see Jimenez, Hinojosa, & Montoro, 2020, for a review). However, the differences between levels of processing are usually quite subtle and not consistently observable in the same awareness measures. They are sometimes observed as a higher use of intermediate ratings in a subjective awareness scale in the low-level task compared with the higher level task (Binder et al., 2017; Anzulewicz et al., 2015) or, conversely, as a significant difference in accuracies with no differences in the awareness scale distribution (Derda et al., 2019; Windey et al., 2014). Differences in RTs have also been found between levels of processing and used as evidence of the LoP effect by different studies (Derda et al., 2019; Anzulewicz et al., 2015). Nonetheless, there is also evidence pointing to an absence of differences between levels of processing on awareness scale distributions or accuracy performances (Jimenez, Grassini, Montoro, Luna, & Koivisto, 2018).

Neuroimaging studies on the LoP hypothesis are still scarce. Binder et al. (2017) conducted an fMRI study and found that the LoP manipulation influenced brain activity in posterior visual areas. The authors suggested that effects on awareness might be mediated by an attentional top–down modulation of activity in visual regions. Jimenez et al. (2018) explored the ERP correlates of different degrees of awareness associated to different depths (low: stimulus localization; high: letter/number identification) of stimulus processing. Their results showed the two typical correlates of visual awareness (see Förster, Koivisto, & Revonsuo, 2020, for a review): an enhanced posterior negativity in the N200 time window (visual awareness negativity [VAN]) and an enhanced positivity in the P3 time window (late positivity [LP]). Interestingly, the LoP manipulation showed that awareness levels modulated N200/VAN amplitudes in a graded manner only for the low-level task, whereas P3/LP amplitudes were modulated in a graded manner for both low- and high-level tasks. The finding that VAN was sensitive to the LoP manipulation was consistent with task effects occurring in the visual cortex and probably mediated by attention to task-relevant features. In the LP time window, amplitudes in both tasks correlated more directly with graded awareness and behavioral accuracy. Alternative results were found in a recent ERP study by Derda et al. (2019). Participants performed either a low-level (color identification) or high-level (magnitude) task and informed their subjective visibility of the stimulus on a 4-scale rating scale (i.e., Perceptual Awareness Scale [PAS]; Ramsoy & Overgaard, 2004). Their results showed that VAN amplitudes correlated with PAS ratings at both levels of processing, the amplitudes being more negative for increasing levels of awareness. In the LP time window, awareness levels modulated LP amplitudes only in the high-level magnitude task, with more positive amplitudes for increasing levels of awareness. Overall, results by Jimenez et al. (2018) and Derda et al. (2019) produced a complex pattern of results.

In this study, we aimed to gather further evidence on the behavioral and electrophysiological patterns associated to different depths (i.e., low and high) of stimulus processing by introducing line drawings and semantic categorizations for the first time within LoP research. The low-level stimulus processing consisted in a color identification task (a common low-level condition within LoP studies), whereas the high-level stimulus processing consisted of a category discrimination task, a task commonly regarded as postperceptual since it involves further cognitive processing (i.e., decision-making) on the perceived stimulus (Cohen, Ortego, Kyroudis, & Pits, 2020; Rutiku & Bachmann, 2017; Koivisto, Salminen-Vaparanta, Grassini, & Revonsuo, 2016). Thus, the tasks required the identification of either the color or the semantic category, without the confounding magnitude comparison element present in many previous studies on LoP (Derda et al., 2019; Binder et al., 2017; Anzulewicz et al., 2015; Windey et al., 2013). In the low-level color task, participants had to report whether the stimulus was “red/reddish” (three different red hues were used) or “blue/bluish” (three different blue hues were used). In the high-level categorization task, the observers had to inform whether the stimulus was an object (three different objects were used) or an animal (three different animals). Subjective awareness of the target was reported on a 4-scale rating scale adapted from the PAS (Ramsoy & Overgaard, 2004). Importantly, to avoid “criterion shifts” confounds in the use of the awareness scale between tasks (i.e., participants variably reporting their subjective awareness based either on the low- or high-level stimulus representations in either of the task), explicit instruction were given to participants on what “clarity” meant in each task. Therefore, participants informed exclusively their subjective awareness of color in the low-level task and their subjective awareness of stimulus category in the high-level task.

Following the predictions by the LoP hypothesis, we expected a higher number of intermediate (i.e., ratings 2 and 3) ratings on the awareness scale for the low-level task, together with a linear (i.e., proportional) increase in accuracy levels for increasing visibilities. We would expect the higher level task to produce a more dichotomously distributed subjective awareness ratings and a more nonlinear increase on accuracy levels. As for the ERPs, we expected significant differences on awareness amplitudes for the low-level task at VAN time window, in line with previous studies showing this component to correlate with the early elementary sensation of the stimulus (Jimenez et al., 2018; Koivisto, Grassini, Salminen-Vaparanta, & Revonsuo, 2017),
whereas the LP would index the awareness of high-level stimulus representation (i.e., the category of the stimulus; high-level task), in line with previous results suggesting that each ERP component reflect different aspects of visual awareness (Koivisto et al., 2016, 2017; Pitts, Padwal, Fennelly, Martinez, & Hillyard, 2014). Specifically, we expected the amplitudes to be more negative for the aware compared with less aware and/or unaware conditions at VAN in the low-level task and differences in LP amplitudes between aware (more positive) and unaware/less aware conditions in the high-level task.

METHODS

Participants

Twenty-six students (22 women, age range = 18–28 years, \( M = 22 \) years; \( SD = 3.51 \) years) from Universidad Complutense de Madrid with normal or corrected-to-normal vision took part in this study. The experiment was conducted with the understanding and written consent of each participant, in accordance with the Declaration of Helsinki, and accepted by the local ethics committee.

Stimuli and Apparatus

The target stimuli were six drawings (three objects: cap, iron, umbrella; three animals: frog, mouse, pig) adapted from the Snodgrass and Vanderwart (1980) picture set (see Figure 1A). The different drawings’ length and height were matched to 300 \( \times \) 150 pixel using Adobe Photoshop, and they subtended a visual angle of 6.96° in length and 3.44° in height. All filled-in parts in the different drawings were removed, so the target stimuli were all line drawings. Two versions of each target were created by rotating them on the horizontal axis (to the left if they originally faced right and vice versa) to avoid orientation-based target discriminations. All drawings could appear in six different colors, either “blue/bluish” (RGB: 178, 210, 255; 191, 239, 255; 196, 201, 255) or “red/reddish” (RGB: 255, 204, 221; 255, 191, 191; 255, 204, 178), for a total of 72 different stimuli. The stimuli were selected from a preliminary pilot study (\( n = 10 \)) based on accuracy performances, in which a larger set of eight stimulus per category (objects: airplane, anchor, balloon, bow, cap, iron, pot, and umbrella; animals: croc, duck, frog, mouse, pig, snail, squirrel, and turtle) was used. The six final drawings selected for the study were those that produced the same difficulty in both tasks, as measured by mean accuracy levels. A number of catch trials (blank stimulus) were introduced to verify that participants were correctly following the instructions.

The masks consisted of straight and curved colored lines of the same hues used for the target stimuli and were of the same length and height (300 \( \times \) 150 pixels). Two different masks were created, which were rotated in the horizontal and vertical axes, for a total of eight different masks (see Figure 1B).

Figure 1. (A) Example of targets. (B) Example of masks.
Procedure and Design

Two tasks (color identification and category) were performed by participants in two different blocks in a counterbalanced order. In both tasks, each trial began with a fixation point in the center of the screen for 500 msec, followed by a blank screen for a variable duration (between 307 and 507 msec, in one frame rate—13 msec—differences). After that, the target (or a blank screen in catch trials) appeared for an individually calibrated duration. Each target stimulus appeared equally often, but in random order. The target was followed by the mask for a variable time (see Figure 2A and B) to maintain the duration of the target-mask combination fixed at a duration of 706 msec. In the forced-choice color identification task, participants indicated the color of the target (blue/bluish vs. red/reddish) by clicking with the mouse on one of two horizontally oriented rectangles (see Figure 2A). In the forced-choice category identification task, participants indicated whether the target was an object or an animal by clicking with the mouse on one of two horizontally oriented rectangles (see Figure 2B). The forced-choice task was followed by the subjective awareness screen, in which participants indicated their subjective awareness by clicking with the mouse on one of the four possible rating-level alternatives (the names of the rating levels were displayed in four horizontally oriented rectangles, where 1 = no perception, 2 = weak perception, 3 = almost clear perception, and 4 = clear perception). Crucially, and to avoid possible “criterion shifts” in the use of the awareness scale between tasks, the meaning of “clarity” was exhaustively defined for each task. Specifically, in the color task, participants were informed that their subjective visibility should be reported exclusively based on their perception of the color of the stimulus (i.e., rating 1 = no perception of the stimulus, rating 2 = weak perception of the color of the stimulus, rating 3 = almost clear perception of the color of the stimulus, and rating 4 = clear perception of the color of the stimulus), whereas in the category task, they should exclusively inform their subjective awareness of the category of the target (i.e., rating 1 = no perception of the stimulus, rating 2 = weak perception of the category of the stimulus, rating 3 = almost clear perception of the category of the stimulus, and rating 4 = clear perception of the category of the stimulus). A blank intertrial screen then appeared for 800 msec. Each task started with a practice block of five trials, followed by four blocks of 97 trials each, yielding a total of 388 experimental trials in each task, of which 100 were catch trials.

Before the experimental session, the duration of the stimulus was calibrated individually so that no perception (rating 1) of the stimulus was reported in 20%–40% of the trials. During the calibration phase, the procedure was the same as that in the experimental phase, except that no forced-choice task had to be performed. Participants were informed that one target stimulus would appear in the screen followed by a mask (all possible stimuli and the sequence of events was shown to each participant beforehand), and they had to inform their subjective awareness on the 4-point scale. Because the aim of the calibration phase was to find a specific number of trials where participants did not see the stimulus at all (i.e., 20%–40% of rating 1), no explicit instructions were given to participants in relation

Figure 2. Sequence of events: (A) color task and (B) category task.
to the level of processing of the stimulus. Specifically, participants were instructed to use rating 1 (no perception) when they did not see the stimulus at all, rating 2 (weak perception) when they had a weak perception of a stimulus appearing (i.e., seeing something), rating 3 (almost clear perception) when they saw the stimulus almost clearly, and rating 4 (clear perception) when they saw the stimulus with total clarity. The calibration started with a short block of 10 trials, followed by up to three further blocks of 30 trials each. In the initial short block, the target duration was 66 msec. If the number of no perception rating was lower than 4, the duration was decreased in 1 frame rate (13 msec), and thus, the target was presented for 53 msec. If the number of no perception rating was 4 or higher in this first block, the next block was performed with the same duration of 66 msec. After the initial short block, the subsequent longer (30 trial) blocks were conducted until the participant informed no perception of the target in 20%–40% of the trials (i.e., between 6 and 12 trials out of 30). If participants reported more than 40% of no perception rating in one of these blocks, the target duration was increased in 1 frame rate (i.e., 13 msec) in the next block. If a participant reported less than 20% of no perception rating in a particular block, the target duration was decreased in 1 frame rate in the following block. If a participant reported between 20% and 40% of no perception rating, the calibration ended, and the duration of this block was used in the experimental session (this scenario occurred with 13 of the 26 participants). If a participant reached the last block and reported more than 40% of no perception rating, the duration in the actual experiment was increased in 1 frame rate (eight participants). If a participant reached the last block and reported less than 20% of no perception rating, the duration in the actual experiment was decreased in 1 frame rate (five participants). Hence, the durations used in the actual experiment were as follows: 66 msec (1 participant), 53 msec (5 participants), 40 msec (11 participants), 27 msec (8 participants), and 13 msec (1 participant).

EEG Recordings and Preprocessing

The EEG signal was continuously recorded using a cap with 64 Ag/AgCl small size electrodes (50–55 mm) mounted in an electrode cap (Quick-Cap, Neuroscan, Inc.), arranged according to the International 10–20 system. All electrodes were online referenced to the left mastoid. Participants were asked to avoid sudden muscle movements, changes of posture, jaw clenching, or lateral eyes movements. Bipolar horizontal and vertical electrooculograms were also recorded to monitor eye movements and blinks. Electrode impedances were kept below 10 kΩ. Recordings were amplified using Neuroscan SynAmplifiers amplifiers, continuously digitized at a sample rate of 1000 Hz, and filtered online with a 0.01–100 Hz band-pass filter.

EEG data were analyzed with the Fieldtrip software package (www.ru.nl/fcdonders/fieldtrip/), implemented in a MATLAB environment (The MathWorks). The continuous sets of raw data were downsampled to 250 Hz, referenced to the average of the two mastoids and segmented into −200 to 1200 msec epochs around the presentation of the stimulus. An infomax independent components analysis (Makeig, Jung, Bell, Ghahremani, & Sejnowski, 1997) was then performed to eliminate the eye blink activity (Jung et al., 2000). Finally, epochs contaminated with gross artifacts were rejected, following a visual inspection criterion. The signal was downsampled filtered with a low cutoff at 30 Hz and averaged separately for each condition and participant.

Based on well-established previous evidence (e.g., Förster et al., 2020; Tagliabue et al., 2019; Koivisto & Revonsuo, 2010) and a visual inspection of grand-averaged ERPs and scalp maps (see EEG section), VAN (the negative amplitude difference between aware and unaware trials) was most visible in 180–500 msec time windows in occipital electrodes (O1, PO3, P07, O2, PO8, PO4). The 180–300 time window in occipital electrodes corresponds to that observed in previous studies (see Förster et al., 2020; Koivisto & Revonsuo, 2010, for reviews) and was therefore selected. The delay between 300 and 500 msec has not been usually observed; therefore, we took both the time windows (180–300 msec vs. 300–500 msec) as a variable in the analyses. The LP (the positive amplitude difference between aware and unaware trials) peaked in the 700–1000 msec time window in central–parietal electrodes (P2, PZ, P1, CPZ, CP2, CP1).

Because of the small number of trials where participants used PAS4 (color task: $M = 16$, $SD = 39.8$; category task: $M = 46.7$, $SD = 56.4$), PAS3 and PAS4 were pooled in both tasks, for which three different ERPs were obtained (i.e., PAS1 = no perception, PAS2 = weak perception, PAS3 = almost clear/clear perception). Conditions with >29 trials were only introduced to the data analyses. Because of a technical error during the EEG recording process, data from one participant were not included in the final analysis.

Data Analyses

Awareness reports, accuracy, and ERP data analyses were carried out on RStudio (Version 1.3.959, RStudio, 2020) based on the R programming language (Version 4.0.2, R Core Team, 2020) using linear mixed effects models to avoid the listwise deletion of missing values (missing data for unreported awareness levels). Linear mixed-effect models were fitted to the data through the lme4 package (Bates et al., 2018), and statistical significance was assessed with the lmerTest package (Kuznetsova, Brockhoff, & Christensen, 2017). Because we were interested in how the level of processing and awareness interact, Task (two levels), PAS (three levels), and their interactions were always introduced as fixed effects to the models. Task and PAS were coded as factors, with color condition and PAS1 as the reference categories. To explore whether differences in PAS ratings between tasks could be due to accuracy differences, accuracy levels were introduced as covariate
into the model \([\text{PASProportion} \sim \text{Task} \times \text{Pas} + \text{ACC} + (\text{Pas} + \text{Subject})]\).

Accuracy data were analyzed as binomial (correct = 1, incorrect = 0) using the generalized linear mixed model and the binomial function \([\text{accuracy} \sim \text{PAS} \times \text{Task} + (\text{PAS} | \text{Participant})\), family = binomial]. The model with random intercept and PAS slope for participants as random effect did not converge. The model was therefore fitted with the \text{bobyqa} optimizer (Powell, 2009). This mode had a better fit than the model with only random intercept for participants. To explore the relation between accuracy and subjective rating categories, trend (i.e., polynomial) analyses were carried out (see Jimenez et al., 2018, 2019, for similar analyses). Specifically, a significant linear trend shows that accuracy performances increase proportionately across the different awareness categories, thus suggesting a graded pattern of awareness. If quadratic (or cubic) trends test significant, however, they signal at least one curve in the pattern, therefore showing a non-proportional change in objective accuracy levels across awareness categories (Maxwell & Delaney, 2004). For analyzing the trends, PAS was coded as an ordered factor.

For the ERP data, we first compared the fit of different models with the maximum likelihood estimation as false premise (REML = F) and Akaike information criterion. For both VAN and LP, the models with random intercept and PAS slope for participants as random effect had a better fit than the models with random intercept for participants. In addition, for VAN, the model with VAN time window (180–300 vs. 300–500 msec) as a fixed effect (in addition to task and PAS) fitted the data better than the model including also the interactions of the time window as fixed effects. Thus, the best fit model for VAN included VAN time window as a fixed effect, PAS and Task with their interactions as fixed effects, and random PAS slope for participants as random effect \([\text{amplitude} \sim \text{PAS} \times \text{Task} + \text{Time window} + (\text{PAS} | \text{Participant})\). For the LP, the corresponding best fitting model was \([\text{amplitude} \sim \text{PAS} \times \text{Task} + (\text{PAS} | \text{Participant})\). These final models were run with the restricted maximum likelihood estimation.

**RESULTS**

**Behavioral**

In the color identification task, *weak perception* (rating 2) was the most used rating (46% of the trials; see Table 1 and Figure 3A). *Almost clear perception* (rating 3) was also frequently used (18%), intermediate ratings (i.e., aggregated ratings 2 and 3) being reported in 64% of the trials (see Figure 3B). When catch trials were presented, participants reported *no perception* in most of the trials (rating 1: 89% of the trials, rating 2: 10%, rating 3: 1%, rating 4: 0%), thus showing that participants used the scale correctly.

In the category discrimination task, *weak perception* (rating 2) was reported in 32% of the trials, and *almost clear perception* (rating 3) was also frequently used in 18% of the trials. In combination, intermediate ratings were reported in 50% of the trials (see Figure 3B). When catch trials were presented, participants reported *no perception* in most of the trials (91% of the trials, rating 2: 9%, rating 3: 1%, rating 4: 0%).

To explore differences in the distribution of the awareness ratings between tasks, a linear mixed-effect model was fitted to the data. Task (two levels) and PAS (four levels) were introduced as fixed effects, and Participant was introduced as the random effect. Because overall accuracy differed between the tasks (color task: \(M = .69, SD = .15\); category task: \(M = .75, SD = .17\), \(t(25) = -2.95, p = .007\), accuracy levels were introduced as covariate into the model to explore whether the differences in PAS ratings between tasks would be due to task difficulty rather than the level of processing of the tasks. Results showed that the effect for Task was not significant (\(\beta = 2.92, SE = 3.36\), \(t(103.82) = 0.87, p = .386\), thus suggesting an absence of differences on the use of PAS1 between tasks. In the color task, PAS2 (\(\beta = 17.29, SE = 7.34\), \(t(33.59) = 2.36, p = .024\), and PAS4 (\(\beta = -26.51, SE = 8.87\), \(t(38.71) = -2.96, p = .005\), were significant, showing that these awareness ratings were used more (PAS2) and less (PAS4) frequently than PAS1 (see Figure 3A). In addition, Task × PAS2 (\(\beta = -16.53, SE = 4.73\), \(t(103.60) = -3.49, p < .001\), and Task × PAS4 (\(\beta = 15.84, SE = 5.47\), \(t(107.80) = 2.89, p = .004\), significant interactions suggested that, in the category task, PAS2 ratings were used less often and PAS4 ratings were more frequently used than in the color task (see Figure 3A). Interestingly, Accuracy was not significant (\(\beta = -5.16, SE = 7.48\), \(t(123.50) = -0.69, p = .491\), suggesting that differences in the use of the awareness scale were not explained by the difficulty of the task.

To further explore differences in the use of the awareness scale, \(t\) test for intermediate (ratings 2 and 3) and scale-end (ratings 1 and 4) reports were conducted. Results showed

| Task       | Awareness | Accuracy |
|------------|-----------|----------|
| Color      | 30%       | .52      |
|            | 46%       | .71      |
|            | 18%       | .85      |
|            | 6%        | .93      |
| Category   | 32%       | .56      |
|            | 32%       | .76      |
|            | 18%       | .94      |
|            | 18%       | .99      |

Downloaded from http://direct.mit.edu/jocn/article-pdf/33/7/1295/1921250/jocn_a_01712.pdf by guest on 28 June 2021
significant differences for both comparisons between tasks: intermediate, \( t(25) = 3.04; p = .006 \), and scale-end, \( t(25) = -3.38; p = .006 \) (see Figure 3B). Thus, in the lower level task, intermediate reports were used more frequently (mean use low-level task: 184 trials; mean use high-level task: 143 trials), whereas scale-end ratings were more frequently used in the high-level task (mean use high-level task: 144 trials; mean use low-level task: 104 trials), in line with the predictions by the LoP. Note, however, that intermediate ratings were significantly used in half of the trials (50%) in the high-level task.

Individual accuracy performances at different awareness levels were analyzed using the generalized linear mixed effects model to avoid the listwise deletion of missing values (missing accuracies for unreported awareness levels). A significant effect for Task (\( \beta = 0.18, SE = 0.06, z = 2.83, p = .004 \)) suggested different accuracies between tasks at PAS1, indicating that accuracy was higher in the category than in the color task at PAS1. Significant PAS2 (\( \beta = 0.98, SE = 0.15, z = 6.25, p < .001 \)), PAS3 (\( \beta = 2.19, SE = 0.29, z = 7.35, p < .001 \)), and PAS4 (\( \beta = 3.61, SE = 0.57, z = 6.31, p < .001 \)) effects showed that accuracies at these awareness levels were all different from PAS1 accuracy in the reference condition (i.e., color discrimination task; see Table 1 and Figure 3C). Task × PAS3 (\( \beta = 1.05, SE = 0.17, z = 5.99, p < .001 \)) and Task × PAS4 (\( \beta = 1.74, SE = 0.49, z = 3.51, p < .001 \)) significant interactions suggested that the category versus color difference (that was already present at PAS1) was even stronger at PAS3 and PAS4. Refitting the model with the category condition as the reference showed that accuracies for PAS2 (\( \beta = 0.98, SE = 0.15, z = 6.25, p < .001 \)), PAS3 (\( \beta = 0.98, SE = 0.15, z = 6.25, p < .001 \)), and PAS4 (\( \beta = 0.98, SE = 0.15, z = 6.25, p < .001 \)) were significantly different than PAS1 accuracy (see Table 1 and Figure 3D).

Accuracy was introduced as an ordered factor to the model to conduct trend analyses for each task. Both color and category tasks produced significant linear trends (color: \( \beta = 2.69, SE = 0.39, z = 6.77, p < .001 \); category: \( \beta = 4.09, SE = 0.42, z = 9.67, p < .001 \)), suggesting that accuracy levels increased proportionately across the awareness scale in both tasks. In addition, quadratic and cubic trends approached significance in the higher level category task (quadratic: \( \beta = 0.56, SE = 0.29, z = 1.92, p = .054 \); cubic: \( \beta = -0.31, SE = 0.16, z = -1.92, p = .054 \)). A visual inspection of categorization performances across the

![Figure 3](http://direct.mit.edu/jocn/article-pdf/33/7/1295/1921250/jocn_a_01712.pdf)
awareness scales showed that accuracy levels increased somehow more abruptly from awareness rating 2 to awareness rating 3, consistent with the interactions found in non-trend analysis.

Overall RTs between conditions were very similar (color: $M = 2312$ msec, $SD = 461.90$; category: $M = 2238$ msec, $SD = 295.99$), even though participants were not instructed on the speed of their response. A paired-sample $t$ test showed no differences on RTs between tasks, $t(24) = 0.74$, $p = 0.464$.

In summary, the levels of processing manipulation produced significant differences in both the awareness rating distributions and accuracy performances between tasks. Specifically, intermediate ratings were used more frequently in the low-level task and extreme ratings in the high-level task (yet they also were significantly used in this task), mainly due to rating 2 (weak perception) being used predominantly in the low-level task and rating 4 (clear perception) more used in the high-level task. Accuracy levels, on the other hand, increased following a linear trend at stimulus lower level of representation (i.e., color task) and followed a more nonlinear pattern at the higher level category task, accuracies increasing more abruptly from rating 2 to rating 3 and possibly showing a ceiling effect at rating 3, results overall following the LoP.

**EEG**

Based on the visual inspection of grand averages and scalp maps (see Figures 4 and 5), data for VAN (180–500 msec time window; occipital electrodes: O1, PO3, P07, O2, PO8, PO4) and LP (700–1000 msec time window, central–parietal electrodes: P2, PZ, P1, CPZ, CP2, CP) components were analyzed for stimulus-present trials. To test whether amplitudes differed between VAN and LP time windows, we first ran the model introducing ERP component (i.e., VAN vs. LP) as a factor. Results showed a main effect for Component ($\beta = 6.66$, $SE = 0.69$), $t(276.25) = 9.63$, $p < .001$, showing that amplitudes were more positive in LP than VAN. A significant two-way interaction, Task × Component ($\beta = -2.39$, $SE = 0.95$), $t(276.25) = 2.51$, $p = 0.012$, and a further three-way interaction, Task × Component × PAS3 ($\beta = 3.31$, $SE = 1.40$), $t(276.25) = 2.36$, $p = 0.019$, suggested specific divergences between...
the two awareness components, for which separate statistical analyses were carried out.

\textit{VAN: 180–300 and 300–500 msec}^1

The effect for Task was not significant ($\beta = -0.24, SE = 0.27$), $t(175.93) = -0.88, p = .377$, suggesting an absence of differences on PAS1 amplitudes between tasks. In the color task, the effects of PAS2 ($\beta = -0.87, SE = 0.26$), $t(103.85) = -3.25, p = .001$, and PAS3 ($\beta = -0.92, SE = 0.42$), $t(42.75) = -2.159, p = .036$, were significant, showing that the amplitudes associated to these awareness ratings were more negative than those at PAS1 (see Figure 6). In addition, overall amplitudes at the second

\textbf{Figure 5.} Topography maps for each awareness condition (PAS1, PAS2, and PAS3) and component (VAN1, VAN2, and LP) showing occipital (VAN1 and VAN2) and central-parietal (LP) electrodes.

\textbf{Figure 6.} ERP amplitudes associated to different PAS ratings in the VAN time window (O1, PO3, P07, O2, PO8, PO4 electrodes). Error bars represent 95% confidence interval.
VAN interval (i.e., 300–500 msec) were significantly more positive than the amplitudes at the first VAN interval ($\beta = 0.48, SE = 0.14$), $t(174.84) = 3.37, p < .001$. Task × PAS2 ($\beta = 1.19, SE = 0.34$), $t(178.38) = 3.44, p < .001$, and Task × PAS3 ($\beta = 0.84, SE = 0.39$), $t(179.69) = 2.13, p = .034$, significant interactions suggested that awareness had a difference effect depending on the task. Refitting the model with the category condition as the reference did not find any effect for PAS in the category task. To evaluate further differences between PAS amplitudes in the color task, the model was rerun introducing PAS2 as reference. Results showed that PAS1 differed from PAS2 ($\beta = 0.87, SE = 0.26$), $t(103.85) = 3.25, p = .001$, but PAS3 did not differ from PAS2 ($\beta = -0.04, SE = 0.36$), $t(38.56) = -0.13, p = .898$.

**LP: 700–1000 msec**

In the LP time window, the effect for Task was significant ($\beta = -2.39, SE = 1.11$), $t(80.90) = -2.16, p = .033$, showing that the amplitude of PAS1 in the category task was less positive than PAS1 amplitude in the color task (see Figure 7). In the color task, neither the effects of PAS2 nor PAS3 were significant, indicating an absence of differences on the amplitudes associated to the different conditions.
awareness categories. A Task × PAS3 (β = 4.11, SE = 1.66), t(83.18) = 2.48, p = .015, interaction suggested that awareness had a difference effect, depending on the task. Refitting the model with the category condition as the reference showed a significant effect for PAS3 (β = 4.63, SE = 1.17), t(71.37) = 3.93, p < .001, but not for PAS2 (β = 1.75, SE = 1.01), t(82.23) = 1.73, p = .087, indicating that only the amplitude associated to PAS3 was statistically significantly more positive than that at PAS1 in the category task.

### DISCUSSION

In this work, we aimed to gather further evidence on the behavioral and electrophysiological patterns associated with different levels of stimulus processing by introducing line drawings and semantic categorization for the first time in LoP research. Importantly, to avoid “criterion shifts” confounds in the use of the awareness scale between tasks (i.e., participants variably reporting their subjective awareness based either on the low- or high-level stimulus representations in either of the task), explicit instruction was given to participants on what “clarity” meant in each task. Stimulus’ low-level of processing was conceptualized as identifying its color, whereas high-level of processing consisted of discriminating whether it belonged to animal or object categories. Interestingly, the LoP manipulation produced significant differences in both behavioral and ERP results.

Behavioral data showed that the levels of processing manipulation produced significant differences in both the awareness rating distributions and accuracy performances between tasks. Specifically, intermediate ratings were used more frequently in the low-level task and scale-end ratings more frequently in the high-level task, mainly due to rating 2 (weak perception) being used predominantly in the low-level task and rating 4 (clear perception) more used in the high-level task. Accuracy levels, on the other hand, increased, following a linear trend at stimulus lower level of representation (i.e., color task) and followed a more nonlinear pattern at the higher level category task, results overall in agreement with the LoP. Whereas previous behavioral evidence has already shown support for the LoP postulates (most significantly when comparing a low-level color task and a more complex high-level numerical judgment task; see Jimenez et al., 2020, for a review), the differences between levels of processing has been usually weak (or even absent, see Jimenez et al., 2018) and not consistently observed in the same awareness measures. They were sometimes noticeable as a higher use of intermediate ratings in the subjective awareness scale (Binder et al., 2017; Anzulewicz et al., 2015) or as significant differences in accuracies with no differences in the awareness scale distribution (Derda et al., 2019; Windey et al., 2014). In this study, significant evidence in both subjective and objective measures of awareness was found when comparing a low-level color task with a high-level category discrimination task, in line with the LoP postulates. Note, however, that the high-level task also produced a significant number of middle ratings (50% of the reports) and intermediate accuracies related to weak perception reports (i.e., rating 2 = 0.76). Overall, it might be argued that the low-level stimulus processing emerges more gradually to awareness than higher level stimulus perception, yet high-level stimulus processing would not be strictly dichotomous, rather it would transition to awareness less gradually than low-level stimulus perception. Note, however, that the LoP predictions usually produce a rather qualitative result interpretation (e.g., higher number of intermediate ratings, a more dichotomously distributed awareness reports). Yet, how high is high enough to evidence for/against the postulates of the LoP hypothesis? At which specific point an awareness distribution becomes dichotomous? We consider that, to allow for an explicit confirmation or refutation of the LoP postulates in future research, these postulates should be clearly defined in a quantitative manner.

The electrophysiological recordings revealed two correlates of visual awareness: enhanced posterior negativity in the N200 time window (VAN) and enhanced positivity in the P3 time window (LP). In this study, nevertheless, both VAN and LP components were delayed compared with other studies (see Förster et al., 2020, for a review). Note, however, that specific conditions between studies can influence the presence and characteristics of the different components associated to visual awareness (Rutiku, Aru, & Bachmann, 2016). For instance, the latency of processes correlating with consciousness may shift depending on stimulus predictability (Melloni, Schwiedrzik, Müller, Rodriguez, & Singer, 2011). Furthermore, the specifics of the experimental setting or task requirements can influence ongoing prestimulus activity, which may have an effect on subsequent stimulus perception and possibly on the related...
markers of awareness (Busch, Dubois, & VanRullen, 2009; Mathewson, Gratton, Fabiani, Beck, & Ro, 2009). In this study, VAN onset at 180 msec was consistent with previous evidence, but the differences between aware (i.e., PAS3) and unaware/less aware (i.e., PAS1/PAS2) amplitudes continued until about 500 msec. The P3/LP, on the other hand, started to build around 500 msec (i.e., immediately after VAN as usual; e.g., Koivisto & Revonsuo, 2010) and peaked in the 700–1000 msec time window. It might be argued that, in a scenario of a delayed P3 (as visible in this study), VAN would show up for a longer time in the ERPs. Following this explanation, strong P3s found in previous studies (Derda et al., 2019; Jimenez et al., 2018; Koivisto et al., 2016, 2017; Koivisto & Grassini, 2016; Salti, Bar-Haim, & Lamy, 2012; Del Cul et al., 2007) would override VAN amplitudes, because the activities from different sources are summed to the scalp electrodes. The delayed LP might be explained by two further factors. First, participants responded using the computer mouse clicking in the corresponding response option, as a consequence the observed RTs (i.e., ~2 sec in both tasks) being significantly longer than in previous studies. In this regard, the long target + mask duration (~750 msec) might have contributed to the delay of P3/LP if participants waited until mask offset to start higher cognitive processes. In addition, the P3 found in this study closely resembles a later ERP component found within perceptual decision-making research (Tagliaubre et al., 2019), called the centroparietal positivity (CPP). The CPP has revealed to be a correlate of the accumulation of sensory evidence, and it is closely related to subjective perceptual experience (Tagliaubre et al., 2019).

Interestingly, the electrophysiological correlates associated to the different degrees of awareness showed a dissociation between low- and high-level stimulus processing. In the VAN time window, awareness modulated the ERP waves only in the low-level color task, the amplitudes associated to increasing levels of awareness being significantly more negative than the amplitude associated to PAS1. Specifically, PAS2 and PAS3 amplitudes were significantly more negative than PAS1, but PAS3 did not differ from PAS2 (PAS3 = PAS2 < PAS1). In the high-level category task, awareness did not modulate VAN amplitudes. This pattern of results is consistent with Koivisto et al. (2017), where VAN was sensitive for conscious detection but not for conscious identification (which correlated with a later widespread activity), suggesting that VAN would be sensitive to low-level awareness per se. In addition, the fact that VAN was sensitive to the LoP manipulation converges with previous fMRI findings by Binder et al. (2017) and Jimenez et al. (2018), suggesting that the level of processing effect occurs in the visual cortex. Indeed, in Jimenez et al. (2018), the LoP manipulation showed that awareness levels modulated VAN amplitudes in a graded manner only for the low-level task (PAS3 < PAS2 < PAS1), whereas in the high-level task VAN amplitudes followed a nonlinear pattern (PAS3 = PAS2 < PAS1). Note, however, that this specific pattern of results is not in line with the present findings, the observed differences between the studies possibly explained by differences in the LoP manipulation. Whereas stimulus processing was defined as perception of “energy” (stimulus localization; low-level task) and “letters/numbers” (stimulus identification, higher level task) in Jimenez et al. (2018), it was operationalized as color discrimination (low-level) and category discrimination (high-level) tasks in this study. Furthermore, Jimenez et al. (2018) used a noncentral stimulus presentation (i.e., four possible locations equidistant to the central fixation), an experimental design that possibly involved shifts of attention as an additional source of variability between conditions. Consequently, attentional involvement probably modulated the ERP waves through enhanced feedback activity in early visual areas (see Carrasco, 2011, for a review).

In Derda et al. (2019), the LoP manipulation did not produce significant differences between conditions in ERP amplitudes at VAN time window, mean amplitudes in their study correlating with PAS ratings in both conditions. Thus, the results by Derda et al. neither converge with previous (Jimenez et al., 2018; Binder et al., 2017) nor present results, which show VAN as an index of low-level awareness of the stimuli. These results might be explained by the LoP manipulation in Derda et al. producing very subtle differences between tasks as observed in the behavioral measures (i.e., no significant differences were found in the mean visibility rating between conditions).

In the LP time window, awareness modulated the ERP waves only in the higher level category task, the amplitudes associated to increasing levels of awareness being significantly more positive. Specifically, PAS3 amplitude was more positive than PAS2 and PAS1 amplitudes, which did not differ from each other (PAS3 > PAS2 = PAS1). Interestingly, this pattern of results is in line with Derda et al. (2019), showing that mean amplitudes in the LP time window correlate with PAS only in the high-level condition, but not with results from Jimenez et al. (2018), where LP amplitudes were modulated in a graded manner for both low- and high-level tasks. Differences between studies might be possibly due to the nature of the tasks involved. In this study, as well as in Derda et al. (2019), the low-level tasks were operationalized as a color discrimination task, and the high-level tasks involved postperceptual judgments (magnitude judgments and category discriminations), whereas in Jimenez et al. (2018), the low-level task involved the location of the target and the high-level task involved the identification of letters or numbers, both being perceptual tasks. The findings in Jimenez et al. (2018) could be reconciled with the levels of processing framework by proposing that the more dichotomous nature of visual experience would only emerge at the highest level of stimulus processing, that is, with stimuli and tasks requiring access to meaning.

Indeed, it has been argued that conscious access to higher properties of the stimulus and postperceptual or decision-making processes correlate with the later P3 component (Koivisto et al., 2016, 2017; Pitts et al., 2014;
Koivisto & Revonsuo, 2010). Recent evidence using no-report paradigms (Cohen et al., 2020) and the manipulation of response criterion (Mazzi, Mazzeo, & Savazzi, 2020) provide further evidence on this account, showing that LP reflects the contribution from postperceptual processes related to response requirements. In Cohen et al. (2020), a standard visual masking paradigm was combined with the recently developed “no-report” paradigm (see Tsuchiya, Wilke, Frässle, & Lamme, 2015). In the visual masking paradigm, observers saw visible/invisible (depending on masking manipulation) images of animals and objects. On half of the trials, they reported the contents of their perceptual experience (i.e., report condition), whereas on the other half of trials, they avoided any report of their experience (i.e., no-report condition). The authors examined how visibility interacted with reporting by measuring the P3b ERP, one of the candidate ERP components of conscious processing. Overall, results showed a robust P3b in the report condition, but no P3b whatsoever in the no-report condition. In Mazzi et al. (2020), the authors manipulated the response criterion, which affects how a percept is translated into a decision. Specifically, participants performed an orientation discrimination task and were asked to shift their response criterion (inducing a “liberal” or “conservative” bias in Experiment 1, whereas asking participants to follow their natural response criterion in Experiment 2) across sessions. Following this manipulation, the resulting modulation would concern those ERP components not exclusively associated to the subjective conscious experience itself but also the processes accompanying it. Electrophysiological results showed that N1 and P3 were sensitive to the response criterion adopted by participants. When the data were considered independently from the response criterion, both VAN and LP correlated with awareness. Crucially, the LP component was also modulated by the interaction of awareness and response criterion, whereas VAN was not. Thus, in line with previous literature (Koivisto et al., 2016, 2017; Pitts et al., 2014; Koivisto & Revonsuo, 2010), these findings provide additional evidence supporting the hypothesis that VAN tracks the emergence of visual awareness, whereas LP reflects the contribution from postperceptual processes related to response requirements.

Whether the results by Cohen et al. (2020) and Mazzi et al. (2020) are in line with LP indexing postperceptual processes, the present results and those found in Derda et al. (2019) suggest that LP is also sensitive to decision-making processes. Indeed, both in Derda et al. (2019) and here, LP is sensitive to high-level stimulus representation, but not to low-level stimulus processing. As indicated in Derda et al. (2019), this sensitivity of LP amplitude to fluctuations in visibility in the high-level condition might be interpreted as an indicator of processes representing more abstract, semantic features of stimuli. This would be in line with the LP reflecting sensory accumulation and decision-making processes (i.e., CPP), a process that is neither specific to any particular sensory modality, feature, or motor requirements, and it is believed to reflect evidence accumulation at an intermediate, abstract level of processing (Tagliafame et al., 2019).

One might argue from the observed ERP results, however, that neither VAN nor LP does correlate with awareness. Indeed, the absence of significant differences in ERP amplitudes at VAN time window in the high-level category task could be interpreted as VAN not being a correlate of consciousness. Following the same logic, an analogical argument could be presented for the LP, because no significant differences were found between awareness levels at this time window in the low-level color task. We consider, however, that the interpretation of the present results should be done in the light of previous evidence on the neural correlates of consciousness, as well as taking into account the particularities of the LoP framework implemented here. To our knowledge, this is the first time that the meaning of “clarity” is exhaustively defined at each level of stimulus processing. Thus, in the color task, participants’ subjective visibility was reported exclusively based on their perception of the color of the stimulus, whereas in the category task, participants were instructed to inform their subjective awareness of the category of the target. Taking this important methodological factor into account, the overall pattern of results would suggest that awareness of low-level features is reflected in VAN, whereas awareness of high-level features is reflected in LP.

The complexity of this process is best illustrated by the microgenetic approach (Bachmann, 2000). According to this, the formation of a percept is a gradual process evolving from an initial basic content, which then “matures” by acquiring systematically more qualities to the preceding version of the percept through a time-consuming process (Aru & Bachmann, 2017). Interestingly, conscious experience of the same object would change considerably over time (see also Breitmeyer, 2014; Pitts et al., 2014; Hegdé, 2008). More recently (Aru & Bachmann, 2017), the possibility of two different processes has been proposed within the microgenesis of conscious perception: a “perceptual microgenesis,” where the conscious experience would emerge fast and decay fast (possibly equal to iconically-memorized decay), and an “immediate memory-based microgenesis,” where the conscious experience of the same stimulus would form slower than in perceptual microgenesis and would decay later (Sligte, Scholte, & Lamme, 2008). These two different microgeneses would fit the distinction between “phenomenal consciousness,” the subjective experience of having qualitative experiences (such as the perception of color; Block, 1995) and “reflective” or “access consciousness,” a system’s central access to the contents of phenomenal consciousness, contents that have been attentionally selected for further cognitive processing (such as category discriminations in the current study) in working memory. Interestingly, VAN has been proposed to be the electrophysiological correlate of phenomenal visual consciousness, a sufficient
electrophysiological signature of a visual stimulus producing a corresponding subjective experience; whereas the later timing of LP suggest that it indexes the stimulus information accessing reflective consciousness ( Förster et al., 2020; Koivisto & Revonsuo, 2010). Following this distinction, both early and late ERP components (i.e., VAN and LP) might reflect not only the different dimensions of consciousness (i.e., phenomenal vs. reflective) but the different stages of the microgenetic process, the former associated to the “perceptual microgenesis” whereas the latter associated to “immediate memory-based microgenesis.”

Finally, we compared the ERPs associated to an absence of stimulus perception when a stimulus was present (i.e., StimulusPresent-PAS1) or absent (i.e., CatchTrials-PAS1), for which task manipulation did not show any effect. We consider this result is important in showing an absence of bias effects between tasks. Regarding the direct comparison between StimulusPresent-PAS1 and CatchTrials-PAS1, different scenarios could be initially predicted. First, an absence of differences between the two conditions might suggest that StimulusPresent-PAS1 reflects an absolute unawareness of the stimulus. Yet, differences in brain function between CatchTrials-PAS1 and StimulusPresent-PAS1 could be also expected, possibly due to physical differences between the two conditions. Indeed, this is what we found: Stimulus-present amplitudes were more positive/less negative than stimulus-absent amplitudes in both tasks, which was only statistically significant at VAN time window. One of the plausible interpretations of these results (note that the interpretation of these results is only tentative, because this study was not designed to explore differences in ERPs between stimulus-present and stimulus-absent trials) would be that differences in brain activity between stimulus-present and catch trials are due to differences in physical stimulation between conditions, as already mentioned; additionally, a more positive wave for the StimulusPresent-PAS1 condition could possibly reflect an unconscious processing of the stimulus; finally, it might be suggested that the stimulus-present condition reflects some residual awareness of the stimulus. Even if the difference between these two conditions might be due to residual awareness of the stimulus (e.g., due to a conservative response criterion in StimulusPresent-PAS1 trials), that would not be of relevance to this study because our aim was to explore the relation between different degrees of visual awareness at different levels of stimulus processing; hence, an absolute absence of awareness at StimulusPresent-PAS1 condition is not a necessary assumption.

In summary, in this study, we showed that the LoP manipulation produced differences between the low- and high-level stimulus processing at both behavioral (i.e., awareness rating distribution and accuracy performances) and electrophysiological levels. On a behavioral level, it might be argued that the low-level stimulus processing emerges “more gradually” to awareness than higher level stimulus perception. On an electrophysiological level, the ERP analyses showed a double dissociation between awareness and the LoP manipulation: Awareness modulated the ERP waves only in the low-level color task at VAN time window, whereas it modulated the ERP waves only in the higher level category task in the LP time window. These findings are compatible with a two-stage microgenesis model of conscious perception (Aru & Bachmann, 2017), where an early elementary phenomenal sensation of the stimulus (i.e., the subjective perception of color) would be indexed by VAN and the stimulus’ higher level properties (i.e., the category of the target) would be indexed by the LP in a later latency range.
awareness. *Journal of Neuroscience*, 31, 1386–1396. DOI: https://doi.org/10.1523/JNEUROSCI.4570-10.2011, PMID: 21273423, PMCID: PMC6623527

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, 700–708. DOI: https://doi.org/10.1016/j.concog.2006.04.002, PMID: 16725347

Pitts, M. A., Padwal, J., Fennelly, D., Martinez, A., & Hillyard, S. A. (2014). Gamma band activity and the P3 reflect post-perceptual processes, not visual awareness. *Neuroimage*, 101, 337–350. DOI: https://doi.org/10.1016/j.neuroimage.2014.07.024, PMID: 25063751, PMCID: PMC4169212

Powell, M. J. D. (2009). The BOBYQA algorithm for bound constrained optimization without derivatives (Report No. DAMTP 2009/06). Cambridge, United Kingdom: Centre for Mathematical Sciences, University of Cambridge.

Pretorius, H., Tredoux, C., & Malcolm-Smith, S. (2016). Subjective awareness scale length influences the prevalence, not the presence, of graded conscious states. *Consciousness and Cognition*, 45, 47–59. DOI: https://doi.org/10.1016/j.concog.2016.08.007, PMID: 27567528

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

Rutiku, R., Aru, J., & Bachmann, T. (2016). General markers of conscious visual perception and their timing. *Frontiers in Human Neuroscience*, 10, 23. DOI: https://doi.org/10.3389/fnhum.2016.00023, PMID: 26869905, PMCID: PMC4740392

Rutiku, R., & Bachmann, T. (2017). Juxtaposing the real-time unfolding of subjective experience and ERP neuromarker dynamics. *Consciousness and Cognition*, 54, 3–19. DOI: https://doi.org/10.1016/j.concog.2017.05.003, PMID: 28554797

Sali, M., Bar-Haim, Y., & Lamy, D. (2012). The P3 component of the ERP reflects conscious perception, not confidence. *Consciousness and Cognition*, 21, 961–968. DOI: https://doi.org/10.1016/j.concog.2012.01.012, PMID: 22341937

Sandberg, K., Timmermans, B., Overgaard, M., & Cleeremans, A. (2010). Measuring consciousness: Is one measure better than the other? *Consciousness and Cognition*, 19, 1069–1078. DOI: https://doi.org/10.1016/j.concog.2009.12.013, PMID: 20135167

Sekar, K., Findley, W. M., Poeppel, D., & Llinás, R. R. (2013). Cortical response tracking the conscious experience of threshold duration visual stimuli indicates visual perception is all or none. *Proceedings of the National Academy of Sciences U.S.A.*, 110, 5642–5647. DOI: https://doi.org/10.1073/pnas.1302229110, PMID: 23509248, PMCID: PMC3619304

Sergent, C., & Dehaene, S. (2004). Is consciousness a gradual phenomenon? Evidence for an all-or-none bifurcation during the attentional blink. *Psychological Science*, 15, 720–728. DOI: https://doi.org/10.1111.j.0956-7976.2004.00748.x, PMID: 15482443

Sligte, I. G., Scholte, H. S., & Lamme, V. A. F. (2008). Are there multiple visual short-term memory stores? *PloS One*, 3, e1699. DOI: https://doi.org/10.1371/journal.pone.0001699, PMID: 18301775, PMCID: PMC2246053

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, 6, 174–215. DOI: https://doi.org/10.1037/0278-7393.6.2.174, PMID: 7373248

Tagliaabe, C. F., Veniero, D., Benwell, C. S. Y., Cecere, R., Savazzi, S., & Thut, G. (2019). The EEG signature of sensory evidence accumulation during decision formation closely tracks subjective perceptual experience. *Scientific Reports*, 9, 4949. DOI: https://doi.org/10.1038/s41598-019-41024-4, PMID: 30894558, PMCID: PMC6426990

Tsuchiya, N., Wilke, M., Frässle, S., & Lamme, V. A. F. (2015). No-report paradigms: Extracting the true neural correlates of consciousness. *Trends in Cognitive Sciences*, 19, 757–770. DOI: https://doi.org/10.1016/j.tics.2015.10.002, PMID: 26585549

Windey, B., & Cleeremans, A. (2015). Consciousness as a graded and an all-or-none phenomenon: A conceptual analysis. *Consciousness and Cognition*, 35, 185–191. DOI: https://doi.org/10.1016/j.concog.2015.03.002, PMID: 25804704

Windey, B., Vermeiren, A., Atas, A., & Cleeremans, A. (2014). The graded and dichotomous nature of visual awareness. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences*, 369, 20130282. DOI: https://doi.org/10.1098/rstb.2013.0282, PMID: 24639587, PMCID: PMC3665170

Windey, B., Gevers, W., & Cleeremans, A. (2013). Subjective visibility depends on level of processing. *Cognition*, 129, 404–409. DOI: https://doi.org/10.1016/j.cognition.2013.07.012, PMID: 23978782