Is auditory awareness negativity confounded by performance?

Rasmus Eklund⁎, Billy Gerdfeldter, Stefan Wiens

Gösta Ekmans Laboratorium, Department of Psychology, Stockholm University, Stockholm, Sweden

ARTICLE INFO

Keywords:
Auditory awareness negativity
Late positivity
Consciousness
Pitch discrimination
Controlling for performance

ABSTRACT

Research suggests that the electrophysiological correlates of consciousness are similar in hearing as in vision: the auditory awareness negativity (AAN) and the late positivity (LP). However, from a recently proposed signal-detection perspective, these correlates may be confounded by performance, as the strength of the internal responses differs between aware and unaware trials. Here, we tried to apply this signal-detection approach to correct for performance in an auditory discrimination and detection task (N = 28). A large proportion of subjects had to be excluded because even a small response bias distorted the correction. For the remaining subjects, the correction mainly increased noise in the measurement. Furthermore, the signal-detection approach is theoretically problematic because it may isolate post-perceptual processes and eliminate awareness-related activity. Therefore, we conclude that AAN and LP are not confounded by performance and that the contrastive analysis identifies both as correlates of awareness.

1. Introduction

In research on awareness, a central goal is to isolate the neural activity that is associated with whether or not an individual becomes aware of a stimulus. In a typical task, stimulus parameters are manipulated until some stimuli are reported as aware whereas others are not (Kim & Blake, 2005). If neural activity is also recorded during stimulus presentation, the contrastive analysis of the difference in neural activity between awareness and unawareness reflects the neural correlates of consciousness (NCC, Aru, Bachmann, Singer, & Melloni, 2012). The logic of this contrastive analysis is that if the stimulus is physically identical across repeated stimulations, the only difference in neural activity between awareness and unawareness should reflect conscious processing (Frith, Perry, & Lumer, 1999).

This contrastive analysis has mainly been used to investigate the NCC in vision. From recordings of electroencephalography (EEG), two event-related potentials (ERPs) have been reported consistently as NCC (Koivisto & Revonsuo, 2010). The earlier of these ERPs is the visual awareness negativity (VAN), a negative difference wave about 200 ms after visual onset (Ojanen, Revonsuo, & Sams, 2003). The topography of VAN has a negative peak at occipital electrodes, and source localization suggests sources in early visual and lateral occipital cortex (Koivisto & Revonsuo, 2010). The later ERP is the late positivity (LP), a positive difference wave about 300 ms after visual onset with a positive peak at parietal electrodes (Wilenius & Revonsuo, 2007), with potential sources in fronto-parietal and temporal areas (Koivisto & Revonsuo, 2010).

Importantly, when we applied the contrastive analysis in hearing, we showed that the NCC in hearing resemble the VAN and LP in vision (Eklund & Wiens, 2019; Eklund, Gerdfeldter, & Wiens, 2019). In our studies, auditory tones were presented at the individual awareness threshold: The level at which a subject reports being aware of the tones half of the time. For the tones at threshold, subjects reported that they were weakly aware of half of the tones. Critically, the contrastive analysis between awareness and unawareness...
revealed a negative difference wave at around 200 ms after tone onset at frontocentral electrodes, and a positive difference wave 300 ms after tone onset at parietal electrodes (LP). For the early negativity, we named it auditory awareness negativity (AAN) because it resembles the VAN in vision. Taken together, the results from the contrastive analyses in vision and hearing are consistent with theories that early and late NCC are general across different sensory modalities (Dykstra, Cariani, & Gutschalk, 2017; Meyer, 2011; Snyder, Yerkes, & Pitts, 2015).

However, Lamy, Salti, and Bar-Haim (2009) argue that the contrastive analysis between awareness and unawareness is confounded by performance. This argument can be understood with the help of the high threshold model, illustrated in Fig. 1. As described by Morales, Chiang, and Lau (2015), the high threshold model includes two thresholds: A processing threshold and an awareness threshold. Stimuli below the processing threshold are not processed at all whereas stimuli above the processing threshold are processed either unconsciously (if they fall below the awareness threshold) or consciously (if they fall above the awareness threshold). In support of processing above the processing threshold but below the awareness threshold (i.e., unconscious processing), previous research suggests that stimuli below the awareness threshold can be responded to correctly (Dehaene & Changeux, 2011). That is, even though subjects rate that they are unaware of the stimuli, they can perform better than chance when responding to them.

According to Lamy et al. (2009), stimuli rated as unaware mix together stimuli that were processed unconsciously (i.e., above the processing threshold but below the awareness threshold) and stimuli that were not processed at all (i.e., below the processing threshold). As a consequence, the contrastive analysis is confounded by performance because it captures the difference between conscious processing and a combination of unconscious processing and no processing.

To remove the confounding influence of stimuli that are not processed at all, Lamy et al. (2009) propose that the contrastive analysis should consider behavioral performance; that is, whether the stimulus was discriminated correctly. If performance to a stimulus is correct, it means that the stimulus was processed or it was guessed correctly. If only correct trials that are not guesses are considered, stimuli that are processed and rated as aware would be compared with stimuli that are processed and rated as unaware. By correcting for performance, this contrastive analysis would correctly isolate awareness.

To apply this correction, Lamy et al. (2009) used a discrimination and detection task with backward masked stimuli presented in the four visual quadrants. Subjects indicated in which quadrant the stimulus appeared and then rated their awareness of the stimulus. The interval between the stimulus and the mask was adjusted so that stimuli were reported as aware on 50% of the trials. Results showed that task performance across all trials rated as unaware was well above chance. This finding was taken as support of unconscious processing, that is, processing above the processing threshold but below the awareness threshold.

Lamy et al. (2009) applied the correction in two steps. In a first step, neural activity to stimuli that were correctly localized and reported as aware was compared with neural activity to stimuli that were correctly localized and reported as unaware. In a second step, Lamy et al. mathematically adjusted ERPs in the unaware correct condition for trials that were correct by chance. These trials are assumed to not be guided by unaware processes because processing fell below the processing threshold. The influence of these correct-by-chance trials was mathematically removed, leaving only chance-free trials in the unaware correct condition (i.e., trials that were localized correctly, presumably driven by unconscious processing). According to Lamy et al., the difference between aware correct trials and unaware correct (chance-free) trials isolates awareness.

When Lamy et al. (2009) conducted the contrastive analysis and controlled for performance, there was no statistically significant effect for VAN, but there was for LP. Therefore, Lamy et al. argued that VAN reflects preconscious processing and that LP is the proper NCC. However, in a subsequent study (Koivisto & Grassini, 2016), both VAN and LP remained after correcting for performance. In our own study (Eklund & Wiens, 2018), we used a discrimination task and also found both VAN and LP in the contrastive analysis of aware correct versus unaware correct trials.

Critically, there are theoretical problems with the correction proposed by Lamy et al. (2009), as discussed by Morales et al. (2015). First, the correction assumes that stimuli fall above or below a processing threshold, as illustrated in the high threshold model.
(Fig. 1). In contrast, signal detection theory (SDT; Macmillan & Creelman, 2005) implies that the notion of a processing threshold is incorrect and that processing of a stimulus is not binary (i.e., processed above or below a presumed processing threshold). Instead, the repeated presentation of a physically identical stimulus elicits internal responses that vary in strength on a normal distribution. On every trial, the strength of the internal response varies because the internal processing of the stimulus is combined with internal noise. A strong internal response reflects a strong percept, and a weak internal response reflects a weak percept. Although SDT is not explicit about awareness, the threshold for a report of awareness may be reflected by the subject’s response criterion: If the internal response reaches the subject’s criterion for awareness, awareness will be reported (Macmillan, 1986). Alternatively, a stimulus could theoretically be processed with a fairly strong internal response but not reach the threshold that leads to a report of awareness.

Another theoretical problem with the high threshold model is that several studies reported that about 2–3% of the incorrect trials were rated as aware (Eklund & Wiens, 2018; Lamy et al., 2009). In terms of the high threshold model, these trials do not make any theoretical sense: If performance is incorrect, the stimulus was processed below both the processing threshold and the awareness threshold. Thus, incorrect trials should be reported as unaware rather than as aware. According to SDT, however, these trials can be explained as high internal noise resulting in a strong internal response that leads to a report of awareness.

To provide evidence that the correction by Lamy et al. (2009) is incorrect, Morales et al. (2015) conducted a simulation of ERP data for a two-alternative forced choice discrimination task that involved two stimuli (A and B) and dichotomous awareness ratings (yes or no). From the behavioral data, SDT indices of discrimination sensitivity and discrimination criterion were computed together with discrimination awareness thresholds (i.e., the subject reports awareness of stimulus A or B). Internal (perceptual) responses for individual trials were generated by sampling from the SDT distributions. For each trial, the amplitude of the simulated ERP was determined by the strength of the internal response, as it was assumed that “[neural] activation intensity is linearly determined by the internal response” (Morales et al., 2015, p. 7). For aware trials, an additional, awareness-related ERP activity was added to the simulated ERPs. Afterwards, trials were sorted into aware correct, unaware correct, aware incorrect, and unaware incorrect trials. Importantly, when the correction by Lamy et al. was applied to the simulated ERP data, it did not isolate the awareness-related ERP activity and was thus deemed incorrect.

Instead of the correction suggested by Lamy et al. (2009), Morales et al. (2015) applied the perspective of SDT to correct for performance. In the model by Morales et al., the internal response reflects unconscious processing of the physical stimulus properties. To correct for this unconscious processing, the strength of unconscious processing in unaware correct trials needs to be identical to the strength of unconscious processing in aware correct trials: “By scaling up the weaker response in the unaware condition to approximately match the intensity of the stronger response in the aware condition, we can subtract away any activation due to magnitude difference in internal response” (Morales et al., 2015, p. 7). The scaling factor can be calculated from the behavioral data. According to Morales et al., the difference between aware correct trials and re-scaled unaware correct trials should isolate awareness-related ERP activity. In support, when this correction was applied to the simulated ERP data, it isolated the awareness-related ERP activity and was thus deemed incorrect.

Importantly, no previous study in hearing has examined whether AAN remains after correcting for performance. Therefore, the primary goal of the present study was to investigate whether AAN (and LP) can be obtained in the contrastive analysis after correction for performance. Similar to studies in vision, we used an auditory discrimination and detection task in which both perceptual and awareness-related ERP activity were measured. In support, when this correction was applied to the simulated ERP data, it isolated the awareness-related ERP activity and was thus deemed correct.

Another goal of this study was to replicate our previous findings because the data sets (with many channels and potential intervals) are huge and invite analytic flexibility (Luck & Gaspelin, 2017). Because this risk for false positives can be reduced with preregistration (Nosek, Ebersole, DeHaven, & Mellor, 2018), we preregistered the primary hypotheses, method, and analyses (osf.io/w4u7v). The preregistered electrodes and intervals for AAN and LP were selected on the basis of previous findings (Eklund & Wiens, 2019). Deviations from the preregistration are noted below. All data, scripts, and supplementary material are available at a university depository (Wiens and Eklund, 2020).

2. Method

In ERP research, there is a high risk for false positive findings because the data sets (with many channels and potential intervals) are huge and invite analytic flexibility (Luck & Gaspelin, 2017). Because this risk for false positives can be reduced with preregistration (Nosek, Ebersole, DeHaven, & Mellor, 2018), we preregistered the primary hypotheses, method, and analyses (osf.io/w4u7v). The preregistered electrodes and intervals for AAN and LP were selected on the basis of previous findings (Eklund & Wiens, 2019). Deviations from the preregistration are noted below. All data, scripts, and supplementary material are available at a university depository (Wiens and Eklund, 2020).

2.1. Participants

We preregistered to recruit at least 20 subjects. If the Bayes Factor (BF) exceeded 3 or was below $\frac{1}{3}$ for our hypotheses, recruitment would end. Otherwise, recruitment would continue for one more week. This process was repeated until the BF reached the criterion, a maximum of 60 subjects were tested, or at the end of February 2019 (osf.io/w4u7v).

The sample consisted of 28 healthy subjects (10 male; 25 right-handed; age: $M = 26.7, SD = 5.1$). Recruitment stipulated a target age range of 18–40 years, no history of neurological diseases, normal or corrected to normal vision, and normal hearing. Subjects were recruited from local universities and through online billboards and were compensated with either 100 SEK vouchers or course credits. Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Participants provided their written informed consent to participate in this study.

For the preregistered analyses, no subject was excluded on the basis of noisy EEG data. Although we did not preregister any other
exclusion criteria, three subjects were excluded because they lacked trials (less than 5) in one of the critical conditions. Further, one subject was excluded because performance when reporting awareness was below chance (14.5% vs. 50%). Without these trials, it was not possible to conduct the preregistered analyses. For the remaining subjects, mean performance when reporting awareness was 87.9% (SD = 9.5). Thus, a total of four subjects were excluded, and the final sample for the preregistered analyses consisted of 24 subjects (8 male; 21 right-handed; age: $M = 27.0, SD = 5.3$).

2.2. Apparatus and stimuli

Stimuli were two 100-ms tones ($f = 900$ Hz and $1400$ Hz, 5 ms fade-in and fade-out). Tones were presented through in-ear tubephones (ER2; Etymotic Research Inc., IL; www.etymotic.com). Instructions were displayed on a BenQ XL2430T, 24-inch gaming monitor (at 144 Hz, $1920 \times 1080$ resolution). PsychoPy v 1.85.3 (Peirce, 2007) was used to generate tones and to collect behavioral data. Auditory stimuli were sent as four channels with a sound card (RME Babyface). Channels one and two contained the calibrated tones. Channels three and four contained loud tones that were sent only to a Cedrus StimTracker (Cedrus Corporation, San Pedro, CA). The StimTracker detected the auditory onsets and marked them as triggers in the EEG recording. Thus, tone onsets could be detected precisely except for any constant delays from the equipment and cables.

Subjects rated their awareness on a modified version of the perceptual awareness scale (Sandberg, Timmermans, Overgaard, & Cleeremans, 2010). Our modified version had three rather than four levels because this has been used in related research in vision (Koivisto & Grassini, 2016), and we had good experiences with shorter versions (Eklund & Wiens, 2019; Eklund et al., 2019). In the present study, we eliminated the response option of an almost clear experience and allowed subjects to rate if they had no experience, a weak experience, or a clear experience. Thus, subjects had the full range available to rate their experiences. Because our focus was the distinction between no and weak experience, we felt it unnecessary to try to distinguish between weak, almost clear, and clear experiences. However, we kept the clear alternative to minimize the number of trials in which subjects had a weak experience but rated it as no experience because the scale lacked an option reflecting a weak experience.

2.3. Procedure

Subjects performed a discrimination and detection task of tones while seated in front of a computer screen with their chin in a chinrest. Fig. 2 shows the time course of a trial. During each trial, a black fixation cross (0.5 visual degrees) was displayed for 1000 ms. Critical trials contained a tone at the individual awareness threshold. On these trials, a low- or high-pitched tone was played binaurally 500 ms after trial onset. On catch trials, no tone was played. After the fixation cross disappeared, subjects indicated first if they heard a low- or high-pitched tone by using one of two buttons (“down arrow” and “up arrow”, respectively). Second, they rated their subjective awareness of the tone by using one of three buttons (“1”, “2”, and “3” corresponding to “I did not hear a tone”, “I heard the tone weakly”, and “I heard the tone clearly”, respectively). Then, the next trial immediately began. To discourage subjects from using a strategy in the discrimination task (e.g., always guessing low-pitch tone when not hearing the tone), we instructed subjects to always guess the correct tone pitch to the best of their ability and to avoid using any strategy when they did not hear the tone. Subjects were also instructed to focus on performing in the discrimination task and to rate their awareness accurately rather than respond quickly.

The task comprised 800 trials (720 critical and 80 catch). The trials were divided into eight blocks of 100 trials each (90 critical and 10 catch). The order of critical and catch trials was randomized within each set of 20 trials (9 low pitch, 9 high pitch, and 2 catch

![Fig. 2. The time course of a trial. On each trial, a black fixation cross (0.5°) was displayed on a gray background for 1000 ms. On critical and control trials, a tone (low or high pitch) was played binaurally 500 ms after trial onset. On catch trials, no tone was played. Afterward, subjects discriminated pitch (low or high) and rated their subjective awareness of the tone.](image-url)
2.4. EEG recording

Before the experiment began, subjects performed a short practice task. This task was identical to the main task but with clearly audible tones. After the practice task, interleaved staircases were used to calibrate the low-pitch and high-pitch tone to an intensity that the subject reported hearing either weakly or clearly on approximately 50% of the trials (i.e., individual awareness threshold for each pitch). For each pitch, the staircase procedure consisted of three staircases: One staircase started at the threshold estimate obtained from pilot subjects (4 dB), and the other two staircases started at 20 dB above and below this estimate. The staircase procedure was as follows: If the subject reported hearing a tone (weakly or clearly), the stimulus level decreased. If the subject reported not hearing a tone, the stimulus level increased (irrespective of correct pitch discrimination). For each staircase, reversal steps were 8, 8, 4, 4, 2, and 2 dB for the first six reversals and 1 dB for subsequent reversals. Every separate staircase stopped after 12 reversals. Within a set of six trials (two pitches and three staircases each), trial order was random. If a staircase was completed before the others, only the remaining staircases were sampled.

After the calibration, a validation block was run with 100 trials (90 critical and 10 catch trials). The level of the critical low-pitch and high-pitch tones in the validation block was determined from both the convergence of the three staircases per pitch (from visual inspection) and psychometric response functions. If 45 to 55% of the critical trials were rated as aware in the validation block, the experiment began. If a subject had less than 45% of the critical trials rated as aware, another validation block was run at a higher sound level as suggested by the psychometric response function. Similarly, if a subject reported awareness on more than 55% of the critical trials, another validation block was run at a lower sound level. Validation blocks were repeated until the 45–55% aware criterion was met. All subjects met the criterion within five validation blocks. For the final sample (N = 24), mean sound pressure levels were 5 dB (SD = 4) for low pitch and 7 dB (SD = 5) for high pitch.

2.5. ERP preprocessing

The data from the 64 standard positions together with the tip of the nose and the right cheek were processed and analyzed offline with the MNE-Python software (Gramfort et al., 2013, 2014) and R (R Core Team, 2016). The behavioral analyses included all trials, whereas in the EEG data analyses, some trials were excluded (see below). As described in the preregistration, the continuous EEG data were processed offline. After high-pass filtering of the continuous EEG data with a 0.1-Hz Butterworth 4th degree two-pass filter, all electrodes were re-referenced to the tip of the nose, and Fpz was also re-referenced to the cheek electrode (for a combined measure of vertical and horizontal electrooculography). Although it was not mentioned in the preregistration, we downsampled the data to 256 Hz to speed up preprocessing. In the preregistration, we proposed to consider only a small set of electrodes (Fpz, Fz, Cz, Pz, tip of nose, and cheek) during preprocessing. However, we decided that it would be preferable to consider all 66 electrodes (i.e., the 64 standard positions, tip of nose, and cheek) to minimize unnecessary exclusion of subjects and trials. First, noisy electrodes from the small electrode set could be interpolated from neighboring electrodes to retain individual subjects rather than exclude them. Second, because eye blinks can be detected easily with 66 electrodes but not with the small electrode set, we used independent component analysis (ICA) on all electrodes to correct rather than remove trials with eye blinks. Critically, because these corrections were blind to the type of trials, they were unlikely to bias the results. Instead, they improved data quality because they enhanced the chances of retaining subjects and trials. Accordingly, all electrodes were considered in the following preprocessing steps. Electrodes were visually inspected to detect noisy electrodes. Only a few electrodes had to be interpolated (spherical spline interpolation) from neighboring electrodes (M = 0.57, SD = 1.05). For the ICA, pauses in the continuous EEG recordings were removed, noisy channels were interpolated, and a 1-Hz high-pass filter was applied (to remove slow drift). ICA (fastica) was conducted and eyeblink components were selected by manual inspection of their topography. One component was removed per subject.

As preregistered, epochs were extracted from 100 ms before tone onset to 600 ms after tone onset. Each epoch was baseline corrected to the mean of the 100-ms interval before tone onset (~100 to 0 ms). For each subject, maximum amplitude ranges were extracted for individual epochs, and the distribution of these amplitude ranges was inspected. Individual trials that were apparent outliers were excluded. The number of trials removed per subject were M = 23.43 (SD = 26.40), corresponding to M = 2.93% (SD = 3.30). The exclusion thresholds were set for each individual because subjects showed substantial variability in these amplitude ranges. Critically, inspection of trials was blinded to trial type (low pitch, high pitch, and catch), trial performance, and awareness ratings to avoid bias (Keil et al., 2014).

2.6. Electrophysiology: Aware correct versus unaware correct

As preregistered, two event-related potentials (ERPs) were derived from critical trials for each pitch (low and high): Aware correct trials were correctly categorized tones rated as “I heard the tone clearly” or “I heard the tone weakly.” Unaware correct trials were
correctly categorized tones rated as “I did not hear a tone.” Note that for aware trials, clearly and weakly heard tones were combined because subjects rarely rated their awareness as clear (< 2%, see Section 3.1), as in our previous study (Eklund & Wiens, 2019).

For each tone pitch (low and high), a difference wave was calculated by subtracting the unaware correct ERP from the aware correct ERP. The two difference waves for low pitch and high pitch were averaged to a final difference wave of aware correct minus unaware correct. On the basis of our previous study (Eklund & Wiens, 2019), we predicted that this difference wave would be negative between 160 and 260 ms after tone onset (AAN) and positive between 350 and 550 ms after tone onset (LP). Mean AAN amplitudes were computed across Fz and Cz electrodes, and mean LP amplitudes were computed for the Pz electrode.

We conducted Bayesian hypothesis testing with Bayesian one-sample t tests to determine the degree of evidence for or against the alternative hypothesis (Dienes, 2008). The Bayes Factor (BF10) expresses the likelihood of the data given the alternative hypothesis relative to the likelihood of the data given the null hypothesis, whereas the BF01 shows the reverse (Dienes, 2008, 2016; Wagenmakers, Marsman, et al., 2017; Wiens & Nilsson, 2017). Although the BF is a continuous measure of evidence, we used an interpretation scheme to facilitate verbal communication (Wagenmakers, Love, et al., 2017). The BF was calculated with Aladins Bayes Factor in R (Wiens, 2017).

The alternative hypotheses for AAN and LP were modeled as T distributions defined by the observed effects across two previous samples (Eklund & Wiens, 2019). Specifically, we preregistered the following priors for AAN (M = −0.84, SD = 0.27, df = 45, 2-tailed) and LP (M = 2.54, SD = 0.32, df = 45, 2-tailed). In the analyses, the likelihood was modeled as a t distribution. We used BF greater than 3 or less than 1/3 as the cut-off. We also computed the 95% credible intervals (with an uninformed prior) for the mean amplitudes for AAN and LP.

Although we preregistered that we would control for performance, we did not explicitly state the procedure. Because the Lamy et al. (2009) correction is incorrect, we used the Morales et al. (2015) correction. In their SDT model, Morales et al. assume that the internal response represents unconscious processing and that the relatively smaller unconscious processing to unaware correct trials needs to be multiplied by a scaling factor to match the relatively greater strength of unconscious processing to aware correct trials. For example, if the SDT model suggests a scaling factor of 2.5, the ERP amplitude to unaware correct trials needs to be multiplied by 2.5 to match the strength of unconscious processing in these trials to that in aware correct trials.

Because Morales et al. (2015) described their correction for a discrimination and detection task with two stimuli (A and B), the correction could be applied to the present design as follows: The correction considered performance in terms of correct discrimination between a low frequency tone and a high frequency tone, and in terms of awareness of either low or high frequency tone (i.e., discrimination awareness). Trials in which no tone was presented (i.e., catch trials) were irrelevant because these trials are useful for tone detection but not tone discrimination. In contrast, trials were relevant if subjects reported that they were aware of either low or high frequency tone but chose the wrong frequency (i.e., aware incorrect trials). These trials are predicted by SDT (see Introduction) and were incorporated when computing the scaling factors.

For each subject, the scaling factors for each frequency were computed in two steps (Morales et al., 2015). In the first step, discrimination ability (d') and response criterion (c) between frequencies were computed from the proportion of high-frequency tones labeled correctly as high frequency (hits) and the proportion of low-frequency tone labeled incorrectly as high frequency (false alarms). Thus, in the SDT model, high frequency was considered arbitrarily as signal whereas low frequency was considered as noise. This first step ignored subjects’ ratings of awareness and used the obtained SDT indexes of d' and c to generate the internal response curves for each frequency. In the second step, the proportion of aware correct was combined with the proportion of aware incorrect to compute an awareness threshold, separately for each frequency. The resulting z scores were used to define the weights of the internal responses (i.e., areas under the curve) for aware correct and for unaware correct. The ratio of the weights between aware correct and unaware correct defined the scaling factor for each frequency (e.g., 2.5). Last, the ERP to unaware correct trials was multiplied by this scaling factor to match the (weak) unconscious processing to unaware correct trials to the (strong) unconscious processing to aware correct trials.

Importantly, the correction by Morales et al. (2015) will work only if the ERP to unaware correct trials is affected only by unconscious processing. Unfortunately, our results for the preregistered analyses (see Section 3) suggested that the ERP to unaware correct trials was confounded by a low-frequency drift. This drift was visible even in catch trials in which no tone was presented. Specifically, the mean amplitude to unaware correct trials (as well as catch trials) was about 0.5 µV in the AAN-relevant interval. In contrast, the mean amplitude to aware correct trials was −1 µV and thus had opposite polarity. If the mean amplitude to unaware correct trials would be multiplied by a scaling factor (e.g., 2.5), the correction would estimate that the ERP from unconscious processing was 0.5 × 2.5 = 1.25 µV. After correction, the difference between aware correct and unaware correct would be −2.25 µV. Thus, the correction would suggest that the ERP difference is much larger after correction than before correction. Because this result is unreasonable, it illustrates that the contrastive analysis is distorted if the ERP to unaware correct trials is affected by variables other than unconscious processing. To minimize the confounding effect of low-frequency drift, we deviated from the preregistered analysis and used a 1-Hz high-pass filter (rather than a 0.1-Hz high-pass filter) before performance correction.

2.7. Electrophysiology: Aware versus unaware (ignoring performance)

In an exploratory analysis, we ignored performance and computed the contrastive analysis of awareness minus unawareness. Trials were sorted into two conditions: aware (tones rated as “I heard the tone clearly” or “I heard the tone weakly”) and unaware (tones rated as “I did not hear any tone”). All subjects (N = 28) were included because their performance was irrelevant. Because the ERPs for the preregistered analyses showed a low-frequency drift, we applied a high-pass filter with a 1-Hz Butterworth 4th degree two-pass filter. The selection of intervals and electrodes for AAN and LP was data-driven: We viewed the ERPs and topographies and
selected intervals and electrodes that were maximally sensitive to AAN and LP. The interval was between 190 and 290 ms for AAN and between 370 and 520 ms for LP. The mean AAN and LP amplitudes were computed across a set of 15 central-parietal electrodes (C3, C1, Cz, C2, C4, CP3, CP1, CPz, CP2, CP4, P3, P1, Pz, P2, and P4). The BayesFactor package in R (Morey & Rouder, 2018) was used to compute one-sample Bayesian t tests with the alternative hypothesis (or prior) modeled as a Cauchy distribution ($r = 0.707$). This prior is recommended as a default prior for standardized effects and is used in the software JASP (Wagenmakers, Love, et al., 2017; Wagenmakers, Marsman, et al., 2017).

To explore the neural generators of the AAN and LP scalp topography, we performed source analysis with dynamic statistical parametric mapping (Dale et al., 2000) as implemented in the MNE software (Andersen, 2018; Gramfort et al., 2013, 2014). ERP data from the 64 standard electrodes and the tip of the nose were referenced to the average of all electrodes. Because individual magnetic resonance images were not available, a template brain from the MNE software was used to model both the cortex (sources 3.1 mm apart) and the volume conductor (a boundary element method that models brain, skull, and skin separately with unique conductivities). To capture AAN and LP, source localization was performed on the mean ERP difference between aware and unaware critical trials across low- and high-pitch trials at 240 ms after tone onset for AAN and at 400 ms after tone onset for LP. Because source localization was explorative, no significance testing was performed.

3. Results

3.1. Behavior

Table 1 shows the descriptive statistics for the behavioral data ($N = 24$). Tones were presented close to the individual awareness threshold (50%) because close to fifty percent of the critical low-pitch and high-pitch tones were rated as aware. Of the tones rated as aware, most were rated only as weakly heard ($M = 98.8\%$, $SD = 3.4$). Further, results showed that when no tone was presented (i.e., catch trials), subjects rated that they were aware of a tone on about 20% of trials ($M = 18.1\%$, $SD = 13.8$). Notably, these false alarms are irrelevant for the correction procedure by Morales et al. (2015) because it concerns the discrimination between tones (high vs. low frequency) rather than the detection of tones (tone present vs. absent). Nonetheless, the false alarms suggest that subjects reported even faint experiences as aware, consistent with our goal to capture awareness mainly in terms of weakly heard stimuli.

To illustrate the correction procedure by Morales et al. (2015), Fig. 3 shows the signal detection model of a hypothetical subject whose performance matched the mean behavioral performance (see Table 1). The model shows that for each frequency, the weight of the internal response is greater for aware correct than unaware correct. For example, for high frequency, the weight of the aware correct trials is the area under the solid magenta line to the right of the vertical dashed magenta line. Similarly, for high frequency,

![Fig. 3. Signal detection model of a hypothetical subject whose performance matched mean behavioral performance (see Table 1). The model by Morales et al. (2015) was applied to compute discrimination ability ($d'$) and discrimination criterion ($c$) in regards to the discrimination between low and high pitch and to compute the awareness thresholds for each pitch. For this hypothetical subject, $d' = 1.11$, $c = 0.17$, low pitch awareness threshold = $-0.90$, and high pitch awareness threshold = 0.99.](image-url)
the weight of the unaware correct trials is the area under the solid magenta line between the vertical dashed black line and the vertical dashed magenta line. According to Morales et al. (2015), the internal responses by themselves are unconscious processes that have to be matched between aware correct trials and unaware correct trials. To that end, the correction computes scaling factors that are used to upscale the (weak) unconscious processing in unaware correct trials to the (strong) unconscious processing in aware correct trials. For this hypothetical subject, the scaling factor is 4.18 for low frequency and 2.83 for high frequency.

Results further showed that unaware performance was better than chance (i.e., > 50%) for low-pitch tones (M = 65.8%, SD = 12.4, 95% CI [60.6, 71.0]) but not for high-pitch tones (M = 47.1%, SD = 17.3, 95% CI [39.8, 54.4]), and unaware performance was better to low-pitch tones than to high-pitch tones (M = 18.7%, SD = 25.5, 95% CI [7.9, 29.5]). However, this finding does not imply that there was unconscious processing to low-pitch tones. Instead, an overall response bias in discrimination can explain this pattern. As shown in Fig. 3, a typical subject tended to respond with low pitch rather than high pitch (i.e., the criterion is not at zero but positive). As a consequence, for all of the unaware trials (i.e., all trials between the awareness thresholds), relatively more were correct for low pitch than for high pitch. In fact, there is a clear negative relationship between unaware performance for low-pitch and high-pitch tones. For example, if a hypothetical subject responds only with low pitch when unaware, unaware performance is 100% for low pitch but 0% for high pitch. In support, there was a strong negative correlation of unaware performance between low pitch and high pitch across subjects in the present study (r = −0.72, 95% CI [−0.86, −0.47]).

3.2. Electrophysiology

Fig. 4 shows mean ERPs for aware correct, unaware correct, the difference between aware correct and unaware correct trials, unaware incorrect, and unaware catch trials in the preregistered analyses. The green line in panel A shows a negativity with a peak at 200 ms after stimulus onset (AAN), and the green line in panel B shows a positivity with a peak at 400 ms after stimulus onset (LP). AAN and LP were present for both low- and high-pitch tones, as shown in supplementary material (Wiens and Eklund, 2020). Table 2 shows the descriptive and inferential statistics for the preregistered analyses. For the preregistered prior (which was derived from previous, published results), Bayesian one-sample t tests confirmed the presence of AAN (BF10 > 80) and LP (BF10 > 5000), providing very strong evidence for AAN and extreme evidence for LP.

As shown in Fig. 4, the ERPs showed a general slow, positive drift for all conditions (even for unaware catch trials when no tone was presented). Because a slow drift distorts the correction procedure (see Method section), we preprocessed the raw data again but with a 1-Hz high-pass filter (rather than with the preregistered 0.1-Hz filter). We then tried to apply the correction by Morales et al. (2015) to these data. A thorough description of all steps is provided in the supplementary material. In short, the behavioral data for
each subject were used to estimate SDT indices of pitch discrimination performance, that is, discrimination sensitivity ($d'$) and discrimination criterion ($c$), and the awareness thresholds. On the basis of the SDT model for each subject, scaling factors were computed separately for low and high pitch. Results showed that several subjects had very large and even negative factors (e.g., subjects in the present study ($N = 24$) varied in their absolute discrimination criterion ($M = 0.24$, $SD = 0.21$, 95% CI [0.15, 0.33]).

Critically, scaling factors can be greatly distorted even by small response biases, as explained in the supplementary material. To estimate the effects of this distortion, we simulated hypothetical ERP data as in Morales et al. (2015) on the basis of each subject’s $d'$, $c$, and awareness thresholds. Afterward, we corrected these simulated ERPs for each subject. Results were that these corrected ERPs were distorted for the subjects with large scaling factors. However, when these subjects were excluded ($n = 6$), the correction worked reasonably well on the simulated ERPs for the remaining subjects ($n = 18$). Therefore, we applied the correction to the subjects’ actual ERP data in the present study. Their scaling factors ranged between 2.5 and 5.5 for low pitch ($M = 3.7$) and between 2.4 and 4.4 for high pitch ($M = 3.1$).

Fig. 5 shows the actual ERPs before correction (top) and after correction (bottom). Because the correction corrects only the unaware correct ERPs (blue line), any changes from before to after correction are driven by changes in the unaware correct ERPs. Before correction, the unaware correct ERPs were close to zero with little variability (as supported by a small confidence interval). After correction, the unaware correct ERPs overlapped zero but showed larger variability (i.e., large confidence interval). As a consequence, the corrected difference waves for AAN and LP were very noisy. This is further illustrated in Fig. 6 that shows the 95% CIs for the various conditions for the AAN-relevant interval (left panel) and for the LP-relevant interval (right panel). Superimposed are the mean amplitudes for individual subjects. As shown, the correction substantially increased variability from unaware correct trials (UC) to the corrected unaware correct trials (UCc) and thus increased variability in the corrected versions of both AAN and LP. Specifically, the standard deviation was 3.4 times larger in the corrected AAN (AANc) than in the uncorrected AAN, and the standard deviation was 2.4 times larger in the corrected LP (LPc) than in the uncorrected LP. Notably, there was no apparent evidence that the correction affected the means of either AAN or LP, as the 95% CIs for the differences between uncorrected and corrected mean amplitudes overlapped zero.

Notably, our results showed that after removing the low-frequency drift, the ERP to unaware correct trials was almost flat (see Fig. 5). Critically, a flat ERP to unaware correct trials is not a problem for the correction procedure (Morales et al., 2015). In theory, the ERP to unaware correct trials is completely driven by unconscious processing. So, a flat ERP to unaware correct trials suggests that this ERP is unaffected by unconscious processing, and that the ERP to aware correct trials is also unaffected by unconscious processing. Because the present study used threshold stimuli, the ERPs may have been sensitive only to conscious processing. With louder tones, it may well be that the ERPs would show evidence of unconscious processing, and that the correction procedure would adjust for this processing. Nonetheless, a flat ERP to unaware correct trials implies that the contrastive analysis between aware correct trials and unaware correct trials is not confounded by differences in unconscious processing, and that there is no apparent reason to consider performance at all.

Further, we considered to use an SDT model that focuses on detection as an alternative to the Morales et al. (2015) model. In the SDT model by Morales et al., the focus is on the discrimination between two stimuli (here, low and high pitch). In our detection model, the focus is on the detection of low-pitch trials versus catch trials (i.e., noise) and of high-pitch trials versus catch trials (Macmillan, 1986). Thus, this alternative SDT model attempts to correct for the internal response in tone detection rather than tone discrimination. However, because the results showed that the scaling factors were negative for most subjects (22 out of 24), they were not meaningful (as explained in the supplementary material).

Because there was no apparent evidence that the Morales et al. (2015) correction affected AAN and LP, we explored results of the typical contrastive analysis of awareness minus unawareness (that ignores performance) with a data-driven selection of intervals and electrodes. Because performance was irrelevant, all subjects were included ($N = 28$). Fig. 7 shows mean ERPs and topographies for the difference between aware and unaware trials in the exploratory analysis (i.e., awareness minus unawareness ignoring performance). AAN was maximal at central-parietal electrodes between 190 and 290 ms after stimulus onset, and LP was maximal at central-parietal electrodes between 370 and 520 ms after stimulus onset. For the exploratory prior (which was represented by a default Cauchy prior), Bayesian one-sample $t$ tests provided extreme evidence for AAN ($BF_{10} > 10$ million) and LP.
Fig. 5. Mean ERPs ($N = 18$) with 95% confidence intervals to tones at the individual awareness threshold, separately for aware correct (red), unaware correct (blue), and aware correct minus unaware correct (green) trials. (A) and (B) show the original ERPs and (C) and (D) show the ERPs after performance correction. In (A) and (C), auditory awareness negativity (AAN) is apparent in the difference wave (green) across Fz and Cz between 160 and 260 ms after tone onset. In (B) and (D), late positivity (LP) is apparent in the difference wave (green) from Pz between 350 and 550 ms after tone onset. The intervals are marked in gray. The data were referenced to the tip of the nose. In the plots, the data were low-pass filtered at 30 Hz. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Fig. 6. Mean amplitudes ($N = 18$) with 95% confidence intervals for the various conditions. The data were high-pass filtered at 1 Hz. Left: Mean amplitudes for the AAN-relevant interval (across Fz and Cz between 160 and 260 ms after tone onset). Right: Mean amplitudes for the LP-relevant interval (for Pz between 350 and 550 ms after tone onset). Superimposed are the mean amplitudes for individual subjects. AC = Aware correct, UC = unaware correct, UCc = corrected unaware correct, AAN = auditory awareness negativity, LP = late positivity.
BF10 > 19,000), as shown in Table 2. Notably, when we considered only the preregistered (rather than data-driven) selection of electrodes and intervals, results were unaffected for AAN (BF10 > 71 million) and LP (BF10 > 15,000). Taken together, these results provide extreme evidence for AAN and LP.

Fig. 8 shows the results of the source localization. For AAN, source localization suggested activity in bilateral auditory cortices (superior temporal cortex) as one of the neural generators. See the supplementary material for videos of the time course (Wiens and Eklund, 2020).

4. Discussion

The main results were that the preregistered contrastive analysis of aware correct trials minus unaware correct trials showed clear evidence for AAN and LP. Because 25% (6 out of 24) of the subjects showed very large scaling factors, the performance correction could not be applied to these subjects’ ERP data. For the remaining subjects (N = 18), there was no apparent evidence that the performance correction affected AAN and LP. Further, the typical contrastive analysis of aware minus unaware trials (ignoring performance) provided extreme evidence for AAN and LP, and sources for AAN included bilateral auditory cortices.

Results from the contrastive analysis of aware correct trials minus unaware correct trials provided clear evidence for AAN and LP (see Table 2). Although we wanted to apply the correction by Morales et al. (2015), the model assumes that there is no response bias.
towards either low or high pitch. However, simulations showed that the correction can be distorted by even small response biases (as described in the supplementary material). In support, the scaling factors for some subjects (N = 6) were unreasonably large (and even negative). When we simulated hypothetical ERPs for these subjects, their corrected ERPs were distorted. We excluded these subjects and retained only subjects (N = 18) with reasonable scaling factors (between 2.4 and 5.5). We simulated hypothetical ERPs for these subjects and found that the correction worked reasonably well on their simulated ERPs.

However, when we applied the correction by Morales et al. (2015) to the subjects’ actual ERP data in the present study, results provided no apparent evidence that the correction affected either AAN or LP, as the 95% CIs of the difference between corrected and uncorrected AAN (and LP) overlapped zero (Figs. 5 and 6). Notably, these results do not provide convincing evidence for or against an effect of the correction on AAN and LP. Thus, the present findings do not resolve whether AAN and LP are affected by performance. Nonetheless, a reasonable interpretation is that the correction mainly increases noise in the measurement. If the correction worked properly, it should adjust the mean of unaware correct (UC) closer to the mean of aware correct (AC) without affecting its variability. However, Fig. 6 suggests that the correction mainly increased the variability from before (Uc) to after the correction (UCc). As a consequence, the standard deviation was 3.4 times larger for the corrected than uncorrected AAN, and the standard deviation was 2.4 times larger for the corrected than uncorrected LP. Therefore, the correction seems to mainly increase variability and thus the noise in the measurement of AAN and LP. Furthermore, the correction clearly has a practical problem with handling response biases, as an unreasonable large number of subjects (6 out of 24 = 25%) had to be excluded. Importantly, we tried to minimize response biases by instructing subjects to always guess to the best of their ability about tone pitch and to avoid using any strategy when they did not hear the tone (such as always responding low tone when they did not hear a tone). Taken together, the present results suggest that the correction may have practical problems and may also not have any effect on AAN and LP.

For these reasons, we explored results for the AAN and LP in the typical contrastive analysis of awareness minus unawareness (that ignores performance). As shown in Table 2 and Fig. 7, there was clear evidence for AAN as well as LP. As shown in Fig. 8, source localization for AAN suggested sources in bilateral auditory cortices (superior temporal cortex), similar to the results of the source localization in our previous study (Eklund et al., 2019). These results support the conclusions that phenomenal consciousness, that is, what it is like to have an experience, may be indexed by AAN, that AAN is mediated by local recurrent processing in early sensory areas (Lamme, 2010), and that the early NCC are consistent across sensory modalities (Dykstra et al., 2017; Meyer, 2011; Snyder et al., 2015).

Aside from practical problems, the premise of the correction by Morales et al. (2015) suffers from at least two theoretical problems. First, the SDT model by Morales et al. focuses on awareness of the discrimination between two stimuli. In the present study, this discrimination awareness refers to the two tone frequencies: Subjects are aware of how the low-frequency tone (900 Hz) differs from the high-frequency tone (1400 Hz), or vice versa. This does not seem intuitive because subjects may have an experience of detecting a tone without an experience of how this tone differs from other tones. Because the present study used simple tones that clearly differed in frequency (900 Hz and 1400 Hz), it was probably true that when subjects experienced a tone, they were also clearly aware of its frequency. However, this may not be the case with tones that are closer in frequency (e.g., 900 Hz and 910 Hz). These tones could be rather loud so that subjects experience the tones (i.e., detection) without any experience of a difference between the tones (i.e., no discrimination). The distinction between detection and discrimination awareness is particularly relevant for AAN because it is unresolved if it is a neural correlate of detection or discrimination, similarly as for VAN (Koivisto, Grassini, Salminen-Vaparanta, & Revonsuo, 2017).

Because the SDT model by Morales et al. (2015) focuses on discrimination awareness rather than detection awareness, we tried to adapt this model to detection awareness (see supplementary material). To represent detection awareness rather than discrimination awareness, awareness was conceptualized as the difference between experiencing a tone (either low or high frequency) and experiencing no tone (i.e., catch trials). The scaling factor was the ratio of the weight of aware trials and the weight of unaware trials on tone trials. Whereas the weight of aware trials was always positive, the weight of unaware trials was negative for most subjects. As a result, the scaling factors for each frequency were negative for most subjects (22 of 24). Because negative scaling factors flip the polarity of the ERP to unaware trials (but not to aware trials), they are not meaningful. To conclude, the SDT model is limited because it deals only with discrimination awareness and cannot handle detection awareness.

Second, according to Morales et al. (2015), awareness is something beyond the processing captured by the internal response: “When controlling for performance capacity in imaging studies, researchers should focus on controlling for the internal response strength …. In imaging studies of consciousness, this means isolating some kind of further processing which only happens during trials crossing the awareness criteria.” (p. 9). However, these awareness criteria may simply be the thresholds of reported awareness; that is, the levels of the internal response above which the percept is reported as aware (of either low pitch or high pitch). Because Lamme (2010) argues that phenomenal consciousness can occur without a report of awareness and is indexed by local recurrent processing, we propose that the strength of the internal response may reflect the degree of local recurrent processing in sensory areas and thus the level of awareness. In support, awareness varies gradually (Sandberg et al., 2010), and early activity in the visual cortex is associated with the reported strength of awareness, as measured with the perceptual awareness scale (Andersen, Pedersen, Sandberg, & Overgaard, 2016). Further, V1 activity already matches awareness closely (Lamme, Supèr, Landman, Roelfsema, & Spekreijse, 2000; Michel, Chen, Geisler, & Seidemann, 2013; Schwarzkopf, Song, & Rees, 2011; Silvanto, Lavie, & Walsh, 2005). Because in the correction by Morales et al., the level of the internal response below the awareness threshold is scaled to the level of the internal response above the awareness threshold, the correction may not isolate awareness but rather completely remove processing related to awareness. As a consequence, it may isolate post-perceptual processing that is related to a report of awareness. Post-perceptual processing involves activity that is different from awareness-related activity but perfectly correlated with awareness such as planning a report of the experience (Pitts, Metzler, & Hillyard, 2014; Tsuchiya, Wilke, Frässle, & Lamme, 2015).
However, if this were true, the correction by Morales et al. (2015) should have reduced AAN (and maybe LP). This reasoning rests on two assumptions. First, the internal response reflects local recurrent processing, and second, AAN is an index of local recurrent processing. These assumptions link the internal response with AAN: Internal response = local recurrent processing = AAN. So, if the correction eliminates differences in the internal responses to aware correct and unaware correct trials, then AAN should be reduced, if not eliminated. However, the present results do not support this reasoning because AAN (and LP) was apparently unaffected (Fig. 6). One potential explanation could be that even though both assumptions are true, they do not hold for an SDT model that focuses on discrimination awareness (low vs. high tone) rather than detection awareness (tone present vs. absent). This would be consistent with findings in vision that VAN was sensitive to detection rather than discrimination (Koivisto et al., 2017). Accordingly, if AAN represents detection awareness, correcting for discrimination awareness may not have any effect. Because we were unable to modify the SDT model to work for detection awareness, we could not test this explanation.

Another explanation is that the second assumption is incorrect: AAN is not an index of local recurrent processing. Although there is no direct proof that AAN reflects local recurrent processing, the timing of the AAN (about 240 ms after tone onset) and its apparent sources in bilateral auditory cortices are consistent with the idea that AAN reflects local recurrent processing. Critically, even if AAN does not reflect local recurrent processing, results suggest that it is a valid neural correlate of consciousness because it was unaffected by the correction by Morales et al. (2015).

Yet another explanation is that the first assumption is incorrect: The internal response does not reflect local recurrent processing but unconscious processing. For example, the internal response may reflect the feedforward sweep (Lamme & Roelfsema, 2000). Because we used threshold stimuli, the ERPs may not have been sensitive enough to detect unconscious processing. In support, the ERP to unaware correct trials was flat (Fig. 5). So, if the ERP did not detect unconscious processing, the ERP to aware correct trials reflects only conscious processes, and the correction by Morales et al. (2015) would not change the ERPs at all. However, with stimuli above threshold, the ERPs will likely show clear evidence of unconscious processing, and the correction procedure might adjust for this processing. Nonetheless, the flat ERP to unaware correct trials in the present study implies that the contrastive analysis between aware correct trials and unaware correct trials is not confounded by differences in unconscious processing, and that there is no apparent reason to consider performance at all.

To conclude, the results of the typical contrastive analysis of awareness minus unawareness (ignoring performance) provided clear evidence for auditory awareness negativity (AAN) and late positivity (LP) as neural correlates of hearing. Results of the source localization are consistent with recurrent processing theory because AAN may be an indirect measure of local recurrent processing in bilateral auditory cortices. As such, AAN indexes phenomenal consciousness. When we applied an SDT approach to correct for unconscious processing, we encountered a practical problem in handling response biases and a large number of subjects had to be excluded. For the remaining subjects, results suggested that the correction mainly introduced noise in the measurement of AAN and LP. Critically, we argue that the usefulness of the correction is limited because it cannot handle detection awareness and because it may eliminate awareness-related activity to isolate post-perceptual processing. Thus, we conclude that AAN and LP are not confounded by performance and that the contrastive analysis identifies both as correlates of awareness.

Acknowledgement

This work was supported by the Swedish Research Council (Grant 2015-01181) and Marianne and Marcus Wallenberg (Grant 2019-0102). We thank Mats E. Nilsson for helpful discussions.

Author contributions statement

All authors designed the study and wrote the manuscript together. RE programmed the experiment and RE and SW wrote the scripts to analyze the behavioral and EEG data. RE and BG collected the data.

Declaration of Competing Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.17045/sthlmuni.9724280.

References

Andersen, L. M. (2018). Group analysis in MNE-python of evoked responses from a tactile stimulation paradigm: a pipeline for reproducibility at every step of processing, going from individual sensor space representations to an across-group source space representation. Frontiers in Neuroscience, 12, 1–18. https://doi.org/10.3389/fnins.2018.00906.

Andersen, L. M., Pedersen, M. N., Sandberg, K., & Overgaard, M. (2016). Occipital MEG activity in the early time range (< 300 ms) predicts graded changes in perceptual consciousness. Cerebral Cortex, 26(6), 2677–2688. https://doi.org/10.1093/cercor/bhv108.

Aru, J., Bachmann, T., Singer, W., & Melloni, L. (2012). Distilling the neural correlates of consciousness. Neuroscience & Biobehavioral Reviews, 36(2), 737–746. https://doi.org/10.1016/j.neubiorev.2011.12.003.
