General Markers of Conscious Visual Perception and Their Timing

Renate Rutiku[1,2]*, Jaan Aru[2,3] and Talis Bachmann[2]

[1] *Institute of Psychology, Faculty of Social Sciences and Education, University of Tartu, Tartu, Estonia,* [2] *Cognitive Psychology Lab, Institute of Public Law, Faculty of Law, University of Tartu, Tallinn, Estonia,* [3] *Computational Neuroscience Lab, Institute of Computer Science, Faculty of Mathematics and Computer Science, University of Tartu, Tartu, Estonia*

Previous studies have observed different onset times for the neural markers of conscious perception. This variability could be attributed to procedural differences between studies. Here we show that the onset times for the markers of conscious visual perception can strongly vary even within a single study. A heterogeneous stimulus set was presented at threshold contrast. Trials with and without conscious perception were contrasted on 100 balanced subsets of the data. Importantly, the 100 subsets with heterogeneous stimuli did not differ in stimulus content, but only with regard to specific trials used. This approach enabled us to study general markers of conscious visual perception independent of stimulus content, characterize their onset and its variability within one study. N200 and P300 were the two reliable markers of conscious visual perception common to all perceived stimuli and absent for all non-perceived stimuli. The estimated mean onset latency for both markers was shortly after 200 ms. However, the onset latency of these markers was associated with considerable variability depending on which subsets of the data were considered. We show that it is first and foremost the amplitude fluctuation in the condition without conscious perception that explains the observed variability in onset latencies of the markers of conscious visual perception.

Keywords: NCC, consciousness, variability, P300, VAN, ERP, denoising

INTRODUCTION

How long does it take from the moment when a stimulus is presented in the environment until the conscious experience of the stimulus starts to arise? Despite the decades-long quest for the neural correlates of consciousness (NCC) it is not known at what time after stimulus onset they occur. Some results suggest that conscious perception is a relatively late process (Sergent et al., 2005; Del Cul et al., 2007). Others point to the importance of mid-latency markers (Koivisto and Revonsuo, 2010). Still others have found very early correlates for conscious perception (Pins and Ffytche, 2003; Aru and Bachmann, 2009).

One reason for these discrepancies may be that the contrastive method typically used to identify NCC is not only sensitive regarding the actual NCC but also unravels neural processes that precede or follow conscious perception (Overgaard, 2004; Bachmann, 2009; Aru et al., 2012; de Graaf et al., 2012). The contrastive method is supposed to identify markers that are uniquely present or reliably more strongly present in the averaged activity of the condition where a stimulus was consciously perceived compared to the condition where a stimulus was not consciously perceived. However, the markers directly related to conscious perception may not be the only ones that differ between these conditions. Depending on how visual awareness is manipulated and assessed within

a given paradigm, neural prerequisites (NCC-pr) and neural consequences (NCC-co) specific to that paradigm may be misclassified as NCC proper (Aru et al., 2012; de Graaf et al., 2012).

Procedural differences between studies can influence the presence as well as the characteristics of the three types of NCC. If experiments employ restricted categories of stimuli it is hard to tell whether the resulting NCC are markers of only one category or whether they can be generalized to other categories as well. For example, the N170 may be a marker of category specific NCC-pr or even NCC proper only for faces (Navajas et al., 2013; Shafto and Pitts, 2015). It is also known that the latency of processes correlating with consciousness may shift as much as 100 ms depending on stimulus predictability (Melloni et al., 2011). If the stimulus set of a study consists of only a few items then perceptual events inevitably become more predictable and the latencies might shift accordingly (Melloni et al., 2011). Finally, specifics of the experimental setting or task requirements may even influence ongoing pre-stimulus activity which can nonetheless have a robust effect on subsequent stimulus perception and thus possibly also on the related markers of NCC (see Busch et al., 2009 and Mathewson et al., 2009 for two excellent studies on pre-stimulus predictors of conscious perception). Taken together, the fact that a wide variety of different paradigms, stimulus material, recording conditions etc. have been used to study NCC might at least in part explain why many studies have reported largely different signatures and onset times of NCC (see Koivisto and Revonsuo, 2010; Dehaene and Changeux, 2011 for an overview).

However, there might be yet another reason for the discrepancies: maybe the markers of NCC arise at different times even within a study. It has never been tested whether the markers of NCC are reliably identifiable with a specific latency within one study where the paradigm, stimulus material and recording conditions are kept constant. If the same subjects perform the same task over and over again, would contrasting the resulting seen and unseen trials (or representative samples thereof) always lead to comparably similar results in terms of when and where the NCC arise? Looking closer at the rationale behind the contrastive method suggests that this may not be the case. The reliability and onset latency of the markers of the NCC might be affected by a number of different factors even within one study.

It is possible that the latency of the NCC shifts from trial to trial. This would spread out the averaged activity in the condition with conscious perception and the mean onset latency of NCC would become less accurate. A similar effect has been demonstrated for the face-sensitive N170 component if stimulus uncertainty is increased due to added noise (Navajas et al., 2013). In the worst case scenario latency jitter may even hide the NCC from the contrastive analysis altogether. Results from a contrastive analysis may also be influenced by factors not directly related to the NCC. Different noise profiles may accompany the signal in different trials (Arieli et al., 1996). Again, this would influence the onset latency of NCC. One assumes that task-irrelevant noise is mostly averaged out when means are created over trials, but this is of course not completely true and is

particularly problematic if the number of trials differs between the contrasted conditions.

To make matters worse, one cannot even be sure that it is only the signal and noise profiles of the condition with conscious perception that dictate NCC reliability and onset latency. The above described concerns apply to the condition without conscious perception as well. This is because for delineating NCC, trials with conscious perception are compared against those without conscious perception of the target stimulus. Only the significant differences are considered as candidates of NCC (Aru et al., 2012), but the reliability and timing of these significant differences also depends on the trials in the condition without conscious perception.

The last consideration is particularly noteworthy in light of the recent work by Schurger et al. (2015). Their results suggest that the pattern of activity in response to unseen stimuli is less stable within and between trials than the pattern of activity in response to seen trials. They used a measure of representational similarity called directional variance. This measure describes how stable the topographic pattern is within a given time window. Note that although directional variance is more sophisticated than the simple ERP calculation the logic behind it is quite similar. It is the core assumption behind ERPs as well that if activity consistently occurs at the same time over trials then it is preserved after averaging whereas inconsistent activity is averaged out. Thus, if directional variance is higher in trials of the unseen condition (Schurger et al., 2015) then it is prudent to assume that ERPs of the unseen condition should also be more variable. Most importantly, this variability will be reflected in the reliability and onset latency of NCC if the contrastive analysis method is used. In other words, trials from the unconscious condition might directly affect the estimated timing of the ERP changes reflecting the NCC.

Taken together, there are several reasons why NCC as identified by a contrastive analysis may vary even within one single study. In order to arrive at a better understanding of the NCC it would be necessary to know how much each of these factors contributes to the results of a contrastive analysis. Surprisingly, however, it has not yet been thoroughly characterized how much NCC actually vary when only the data from one experiment are considered.

The present study was designed to address the above described issues. To overcome some of the methodological restrictions of previous studies we employed an experimental paradigm where the role of visual categorical restriction and stimulus predictability were reduced. To that end we used many different stimuli with varying characteristics and presented these stimuli on perceptual threshold. We hypothesized that for the described paradigm there is at least one marker that distinguishes consciously perceived trials of our heterogeneous visual stimulus set from the non-perceived trials. We call this the general marker of NCC, gmNCC in short. Note, that with "general" we refer to the content-independent nature of the hypothesized gmNCC, because any single stimulus specific NCC-pr, NCC proper, and/or NCC-co would not have a critical impact on results if so many different stimuli are considered together.

Our first goal was to investigate which EEG correlates qualify as gmNCC in our experimental paradigm. Our second goal was to study the reliability and any possible variability in the onset latency of gmNCC. Our third goal was to characterize the causes of this variability as thoroughly as possible. To achieve these goals, 100 matched subsets of seen and unseen trials were created by repeatedly sampling from the pool of all available trials. This procedure (depicted in **Figure 2**) ensured that objective stimulus content always stayed the same for both conditions while the included trials differed from one matching iteration to another. By performing a contrastive analysis on each of the 100 matched subsets of seen and unseen trials separately and by analyzing variability within these results we show that amplitude variance in the unseen condition has a profound influence on gmNCC onset latency and sometimes obscures the gmNCC altogether. Thus, our research may shed light on the question why different studies have found different markers of NCC or report largely different onset times of these markers.

MATERIALS AND METHODS

Subjects

Twenty-two subjects participated in the EEG experiment. All subjects were healthy and had normal or corrected-to-normal vision. Data from four subjects were not included in the analyses due to a high number of noisy electrodes or too many trials with artifacts. The remaining 18 subjects (eight male) were 18–31 years old (mean = 23.2, median = 22, SD = 3.6). One subject was left-handed. All subjects gave written informed consent prior to participation and received monetary compensation as a reward. The study was approved by the ethics committee of University of Tartu and the experiment was undertaken in compliance with national legislation and the Declaration of Helsinki.

Stimuli

The stimulus set consisted of 70 monochrome drawings. The drawings depicted objects from six different categories. Four categories were further divided into line-drawings and solid forms. Thus, there were 10 different types of stimuli: (1) line-drawings of graphical figures, (2) solid graphical figures, (3) short words, (4) line-drawings of man-made objects, (5) solid forms of man-made objects, (6) line-drawings of faces, (7) line-drawings of animated nature, (8) solid forms of animated nature, (9) line-drawings of inanimate nature, (10) solid forms of inanimate nature. **Figure 1** depicts all 70 stimuli sorted by stimulus type. Stimuli were collected from online databases. Occasionally, stimuli were edited manually to keep the number of filled pixels, i.e., the contrast energy comparable for all solid forms including text and all line-drawings including faces. There were no important reasons why particular stimulus types or exemplars were chosen. The aim was simply to generate a heterogeneous stimulus set that is comparable to many other related studies. Solid forms were included in addition to line drawings so that both high- and low-frequency information would be presented to the subjects.
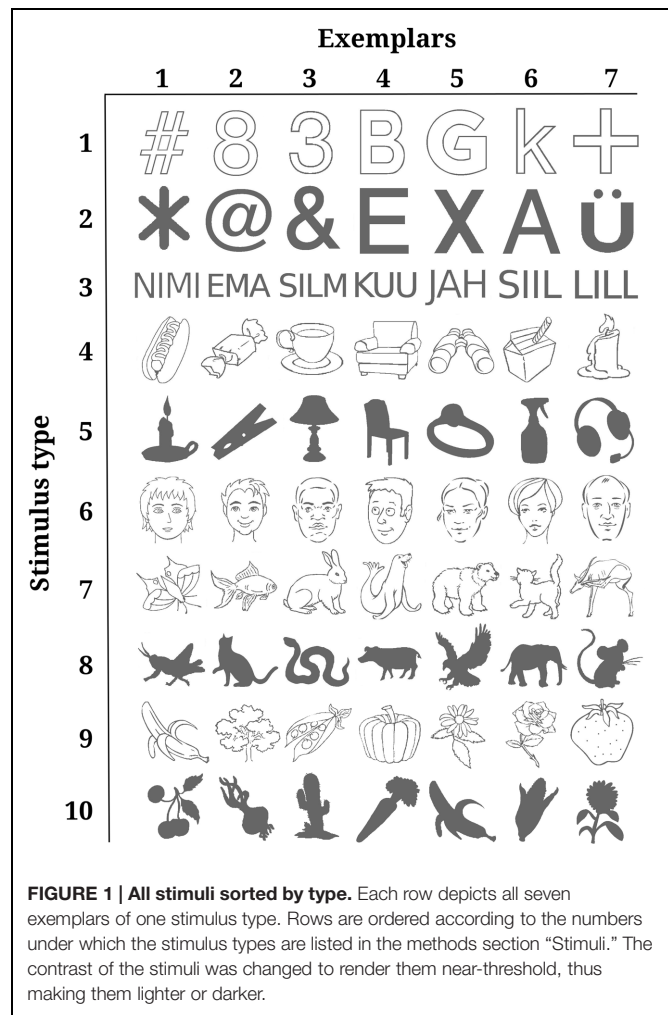

FIGURE 1 | All stimuli sorted by type. Each row depicts all seven exemplars of one stimulus type. Rows are ordered according to the numbers under which the stimulus types are listed in the methods section "Stimuli." The contrast of the stimuli was changed to render them near-threshold, thus making them lighter or darker.

In order to display stimuli at perceptual threshold (i.e., 50% seen responses) their contrast has to be accordingly low. Not all of our stimuli have the same threshold contrast, however. An earlier pilot experiment indicated that for the present stimulus set there are five groups of stimuli with roughly similar threshold contrasts within each group: text, solid graphical figures, line-drawings of graphical figures, solid forms of all other figures and line-drawings of all other figures. Thus, contrast was adjusted separately for each of these five groups with the help of a short pre-experiment prior to the main experiment (see S1 in the Supplementary Material).

Stimuli were presented on a light gray background. Stimulus size was approximately 2.5° of visual angle. Prior to the stimulus a fixation cross was presented. The response screen contained the question "Did you see something?" in the Estonian language. S2 in the Supplementary Material contains more information about the physical characteristics of stimuli.

Task and Design

Subjects were seated in a dark room, 80 cm from the monitor (SUN CM751U; 1024 × 768 pixels; 100 Hz refresh rate). Each session began with a short pre-experiment to determine the

appropriate threshold contrasts for each subject (see S1 in the Supplementary Material), followed by the main experiment. The main experiment comprised 770 trials in total. Each of the 70 stimuli was presented 10 times. There were also 70 catch trials where no stimulus was presented. The order of the trials was fully randomized. Each trial began with the presentation of a fixation cross in the middle of the screen for 500 ms. The fixation cross was followed by a blank screen for 750–1250 ms in order to obtain a clean EEG baseline without the ERP of the fixation cross onset or offset. Then the stimulus was presented in the middle of the screen for one refresh frame, i.e., for 10 ms, followed again by the blank screen. After 1 s the response screen appeared.

Subjects were instructed to fixate on the cross in the middle of the screen, not to blink until the response screen had appeared, and then to report via button press on a standard keyboard whether they perceived a stimulus on a given trial or not. Seen and unseen responses were given with different hands, but the designated hands were balanced across subjects. There was a break after every 154 trials.

EEG Recording and Preprocessing
A Nexstim eXimia EEG-system with 60 carbon electrodes cap (Nexstim Ltd, Helsinki, Finland) was used. All 60 electrodes of the extended 10–20 system were prepared for recording. The reference electrode was placed on the forehead, slightly to the right. The impedance at all electrodes was kept below 15 KΩ. The EEG signal was sampled at 1450 Hz and amplified with a gain of 2000. The bandwidth of the signal was ca. 0.1–350 Hz. As our system only allows one pair of eye-electrodes the horizontal electrooculogram (HEOG) was recorded by placing the respective electrodes a few millimeters from the outer canthi of both eyes. Note that blinks could be easily identified in the EEG of posterior scalp sites because the reference electrode was placed on the forehead.

EEG data was preprocessed with Fieldtrip[1] (version 01-01-2013). Trials were epoched around stimulus onset (−500 to +700 ms), re-referenced to the average reference and baseline corrected with a 100 ms time period before stimulus onset. All trials containing artifacts were identified by visual inspection. Trials containing blinks, eye movements, strong muscle activity or other artifacts were completely removed from the data. Noisy signals were interpolated with the nearest neighbor method (see S3 in the Supplementary Material). The 11.6% of trials were rejected due to artifacts on average (median = 10.7%, SD = 6%, range = 4–26.9%) and 2.6% of the data was interpolated on average (median = 2.6%, SD = 1.8%, range = 0.1–6.3%). Data were filtered with a 30 Hz low-pass zero phase shift Butterworth filter.

Data Analysis
The behavioral analysis was carried out with the R programming language[2] (version 3.1.0). As contrasts had to be readjusted occasionally during the main experiment (see S1 in the Supplementary Material), detection rate also varied in accordance

---

[1]http://fieldtrip.fcdonders.nl
[2]http://www.r-project.org/

with the different levels of contrast. In order to eliminate this accountable variance from the behavioral results only those contrast levels are considered which comprise the most trials. Thus, 93.3% of all available trials are considered (SD over subjects = 9%; SD over types of stimuli = 2.8%). We must note, however, that results are comparable when all available trials are considered.

EEG data was analyzed with Fieldtrip as well as with R.

Trial Matching Procedure
In order to find the gmNCC, to study their reliability and any possible variability in their onset latency within one study 100 different matched sets of seen and unseen trials were constructed per subject. The trial matching procedure serves two goals. First, it guarantees that the two contrasted conditions (seen and unseen) are identical with respect to objective stimulus content. Second, it allows us to repeat the contrastive analysis for 100 objectively equivalent matched sets of trials. Hence one can investigate whether the resulting NCC are also equivalent on every iteration.
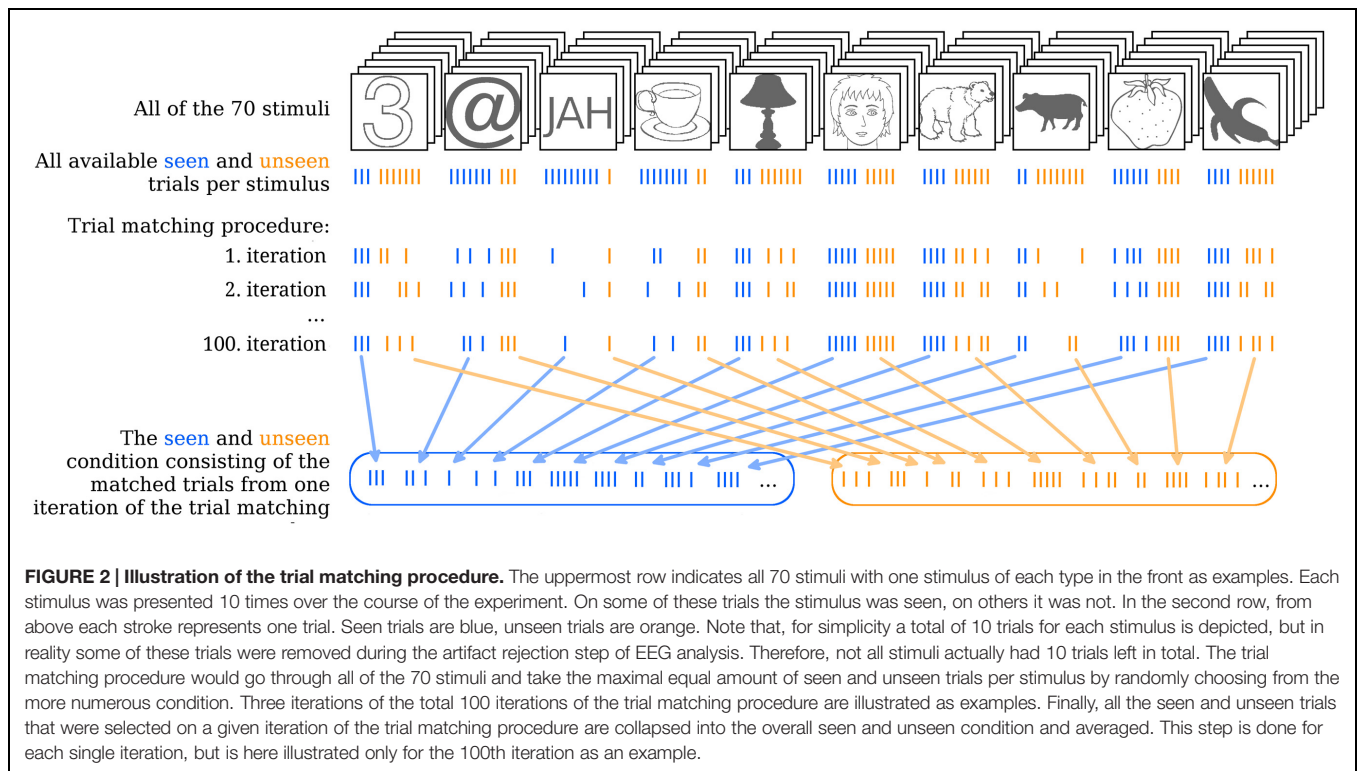
Figure 2 illustrates the trial matching procedure. Each of the 100 sets was composed as follows. For every stimulus an equal number of seen and unseen trials were included in the respective conditions. Thus, the algorithm would select a stimulus (the number "3" in the upper left end of **Figure 2**, for example) and count how many seen and how many unseen trials there are for this stimulus per subject (e.g., 3 vs. 7). It would then take all three seen trials and randomly choose three out of the seven unseen trials. The algorithm would do the same for all 70 stimuli and pool the chosen trials together into their respective conditions for each subject separately. In case the contrast for one particular stimulus had to be readjusted after the first block of the main experiment (see S1 in the Supplementary Material) the algorithm would choose an equal number of seen and unseen trials for each contrast separately.

This random selection of subsets was repeated 100 times for each subject. As a result both the seen and the unseen condition always comprised an equal number of trials for each subject on each iteration of the set matching procedure (m = 122, median = 123, SD = 31, range = 62–177). Furthermore, although there are obvious objective differences between categories of stimuli each matched set of seen and unseen trials consists of the same number of objectively identical stimuli. That holds true for stimulus content as well as for visual parameters (e.g., contrast energy), because the stimulation parameters were fixed throughout the experiment. Importantly, neither stimulus content nor visual parameters could change between matched subsets, i.e., stimulus content and visual parameters were identical not only between seen and unseen sets of trials within one iteration of the trial matching procedure but between all of them – all the 200 sets of trials (100 iterations × 2 conditions) were objectively identical. Hence, physical differences between subsets of stimuli cannot explain any variability in the ERP results from contrastive analyses.

Note that although as a consequence of the trial matching procedure only specific subsets of all available data are considered in the contrastive analyses the amount of trials is still more

FIGURE 2 | Illustration of the trial matching procedure. The uppermost row indicates all 70 stimuli with one stimulus of each type in the front as examples. Each stimulus was presented 10 times over the course of the experiment. On some of these trials the stimulus was seen, on others it was not. In the second row, from above each stroke represents one trial. Seen trials are blue, unseen trials are orange. Note that, for simplicity a total of 10 trials for each stimulus is depicted, but in reality some of these trials were removed during the artifact rejection step of EEG analysis. Therefore, not all stimuli actually had 10 trials left in total. The trial matching procedure would go through all of the 70 stimuli and take the maximal equal amount of seen and unseen trials per stimulus by randomly choosing from the more numerous condition. Three iterations of the total 100 iterations of the trial matching procedure are illustrated as examples. Finally, all the seen and unseen trials that were selected on a given iteration of the trial matching procedure are collapsed into the overall seen and unseen condition and averaged. This step is done for each single iteration, but is here illustrated only for the 100th iteration as an example.

than typically included or even considered necessary for reliable estimates in ERP research (especially for large components such as the P300; see Luck (2005) for a discussion on this topic). After all, even if other studies have used all the recorded trials available to them they are nevertheless also analyzing only a subset of an infinite amount of trials which would maximize the signal-to-noise ratio. Thus the presently employed trial matching procedure should guarantee a sufficiently high signal-to-noise ratio of the experimental conditions and is well comparable to other NCC studies. For example, Sergent et al. (2005) only had a maximum of 96 trials for the seen and unseen condition together, and that was before artifact rejection. Del Cul et al. (2007) had a maximum of 128 trials for the seen and unseen condition together before artifact rejection. Pins and Ffytche (2003) had an estimated average of 100 trials in both seen and unseen conditions. Furthermore, in these studies trial numbers were reported to be only roughly equal between conditions (no further information provided) which may bring its own problems as described in the Section "Introduction."

After the trial matching procedure the seen and unseen conditions comprised 9.6 different types of stimuli on average (median = 10, $SD = 0.6$, range = 8–10), 51.1 different individual stimuli on average (median = 52.5, $SD = 8.2$, range = 31–66) and 1.04 different contrast levels per contrast group on average (median = 1, $SD = 0.1$, range = 1–1.4). Due to the nature of the trial matching procedure some trials (from the less numerous condition) were included in the matched subsets on every iteration (for seen trials: $m = 29\%$, median = 27%, $SD = 14\%$, range = 9–50%; for unseen trials: $m = 23\%$, median = 20%, $SD = 16\%$, range = 5–68%). Other trials were never included

because they could not be matched (for seen trials: $m = 36\%$, median = 39%, $SD = 15\%$, range = 3–66%; for unseen trials: $m = 22\%$, median = 24%, $SD = 15\%$, range = 0–52%). The remaining trials were included in roughly one-third of the 100 matched sets per person (for seen trials: $m = 36\%$, median = 38%, $SD = 7\%$, range = 21–49%; for unseen trials: $m = 38\%$, median = 34%, $SD = 9\%$, range = 25–58%). Thus, the matched sets comprised of 42 and 43% of all available seen and unseen trials on average. The mean within-subject difference between the proportion of seen and unseen trials included in the matched sets from all available seen and unseen trials was only 1%.

For comparisons between the unseen and the catch condition all correctly rejected catch trials and the same matched sets of unseen trials were used. The catch condition comprised 59 trials on average (median = 60, $SD = 5.6$, range = 49–68).

Cluster Permutation Tests

Differences between conditions were analyzed with non-parametric cluster permutation tests as described in Maris and Oostenveld (2007) and implemented in Fieldtrip. The advantage of this method is that it identifies significant differences between conditions as clusters evolving over electrodes and time (see Figure A in S3 in the Supplementary Material for an example). Thus it is well suited to study the onset of significant differences without predefining any electrodes or time periods where the effects might occur (Picton et al., 2000). After averaging the single trials per condition data points (electrode-time pairs) were compared via dependent samples t-tests. Empirical distributions were created using 10 000 random permutations of the data. The maximal sum of t-values belonging to each cluster was

used as the test statistic. Both the entry level for single samples into clusters and the significance threshold for clusters were set at 0.025 for two-sided t-tests. Only clusters lasting longer than 15 ms were considered significant. The choice of a duration threshold originates from the concept of functional microstates (first introduced by Dietrich Lehmann[3]). It is based on the observation that neural activity picked up by EEG exhibits a succession of quasi-stable topographic configurations lasting over tens of milliseconds. Although no precise criterion for the minimal microstate duration exists, 15 ms was taken as a reasonably lenient value. This was necessary because although the data are smoothed by a low-pass filter, it can happen that the cluster permutation test finds small significant blobs consisting of a few electrodes and a few milliseconds in addition to the clear N200 and P300 components (mostly shortly before or after their massive onset/offset). These blobs are impossible to interpret and we have thus discarded them through the adoption of a 15 ms duration threshold.

If not specified otherwise, cluster onsets and offsets were defined as the first/last time points when at least four neighboring electrodes showed significant differences between conditions. See S3 in the Supplementary Material for more information on neighboring electrodes and the cluster formation.

Denoising Single Trials

In order to increase the signal-to-noise ratio for N200 and P300 data were denoised via an algorithm using wavelet decomposition (Quian Quiroga, 2000). This method allows the reconstruction of ERP components on the single trial level. The signal is first decomposed into different wavelets and subsequently reconstructed using only those wavelet coefficients that are relevant for the component of interest. S4 in the Supplementary Material illustrates some denoised trials. Two different sets of wavelet coefficients were used for the reconstruction of the P300 and the N200, but the same sets of coefficients were used for all subjects and all electrodes. All available seen, unseen and catch trials were also always denoised together. For P300, data from electrodes Fcz, C1, Cz, C2, C4, CP1, Cpz, CP2, Pz were denoised. For N200, data from electrodes TP9, TP7, TP10, TP8, P10, P9, O1, Oz, O2, Iz were denoised. These electrodes were selected because results from seen-unseen comparisons with undenoised data indicated that they constitute the most representative electrodes for N200/P300. More specifically, significant differences between conditions occurred first and lasted longest on these electrodes.

It is important to note that this denoising method can also be applied to data with no clear ERP signal (Quian Quiroga and Garcia, 2003). As explained in the introduction and also exemplified in Navajas et al. (2013), there are several reasons why event-related signals may not be apparent from averaged data. This method offers one possibility to find out whether any signal may still be present in the single trials or not.

---

[3]http://www.scholarpedia.org/article/EEG_microstates

Correlation Tests

To explain the variance in gmNCC onset latencies that remained even after denoising, single trial parameters of the two gmNCC (N200 and P300) were extracted from each of the 100 matched sets of trials and correlated with gmNCC onset latencies from the respective contrastive analyses.

First, peak amplitude and peak latency was extracted from the time period of observed variance in the onset latencies of the gmNCC. For each trial, the positive peak between 151 and 268 ms was identified on each of the nine denoised electrodes belonging to the P300 (Fcz, C1, Cz, C2, C4, CP1, Cpz, CP2, Pz). Similarly, negative peaks were identified between 191 and 232 ms for the two denoised electrodes belonging to the N200 (TP7 and P9). These values were averaged per seen and unseen condition for each of the 100 matched sets of trials separately. In addition to mean peak amplitude and mean peak latency, the standard deviation of peak latency was also computed for each matched set. Finally, the six parameters (mean peak amplitude, mean peak latency, and standard deviation of peak latency for both the seen and the unseen condition) were averaged over electrodes and subjects. Thus, a grand average of all six parameters for the N200 and the P300 per matched set was obtained. The grand averages were then correlated with the respective onset latencies of the N200 and the P300 as obtained from the 100 contrastive analyses with denoised data.

In addition to the 12 correlation tests described above 4 confirmatory correlation test were also carried out between averaged ERP parameters and gmNCC onset latencies (see S5 in the Supplementary Material for details). All the p-values ($n = 16$) were corrected for multiple comparisons with the Holm–Bonferroni method.
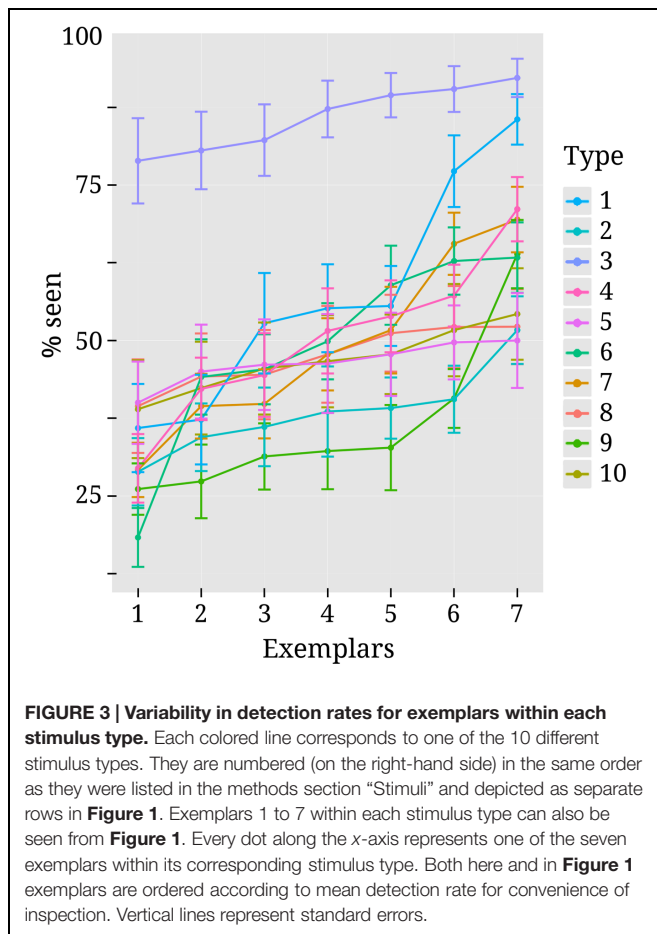
RESULTS

Behavioral Results

The false alarm rate in our study was quite low considering the very faint stimulation. The mean percentage of seen reports for catch trials was 4.2% (median $= 2.9\%$, $SD = 4.5\%$, range $= 0$–17%). Mean detection rate over all stimulus types was close to threshold as intended ($m = 51\%$, median $= 48.6\%$, $SD = 13.8\%$). The high variance in detection rate stems mostly from the fact that contrasts were estimated separately for different types of stimuli. For several subjects, threshold contrast could not be identified equally well for all stimulus types and detection rates were therefore not always clustered evenly around the mean. S6 in the Supplementary Material lists the detection rates for all stimulus types separately and **Figure 3** depicts detection rates for all exemplars within the different stimulus types.

As can be seen from **Figure 3**, detection rates are considerably higher for text stimuli compared to other types of stimuli. This was due to the fact that for 12 out of 18 subjects no precise threshold contrast value could be identified for text stimuli. Depending on the contrast, subjects either perceived close to none of the text stimuli or almost all of them. For those subjects the higher contrast level was selected and this pushed the mean detection rate up. For the other nine stimulus types threshold

FIGURE 3 | Variability in detection rates for exemplars within each stimulus type. Each colored line corresponds to one of the 10 different stimulus types. They are numbered (on the right-hand side) in the same order as they were listed in the methods section "Stimuli" and depicted as separate rows in **Figure 1**. Exemplars 1 to 7 within each stimulus type can also be seen from **Figure 1**. Every dot along the x-axis represents one of the seven exemplars within its corresponding stimulus type. Both here and in **Figure 1** exemplars are ordered according to mean detection rate for convenience of inspection. Vertical lines represent standard errors.

contrasts could be identified more successfully, but there was still variance between individual exemplars. Note, however, that this variability was by and large not systematic across subjects. Most exemplars were perceived above average by some subjects and below average by others.

EEG Markers of Conscious Visual Perception

The first goal of the present study was to identify content-independent general markers of NCC (gmNCC), i.e., markers that distinguish consciously perceived trials of our heterogeneous visual stimulus set from the non-perceived trials. The second goal was to study the reliability and any possible variability in the onset latency of these gmNCC within one study. Importantly, when we refer to the reliability and variability of gmNCC we specifically mean the reliability and variability of significant differences between the seen and unseen condition. We therefore conducted a 100 contrastive analyses on matched subsets of data in order to compare the results with regard to occurrence and timing of the gmNCC. **Figure 4** gives a representative example of the results of one such analysis.

The most reliable difference between the seen and unseen condition was the P300. This component was significant in every one of the 100 contrastive analyses and constituted a cluster of

23 electrodes on average (median $= 23$, $SD = 1$, range $= 21$–25). The onset latency of the P300 component was not as consistent as its occurrence, however. **Figure 5** contains a histogram of all observed onset latencies of the P300. It is obvious that there are two prominent periods of onset. Mean latency of the first onset period was 143 ms after stimulus presentation (median $= 143$, $SD = 6$ ms, range $= 128$–157). Mean latency of the second onset period was 193 ms (median $= 190$, $SD = 13$ ms, range $= 166$–223). The P300 was always significant until the end of the tested time period, i.e., 500 ms.
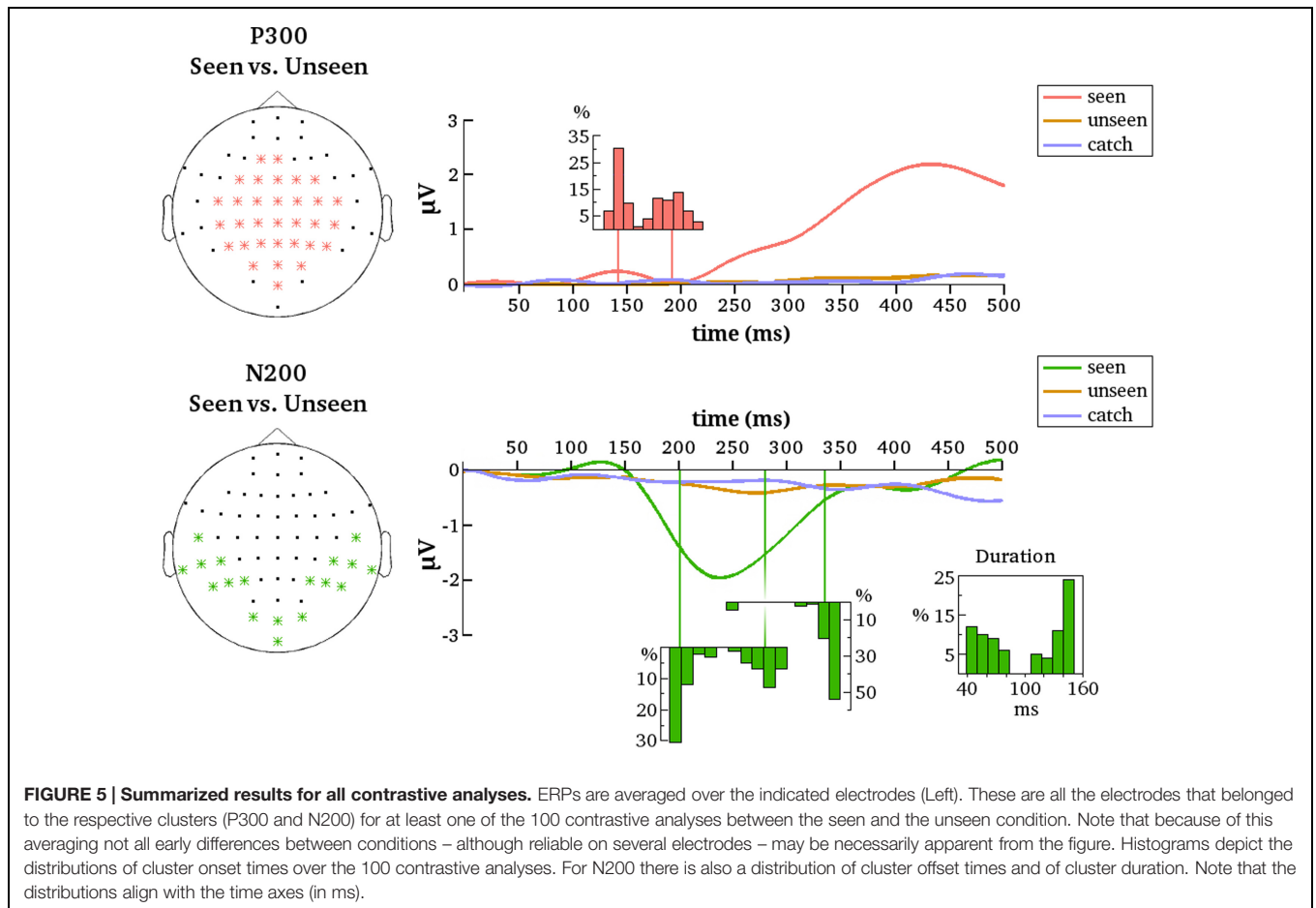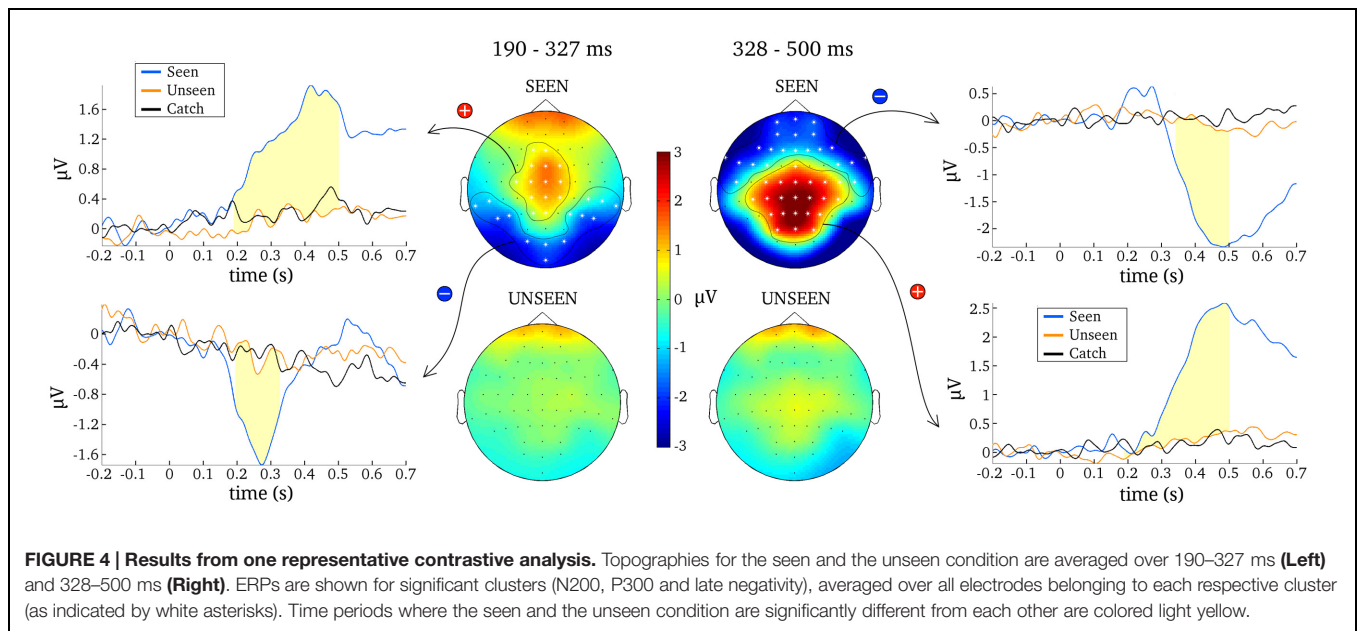
The N200 was significant in only 81 of the 100 contrastive analyses and constituted a cluster of 10 electrodes on average (median $= 11$, $SD = 3$, range $= 4$–15). Thus, in 19% of all cases the contrastive analysis was unable to uncover this gmNCC. Furthermore, even if the N200 was significantly different between the seen and unseen conditions its onset latency nevertheless exhibited considerable variability. As for the P300, there are two prominent periods of onset for the N200. Mean latency of the first onset period was 203 ms (median $= 199$ ms; $SD = 9$ ms, range $= 192$–230 ms). Mean latency of the second onset period was 281 ms (median $= 281$ ms; $SD = 12$ ms, range $= 257$–301 ms). The duration of the N200 was also divided into two groups. The first group lasted 59 ms on average (median $= 57$ ms, $SD = 12$ ms, range $= 40$–82 ms). The second group lasted 137 ms on average (median $= 142$ ms, $SD = 12$ ms, range $= 108$–150 ms). The mean offset of statistical significance was at 336 ms (median $= 341$ ms, $SD = 20$ ms, range $= 245$–350 ms). **Figure 5** contains histograms of the distributions over all iterations.

Finally, the contrastive analyses also yielded a third component in addition to N200 and P300 which we refer to as the late negativity (see S7 in the Supplementary Material for a summary of the respective results). However, the onset latency and topography of this third component suggest that it is probably a consequence of conscious perception (Aru et al., 2012). Another alternative explanation is that N200 together with the primary part of P300 constitutes an early effect of conscious perception while the secondary part of P300 and the late negativity constitute a later effect of conscious experience. We leave this problem out of the scope of the present article, however, and will not concentrate on the late negativity any further.

To test if the above described components are reliably evident only for the seen condition we proceeded by comparing the unseen condition to the catch condition. The 100 matched sets of unseen trials were separately contrasted with all available catch trials where the subjects reported not having seen a stimulus. However, none of the corresponding contrastive analyses yielded any significant differences. Thus, it would seem that a condition where the subject did not perceive a stimulus and a condition where there really was no stimulus are indistinguishable in our present dataset at a statistically significant level.

gmNCC Onset Variability is Partly Explained by Noise

The above described results suggest that the timing of the two gmNCC (N200 and P300) is highly variable even within one study, ranging over 100 ms depending on which trials are

FIGURE 4 | Results from one representative contrastive analysis. Topographies for the seen and the unseen condition are averaged over 190–327 ms **(Left)** and 328–500 ms **(Right)**. ERPs are shown for significant clusters (N200, P300 and late negativity), averaged over all electrodes belonging to each respective cluster (as indicated by white asterisks). Time periods where the seen and the unseen condition are significantly different from each other are colored light yellow.

FIGURE 5 | Summarized results for all contrastive analyses. ERPs are averaged over the indicated electrodes (Left). These are all the electrodes that belonged to the respective clusters (P300 and N200) for at least one of the 100 contrastive analyses between the seen and the unseen condition. Note that because of this averaging not all early differences between conditions – although reliable on several electrodes – may be necessarily apparent from the figure. Histograms depict the distributions of cluster onset times over the 100 contrastive analyses. For N200 there is also a distribution of cluster offset times and of cluster duration. Note that the distributions align with the time axes (in ms).

included in the comparisons. In some cases the N200 was even entirely absent. It follows that some variables characterizing single trials are responsible for the varying results and thus the

third goal of the present study was to identify these variables. As stated in the Section "Introduction," both the signal and the noise profiles of the single trials are potentially involved. It is

thus possible that the above described variability in results is not related to the underlying signal profile of the gmNCC at all, but stems from nuisance factors such as an insufficient signal-to-noise ratio or an unequal noise profile between conditions. In order to rule out this possibility data from representative electrodes were denoised via wavelets and the 100 contrastive analyses were repeated on the same matched sets of trials as before.

After denoising, the onset latency of statistically significant differences again showed considerable variance, albeit with some important differences. The previously observed early period of P300 onsets was effectively not present. Only two from the 100 iterations resulted in P300 onsets earlier than 160 ms. The mean onset latency for the new results was 232 ms (median = 231 ms, SD = 17 ms, range = 151–268 ms). **Figure 6** contains the distribution of all onset latencies after denoising the data. Again, the P300 always remained significant until the end of the tested time period and comprised all the nine electrodes selected for denoising on average (median = 9, SD = 0.1, range = 8–9).

Results also changed for the N200. After denoising only two temporo-parietal electrodes showed significant differences between conditions. We nonetheless decided to go on with the analyses considering clusters starting from two electrodes as significant. For the new results N200 was significant on 97% of the iterations and included three electrodes on average (median = 2, SD = 2, max = 9). The mean onset latency was 208 ms (median = 203 ms, SD = 12 ms, range = 191–232 ms). The mean offset latency was 313 ms (median = 317 ms, SD = 15 ms, range = 261–342 ms). Thus, the mean duration of the N200 was 105 ms (median = 111 ms, SD = 23 ms, range = 39–141 ms). **Figure 6** contains histograms of the respective distributions. Note that the N200 onset and duration displayed a highly negative correlation [r = −0.8, $t(95)$ = −12.95, p = $2.2e − 16$].

To examine if the N200 and the P300 are reliably evident only for the seen condition a separate group of a 100 contrastive analyses comparing the unseen condition to the catch condition were performed. Denoised data were analyzed from the same groups of electrodes as for the seen-unseen comparisons. Recall that no corresponding differences for the undenoised data were found, but perhaps the removal of noise will bring to light some subliminal processing of the stimulus in the unseen condition that was previously missed.

As for the undenoised data, there were no significant differences between the unseen and catch conditions on the central P300 electrodes for the denoised data. Thus, the P300 seems indeed to be reliably evident only for the seen condition. The same is not quite true for N200, however. Results revealed a small but quite consistent negative component on an occipital cluster of electrodes. Note that these are not the same electrodes that were most reliable in the seen-unseen comparison. The occipital cluster was significant on 91% of the iterations and included three electrodes on average (median = 3, SD = 0.21, max = 4). The mean onset latency of statistical significance was 254 ms (median = 256 ms, SD = 17 ms, range = 166–268 ms). The mean offset latency was 302 ms (median = 300 ms, SD = 28 ms, range = 279–491 ms). Thus, the mean duration of the occipital negative cluster was 48 ms (median = 43 ms,

SD = 33 ms, range = 20–235 ms). **Figure 6** contains histograms of the respective distributions. Results thus suggest that a negative cluster on occipital electrodes can reliably differentiate the unseen condition from the catch condition around 250 ms after stimulus onset.

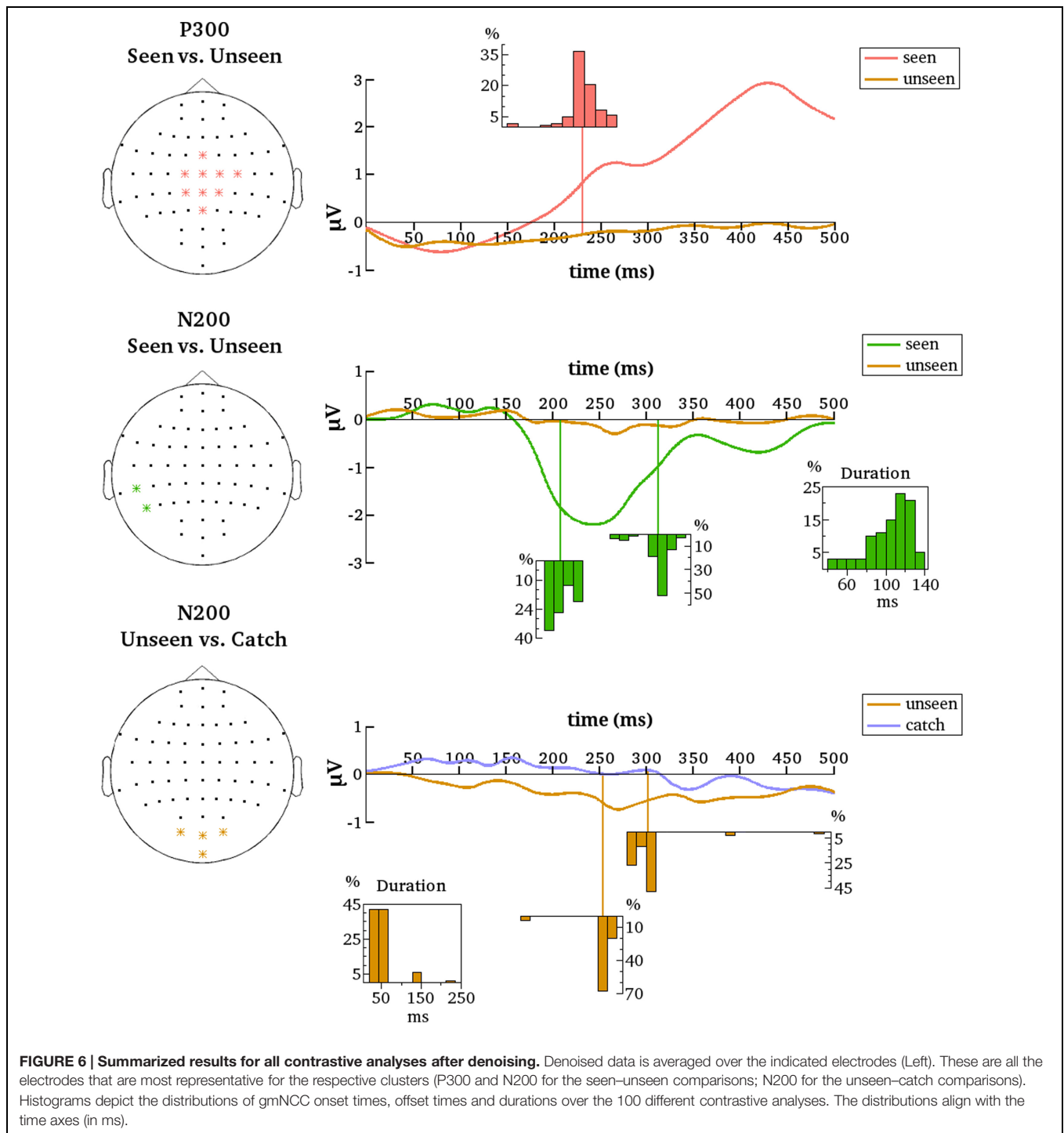## gmNCC Onset Variability Explained by Single Trial Parameters

Results from the previous section indicate that some variability in the gmNCC onset latencies remains even if noise is effectively removed from the data. Thus, some parameters of the gmNCC signal profile must also be involved in the observed variance (see "Introduction" for a description and some theoretical implications of the possible parameters) and it is the third goal of the present study to identify these parameters. Having the list of 100 varying onset times of N200 and P300 one can therefore ask what is different between the matched sets of trials that underlie each of these 100 contrastive analyses.

To answer this question, we first extracted peak amplitude and peak latency of N200 and P300 from the time period of observed variability in onset latencies for both components. Then, grand averages of mean peak amplitude, mean peak latency, and mean latency variance were calculated separately for the seen and the unseen trials and for each of the 100 matched sets, to be subsequently correlated with the 100 different cluster onset latencies (see "Correlation Tests" for more details).

It is important to note that we are presently not analyzing the peaks of the N200 and the P300 components. Because, we are interested in the time period of gmNCC onsets we cannot hope to accurately capture the peaks of the corresponding components in that time window. Our aim is somewhat different. We are trying to understand what happens in the single trials at the time when variance is observed between the 100 contrastive analyses. We are trying to do this by looking at maximal activity in that time window. Because we already have conducted the contrastive analyses, we know that some variables must exist that are responsible for the differences in results. We are now simply taking our analysis one step further by trying to identify these variables.

Figure 7 illustrates the results of all conducted correlation tests. The onset times of P300 correlated significantly neither with mean peak latency of the seen trials (r = −0.05, t = −0.52, p = 1.0) nor with mean peak latency of the unseen trials (r = 0.15, t = 1.54, p = 0.76). The respective correlations with mean latency variance were also not significant (r = 0.04, t = 0.43, p = 1.0 for seen trials; r = −0.02, t = −0.21, p = 1.0 for unseen trials). There was a moderately significant correlation with mean peak amplitude for seen trials (r = −0.3, t = −3.07, p = 0.031), but the most significant correlation was found with mean peak amplitude for unseen trials (r = 0.5, t = 5.67, p = $2.2e − 06$).
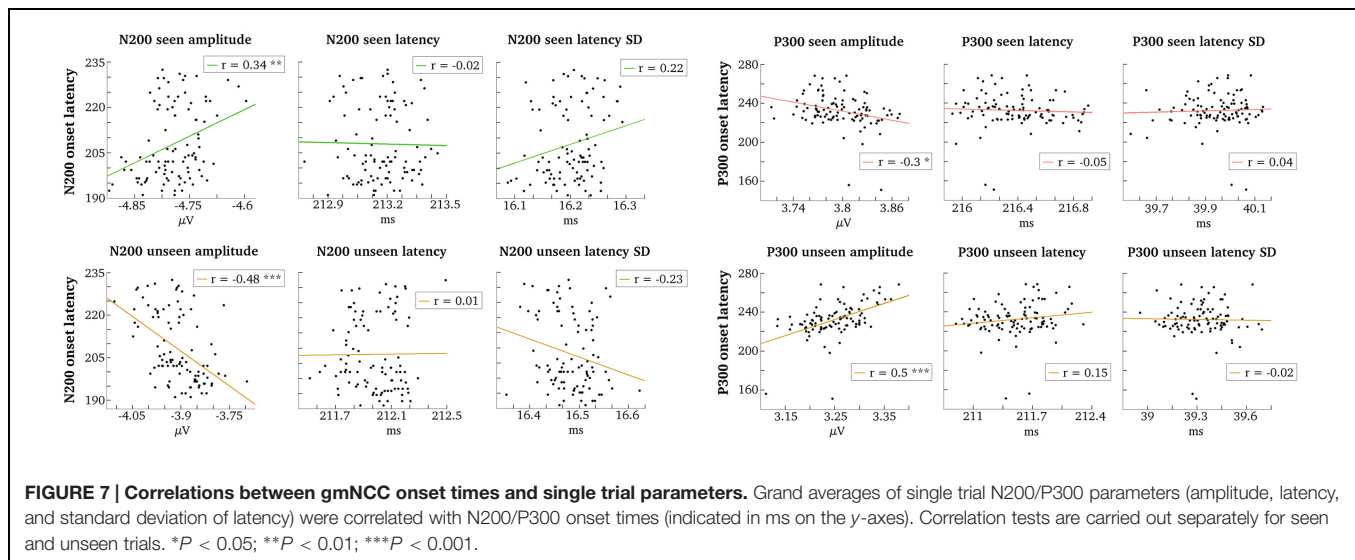
Results were very similar for N200. The onset times of N200 did not correlate significantly with mean peak latency for the seen nor for the unseen trials (r = −0.02, t = −0.18, p = 1.0 and r = 0.01, t = 0.09, p = 1.0, respectively). The correlations with mean latency variance were also not significant (r = 0.22, t = 2.25, p = 0.21 for seen trials; r = −0.23, t = −2.26, p = 0.21 for unseen trials). The correlations with mean peak amplitudes of the seen

FIGURE 6 | Summarized results for all contrastive analyses after denoising. Denoised data is averaged over the indicated electrodes (Left). These are all the electrodes that are most representative for the respective clusters (P300 and N200 for the seen–unseen comparisons; N200 for the unseen–catch comparisons). Histograms depict the distributions of gmNCC onset times, offset times and durations over the 100 different contrastive analyses. The distributions align with the time axes (in ms).

and the unseen trials were again significant ($r = 0.34$, $t = 3.49$, $p = 0.009$ and $r = -0.48$, $t = -5.4$, $p = 7e - 06$, respectively).

To exclude any possible confounds with latency variance and to demonstrate more convincingly the relevance of the amplitude parameter for the observed variability in gmNCC onset times, the above analysis was repeated by first averaging single trials and then extracting peak amplitude. The results are presented in S5 in the Supplementary Material.

Finally, to be sure that the above described results are meaningful and do not derive from the simple fact that any activity in the unseen condition – if at all present – is much weaker compared to the seen condition we repeated all the contrastive analyses and correlation tests, but replaced the post-stimulus time window of the unseen condition with baseline data. Results are described in S8 in the Supplementary Material. These analyses show that both for the P300 and N200 variability in onset

FIGURE 7 | Correlations between gmNCC onset times and single trial parameters. Grand averages of single trial N200/P300 parameters (amplitude, latency, and standard deviation of latency) were correlated with N200/P300 onset times (indicated in ms on the *y*-axes). Correlation tests are carried out separately for seen and unseen trials. $*P < 0.05$; $**P < 0.01$; $***P < 0.001$.

latencies is much decreased compared to the results presented above and no significant correlations with mean peak amplitude of the "unseen" condition (i.e., baseline activity) remain.

We conclude that besides noise the varying onset times of the two gmNCC are first and foremost explained by amplitude variability in the unseen trials, although amplitude variability in the seen trials has an effect as well. If the range of mean peak amplitude values for the seen and the unseen trials in **Figure 7** are compared, it can be noticed that mean peak amplitude of the unseen trials varies over a wider range than mean peak amplitude of the seen trials. Thus, it is not surprising that this variability is reflected in the onset times of significant differences between the seen and the unseen condition.

Importantly, there is no evident connection between gmNCC onset times and latency parameters. And indeed, if one takes a look at the distributions of mean peak P300 and mean peak N200 latencies in **Figure 7**, one can observe that the variability is very small in absolute numbers. It seems that the mean peak latencies of the two gmNCC are very similar across the different matched sets of trials. The distributions of mean latency variance for P300 and N200 in **Figure 7** make it clear that peak latency shifts considerably over single trials, but mean latency variance is again very similar across the 100 different sets of trials.

One of our goals was to study the reliability and any possible variability in the onset latency of gmNCC. The presently reported results bring us closer to an informed answer. We now know that it is not very accurate to only use the onset time of significant differences between the seen and the unseen condition for an estimate of gmNCC latency because this kind of latency estimate may well vary with amplitude fluctuations. In fact, one can disregard a lot of the variance in cluster onset times, because it is caused by possibly irrelevant amplitude fluctuations in the unseen condition. Nevertheless, one cannot fully disentangle the contribution of mean amplitude for the seen trials from the contribution of mean amplitude for the unseen trials. Both vary and both have an effect on the onset time of significant differences.

With this in mind, the best and most reliable estimate for the latencies of our two gmNCCs should not be based solely on results from the contrastive analysis, but also on original latency values for the seen trials in the time window of significance onsets, because these seem to stay surprisingly homogeneous over different sets of trials. Presently, we will use mean peak latency to make the best estimates. For N200 latency the respective estimate is 213 ms (*SD* over subjects = 2.2 ms, range = 210–218 ms). For P300 latency the estimate is 216 ms (*SD* over subjects = 5.5 ms, range = 209–227 ms). Whether the components will be significant at these time points in a given seen–unseen comparison depends a lot on mean amplitudes of the specific selection of the seen and unseen trials included in the comparison. But the mean peak latency of these gmNCC gives an idea of what is going on in the seen trials alone.

General vs. Specific Markers of NCC

Although the aim of this study was to find and describe content-independent general markers of conscious perception it must be noted that not all NCC have to be general. There might exist specific markers of conscious perception which are associated with certain stimulus types only (e.g., N170 for faces; Navajas et al., 2013). Our rationale was to capitalize on a heterogeneous stimulus set so that no stimulus specific markers (whether NCC-pr, NCC proper, and/or NCC-co) could overpower the results and only their common denominators would survive. Thus, we presently did not aim to differentiate between general and specific markers of conscious perception nor to investigate them parametrically. These questions will have to be addressed in future research.

On the other hand, even if a marker is in essence the same for different stimulus types (i.e., it marks the same underlying neural process) its latency and/or amplitude may probably still vary due to stimulus characteristics or perceptual quality. We have conducted some preliminary analysis in this regard as far as the dataset allows. Comparisons between different

subgroups of stimulus types and between stimuli with higher or lower detection rates are presented in S9 in the Supplementary Material. These results do not indicate any influence of stimulus characteristics on N200 and P300 amplitude/latency. However, differences in detection rate seem to be associated with systematic amplitude and/or latency modulations for both N200 and P300.

GENERAL DISCUSSION

The first goal of the present experiment was to find general markers of NCC (gmNCC), that is – markers that distinguish consciously perceived trials of a heterogenous visual stimulus set from the non-perceived trials. The second goal was to study how much these gmNCC vary within one experiment. The third goal was to characterize the causes of this variability. A heterogeneous visual stimulus set was used, presented at a near-threshold contrast. Thus, our paradigm was designed to reduce the influence of stimulus predictability and categorical specificity. One hundred different matched subsets of the resulting seen and unseen trials were contrasted to identify the gmNCC, to study their reliability and variability of their timing. Results indicate that N200 and P300 are the two gmNCC for our paradigm, but their onset latency exhibits considerable variability.

Generality of Various NCC

Differences in the occurrence and onset latency of NCC observed between studies has previously been explained with differences in stimulus material and tasks (Koivisto and Revonsuo, 2010; Dehaene and Changeux, 2011). One explanation is that depending on how visual awareness is manipulated and assessed within a given paradigm, neural prerequisites (NCC-pr) and neural consequences (NCC-co) specific to that paradigm may be misclassified as NCC proper, when the contrastive method is used (Aru et al., 2012; de Graaf et al., 2012). We were able to show, however, that NCC can vary even within one study where the paradigm, stimulus material and recording conditions were kept constant. Admittedly, the present paradigm is also not sufficiently free from possible confounding factors so as to confidently argue that N200 and P300 really are the NCC proper. We can only argue that for our study these ERP components which may be markers of any one of the three subtypes of NCC are general enough so that they do not emerge as related to some narrow visual categorical stimulus group. For this reason, we call them general markers of NCC. The problem is simply that besides general NCC proper there might also exist general NCC-pr or general NCC-co. On the other hand, even with regard to the NCC proper, we should not think that conscious experience marked by it must be invariant and narrowly fixed in time. Conscious experience of the target stimulus need not be indicated by a certain type of strictly defined NCC, but could be understood as a successful evolution of necessarily required neural events over time (see Bachmann, 2000; Navajas et al., 2013; Schurger et al., 2015 for similar arguments).

The P300 component is a well-known marker of conscious perception. It has been found in almost all electrophysiological studies investigating the ERP-correlates of consciousness. Only when the same experimental stimuli are presented repeatedly (Koivisto and Revonsuo, 2008; Sekar et al., 2013) or when one has prior knowledge about the presented stimulus (Melloni et al., 2011) does the P300 increment not occur as a difference between trials with and without conscious perception. As P300 might reflect updating of working memory (Polich, 2007), which is arguably not much needed when the very same stimuli are already encoded in working memory, P300 is not a marker of conscious perception under such experimental conditions (Melloni et al., 2011, but see Rutiku et al., 2015). For the present study stimuli were deliberately unpredictable. Thus, in light of the argumentation presented above it is possible that the P300 is not a gmNCC proper, but rather reflects a general process following the NCC proper, i.e., it represents the NCC-co.

The N200 has also been found as a marker of conscious perception, but not as often as the P300. In many studies the N200 is not reliably different between conditions with and without conscious awareness (Sergent et al., 2005; Del Cul et al., 2007). The present results offer an explanation for these varying results. As the reliability of this marker of conscious perception depends on which single trials are included in the seen as well as the unseen condition it is possible that previous studies have simply missed it. This possibility has also been noted by other researchers (Del Cul et al., 2007). Nonetheless, the present results are different because there is no clear N200 component present in the unseen condition. Studies using stronger stimuli find a well pronounced N200 which is not different between conditions (Sergent et al., 2005; Del Cul et al., 2007). Thus, one might argue that the N200 reflects a general process preceding the NCC proper.

Yet, it seems for the undenoised data that the average onset of P300 occurs somewhat earlier than the average onset of N200. This would be in conflict with the view that N200 reflects a pre-conscious process prior to the NCC proper or NCC-co, which is P300. Another interesting observation is that both components show two periods of onset for the undenoised data. One explanation for these results is that the abnormally distributed results are due to a confounding signal in the measurements (e.g., alpha oscillations) and are actually not a property of the gmNCC *per se*. The current results favor this explanation, because after denoising the relevant single trial data, the discrepant periods of onset disappear. After denoising both components are still reliably associated with conscious perception, but they show one fairly similar period of onset which falls around 200 ms after stimulus presentation. Thus, noise seems to explain a big part of the initial variability in gmNCC onset latencies and the extremely early onset latencies of the P300 in particular.

Despite the fact that the very early period of P300 onsets disappeared after denoising the EEG signal it is noteworthy that P300 still sets on somewhat earlier than is typically estimated in other relevant studies (around 270 ms in Del Cul et al., 2007, for example). One explanation for this discrepancy

may be that we are presently not capturing specifically the onset latency of the P3b subcomponent which is arguably the most relevant P300 subcomponent for conscious perception (Dehaene and Changeux, 2011). P300 also has a somewhat earlier subcomponent – the P3a. It is evident on fronto-central electrodes and is hypothesized to reflect automatic and possibly non-conscious orienting responses (e.g., Muller-Gass et al., 2007). Perhaps in our study a stronger P3a response occurs for the consciously perceived stimuli and this is the earliest critical difference within the P300 that we capture with our contrastive analyses. In that case the earliest part of P300 may just as well reflect a general process of gnNCCpr preceding the NCC proper for our paradigm.

gmNCC Onset Variability Explained

After noise was removed from the data we were able to show that variance in the gmNCC onset latency could be first and foremost attributed to amplitude variance in the unseen condition. Amplitude variance in the seen condition was also associated with the varying gmNCC onsets, albeit to a lesser extent. It is important to note, however, that not only were there no clear N200 and P300 components in the unseen condition, but there really were no clearly pronounced ERP components associated with the unseen condition at all (see Ojanen et al., 2003 for similar results). Thus, the question arises whether this fact in itself could explain the results showing that most of the variance in gmNCC onset latencies came from the unseen condition. To test this possibility we repeated all the analyses after replacing the post-stimulus data of the unseen condition with baseline data. This lead to a marked decrease of variability in gmNCC onset latencies compared to results with actual data and the absence of significant correlations with the amplitude of the unseen condition (i.e., baseline activity) remained. This fact speaks against the possible confound of an unequal signal-to-noise ratio between the seen and the unseen condition in the present study. Furthermore, despite the lack of any clear ERP components in the unseen condition it still exhibited reliable differences with respect to the catch condition on occipital electrodes around 250 ms after stimulus presentation – supporting the assumption that there is a weak signal and thus a weak ERP in the unseen condition. The activity may just be too weak to form a clear component on the ERP.

Although the same occipital electrodes that differentiated the unseen condition from the catch condition sometimes also showed significant differences between the seen and the unseen condition, these were not the most reliable electrodes for the N200 of conscious visual perception. N200 was most reliable on left temporo-parietal electrodes in the present study. Thus, one additional possibility why some previous works have not found the N200 as a marker of conscious perception could be because it is mixed up with other posteriorly recorded components that have similar latencies, but are not necessarily associated with conscious perception.

Taken together, the results reported in this study suggest that signal properties of the unseen condition (amplitude fluctuations in particular) can have a noteworthy impact on the results of a contrastive analysis. Although such effects are generally expected their extent has not been thoroughly investigated in previous studies. However, the present study is comparable to another recent study (Schurger et al., 2015). The authors of this study elegantly showed that the pattern of activity in response to unseen stimuli is less stable within and between trials than the pattern of activity in response to seen trials. Thus, instability may be a property of unconscious neural responses while stability constitutes a hallmark of conscious perception. Our results confirm this assumption, but in addition show that because of this difference in stability comparisons between the seen and unseen condition can yield widely varying results in terms of when and where significant differences begin to occur.

Theoretical Implications

The P300 as a marker of consciousness is most consistent with the theory of a global workspace consisting of multiple areas including frontal, parietal, and temporal cortices (Dehaene et al., 1998, 2003). We cannot say anything certain about the sources of our P300, but since it is a well-studied component one can be fairly confident that a similar multi-focal network is underlying the P300 of the present study.

The N200 is consistent with the visual awareness negativity (Koivisto and Revonsuo, 2003; Wilenius-Emet et al., 2004) concept and the idea of posterior local recurrent activity (Lamme and Roelfsema, 2000). Our N200 component occurs somewhat later than the usual N200 reported previously. This may be due to the faint stimulation. A similar explanation is offered by Sekar et al. (2013). The facts showing that ERP correlates of correct perception have been found at a shorter latency range exemplified by N100-150 (Bachmann, 1994) can be explained as a result of the considerably higher contrast/intensity of the stimuli used, which leads to the speed-up of awareness-related processing and shorter latencies of the negative ERP components reflecting this.

We also did not observe early EEG components in the seen condition (e.g., N100) for the present paradigm. Again, it is likely that these signals are too faint and/or unreliable for the low contrast stimuli used in the present study. This interpretation is backed up by another study (Sekar et al., 2013) where weak stimulation was used. The resulting very small post-stimulus brain response at 100 ms did not differ between conditions at a statistically significant level. Thus, the present results confirm that such early responses do not seem to be markers of direct conscious perception of near-threshold stimuli.

Taken together, our findings show that if a set of heterogeneous stimuli is used, whose identity cannot be predicted by the subject, the two widely reported correlates of consciousness – the N200 and P300 – are reliably observed. However, the onset latencies of these components still showed large variability. Importantly, part of this variability can be attributed to the particular set of trials selected for the condition without conscious perception. These results indicate that any conclusions about the NCC onset timing that are based on data from a single study with its specific stimuli and procedure, are likely to be misleading.

AUTHOR CONTRIBUTIONS

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fnhum. 2016.00023

REFERENCES

Arieli, A., Sterkin, A., Grinvald, A., and Aertsen, A. (1996). Dynamics of ongoing activity: explanation of the large variability in evoked cortical responses. *Science* 273, 1868–1871. doi: 10.1126/science.273.5283.1868

Aru, J., and Bachmann, T. (2009). Boosting up gamma-band oscillations leaves target-stimulus in masking out of awareness: explaining an apparent paradox. *Neurosci. Lett.* 450, 351–355. doi: 10.1016/j.neulet.2008.11.063

Aru, J., Bachmann, T., Singer, W., and Melloni, L. (2012). Distilling the neural correlates of consciousness. *Neurosci. Biobehav. Rev.* 36, 737–746. doi: 10.1016/j.neubiorev.2011.12.003

Bachmann, T. (1994). *Psychophysiology of Visual Masking: The Fine Structure of Conscious Experience*. Commack, NY: Nova Science Publishers.

Bachmann, T. (2000). *Microgenetic Approach to the Conscious Mind*. Amsterdam: John Benjamins.

Bachmann, T. (2009). Finding ERP-signatures of target awareness: puzzle persists because of experimental co-variation of the objective and subjective variables. *Conscious. Cogn.* 18, 804–808. doi: 10.1016/j.concog.2009.02.011

Busch, N. A., Dubois, J., and VanRullen, R. (2009). The phase of ongoing EEG oscillations predicts visual perception. *J. Neurosci.* 29, 7869–7876. doi: 10.1523/JNEUROSCI.0113-09.2009

de Graaf, T. A., Hsieh, P. J., and Sack, A. T. (2012). The 'correlates' in neural correlates of consciousness. *Neurosci. Biobehav. Rev.* 36, 191–197. doi: 10.1016/j.neubiorev.2011.05.012

Dehaene, S., and Changeux, J. P. (2011). Experimental and theoretical approaches to conscious processing. *Neuron* 70, 200–227. doi: 10.1016/j.neuron.2011.03.018

Dehaene, S., Kerszberg, M., and Changeux, J. P. (1998). A neuronal model of a global workspace in effortful cognitive tasks. *Proc. Natl. Acad. Sci. U.S.A.* 95, 14529–14534. doi: 10.1073/pnas.95.24.14529

Dehaene, S., Sergent, C., and Changeux, J. P. (2003). A neuronal network model linking subjective reports and objective physiological data during conscious perception. *Proc. Natl. Acad. Sci. U.S.A.* 100, 8520–8525. doi: 10.1073/pnas.1332574100

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

Koivisto, M., and Revonsuo, A. (2003). An ERP study of change detection, change blindness, and visual awareness. *Psychophysiology* 40, 423–429. doi: 10.1111/1469-8986.00044

Koivisto, M., and Revonsuo, A. (2008). The role of selective attention in visual awareness of stimulus features: electrophysiological studies. *Cogn. Affect. Behav. Neurosci.* 8, 195–210. doi: 10.3758/CABN.8.2.195

Koivisto, M., and Revonsuo, A. (2010). Event-related brain potential correlates of visual awareness. *Neurosci. Biobehav. Rev.* 34, 922–934. doi: 10.1016/j.neubiorev.2009.12.002

Lamme, V. A., and Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends Neurosci.* 23, 571–579. doi: 10.1016/S0166-2236(00)01657-X

Luck, S. J. (2005). *An Introduction to the Event-Related Potential Technique*. Cambridge, MA: MIT Press.

Maris, E., and Oostenveld, R. (2007). Nonparametric statistical testing of EEG– and MEG-data. *J. Neurosci. Methods* 164, 177–190. doi: 10.1016/j.jneumeth.2007.03.024

Mathewson, K. E., Gratton, G., Fabiani, M., Beck, D. M., and Ro, T. (2009). To see or not to see: prestimulus α phase predicts visual awareness. *J. Neurosci.* 29, 2725–2732. doi: 10.1523/JNEUROSCI.3963-08.2009

Melloni, L., Schwiedrzik, C. M., Müller, N., Rodriguez, E., and Singer, W. (2011). Expectations change the signatures and timing of electrophysiological correlates of perceptual awareness. *J. Neurosci.* 31, 1386–1396. doi: 10.1523/JNEUROSCI.4570-10.2011

Muller-Gass, A., Macdonald, M., Schröger, E., Sculthorpe, L., and Campbell, K. (2007). Evidence for the auditory P3a reflecting an automatic process: elicitation during highly-focused continuous visual attention. *Brain Res.* 1170, 71–78. doi: 10.1016/j.brainres.2007.07.023

Navajas, J., Ahmadi, M., and Quian Quiroga, R. (2013). Uncovering the mechanisms of conscious face perception: a single-trial study of the n170 responses. *J. Neurosci.* 33, 1337–1343. doi: 10.1523/JNEUROSCI.1226-12.2013

Ojanen, V., Revonsuo, A., and Sams, M. (2003). Visual awareness of low-contrast stimuli is reflected in event-related brain potentials. *Psychophysiology* 40, 192–197. doi: 10.1111/1469-8986.00021

Overgaard, M. (2004). Confounding factors in contrastive analysis. *Synthese* 141, 217–231. doi: 10.1023/B:SYNT.0000043019.64052.e0

Picton, T. W., Bentin, S., Berg, P., Donchin, E., Hillyard, S. A., Johnson, R. Jr., et al. (2000). Guidelines for using human event-related potentials to study cognition: recording standards and publication criteria. *Psychophysiology* 37, 127–152. doi: 10.1111/1469-8986.3720127

Pins, D., and Ffytche, D. (2003). The neural correlates of conscious vision. *Cereb. Cortex* 13, 461–474. doi: 10.1093/cercor/13.5.461

Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clin. Neurophysiol.* 118, 2128–2148. doi: 10.1016/j.clinph.2007.04.019

Quian Quiroga, R. (2000). Obtaining single stimulus evoked potentials with wavelet denoising. *Physica D* 145, 278–292. doi: 10.1016/S0167-2789(00)00116-0

Quian Quiroga, R., and Garcia, H. (2003). Single-trial event-related potentials with wavelet denoising. *Clin. Neurophysiol.* 114, 376–390. doi: 10.1016/S1388-2457(02)00365-6

Rutiku, R., Martin, M., Bachmann, T., and Aru, J. (2015). Does the P300 reflect conscious perception or its consequences? *Neuroscience* 298, 180–189. doi: 10.1016/j.neuroscience.2015.04.029

Schurger, A., Sarigiannidis, I., Naccache, L., Sitt, J. D., and Dehaene, S. (2015). Cortical activity is more stable when sensory stimuli are consciously perceived. *Proc. Natl. Acad. Sci. U.S.A.* 112, E2083–E2092. doi: 10.1073/pnas.1418730112

Sekar, K., Findley, W. M., Poeppel, D., and Llinás, R. R. (2013). Cortical response tracking the conscious experience of threshold duration visual stimuli indicates visual perception is all or none. *Proc. Natl. Acad. Sci. U.S.A.* 110, 5642–5647. doi: 10.1073/pnas.1302229110

Sergent, C., Baillet, S., and Dehaene, S. (2005). Timing of the brain events underlying access to consciousness during the attentional blink. *Nat. Neurosci.* 8, 1391–1400. doi: 10.1038/nn1549

Shafto, J. P., and Pitts, M. A. (2015). Neural signatures of conscious face perception in an inattentional blindness paradigm. *J. Neurosci.* 33, 1337–1343. doi: 10.1523/JNEUROSCI.0145-15.2015

Wilenius-Emet, M., Revonsuo, A., and Ojanen, V. (2004). An electrophysiological correlate of human visual awareness. *Neurosci. Lett.* 354, 38–41. doi: 10.1016/j.neulet.2003.09.060