

The Role of Attention in Basic Ensemble Statistics Processing

Anton Lukashevich ¹, Heida Maria Sigurdardottir ¹, Nikita Kydryavstsev ², Igor Utochkin ³

¹ Department of Psychology, University of Iceland

² Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University

³ Institute for Mind and Biology, University of Chicago

Author note:

Anton Lukashevich <https://orcid.org/0000-0002-6927-7724>

Heida Maria Sigurdardottir <https://orcid.org/0000-0002-4248-0992>

Igor Utochkin <https://orcid.org/0000-0001-8433-446X>

Funding:

This work was supported by the Icelandic Research Fund (grant number: 228916).

Conflict of interest:

The authors declare no conflicts of interest.

Correspondence should be addressed to Anton Lukashevich, University of Iceland, Sæmundargata 12, 102 Reykjavík, Iceland

E-mail: anl22@hi.is

Abstract

The visual system can represent information about multiple objects in the form of ensemble statistics, such as their mean feature. Although ensemble representation is often considered a strategy to deal with attentional capacity limitations, it is under debate whether it requires attention. We investigated this question using two ERP markers, the P3 which is evoked by change detection of attended stimuli, and visual mismatch negativity (vMMN) which is elicited by automatic sensory discrimination under attention diverted from the critical stimulus. In Experiment 1 (attended ensemble changes), observers attended to ensembles and reported rare changes of their mean orientation, while fixating on a central cross. In Experiment 2 (unattended ensemble changes), participants attended to changes in a central cross and ignored the background ensemble stimuli which also sporadically changed mean orientation. When ensembles were attended, changes in their mean evoked the P3 component, a marker of conscious change detection. When the same ensembles were unattended, no evidence for the vMMN, a marker of automatic discrimination, to ensemble mean changes was found. These results let us suggest that attention is critical for ensemble discrimination.

Keywords: ensemble perception, ERP, attention, vMMN, P3.

Highlights:

- Attention-dependent ensemble processing: Ensemble mean orientation changes are processed only when attended, as indicated by the P3 component, with no vMMN response for unattended changes.
- Distinct neural signatures: ERP components show a clear divide between attended and unattended ensemble statistics, revealing attention's crucial role.
- Challenging automaticity: Our findings suggest that ensemble processing is not automatic, and instead point to a crucial role of attention.

1. Introduction

We are normally exposed to numerous objects at a time in our everyday visual perception. Due to limited attentional and working memory resources, we cannot deeply process all objects simultaneously (Cowan, 2001; Pylyshyn & Storm, 1988). Nevertheless, the visual system utilizes more information than these limitations would suggest. One possible way to accomplish this is by extracting *ensemble summary statistics* from the entire set of objects rather than detailed information about each individual item (Alvarez, 2011; Baek & Chong, 2020). Such ensemble perception can rely on neural substrates distinct from those involved in single object perception (Cant et al., 2015) and access to ensemble properties may precede access to properties of individual items (Epstein & Emmanouil, 2021; Hochstein & Ahissar, 2004). Ensemble perception is not limited to single parameters of a distribution; it encompasses the ability to extract comprehensive statistical properties from a set, including the mean feature (Haberman & Whitney, 2009), feature variability (such as variance or range, Khayat & Hochstein, 2018), and even detailed information about the entire feature distribution, including the probability density function (Chetverikov, 2016; Kim & Chong, 2020). Ensemble perception is not limited to basic, low-level feature dimensions such as orientation and color (Whitney & Yamanashi Leib, 2018) but can extend to complex and even abstract dimensions such as animacy (Leib et al., 2016) and facial expressions (Haberman & Whitney, 2009); for a more comprehensive review, see Corbett et al., 2023). Ensemble perception can occur without phenomenal awareness of the stimuli (Sekimoto & Motoyoshi, 2022) and feature distributions may be encoded implicitly (Hansmann-Roth et al., 2021). Despite showing various signs of automaticity (for example, its implicit and fast character (Han et al., 2021) (Corbett & Oriet, 2011)) and potentially allowing the visual system to overcome its capacity limitations, there is no strong consensus on to what degree ensemble perception relies on limited-capacity mechanisms such as attention.

The current body of literature presents conflicting evidence. Some experiments suggest that even when attention is diverted from a set of items, observers can accurately identify its overall central location and detect global changes in that set, indicating that ensemble information is preserved (Alvarez & Oliva, 2008). Similarly, using a Posner cueing paradigm, no effects of spatial attention were found on the processing of motion ensembles (Talipiski et al.,

2021). Also, mean and range of objects' orientation and size were found to be encoded automatically (Khayat & Hochstein, 2018). Moreover, recent work by Kacin et al. (2023) suggests that ensemble processing ability is at least partially independent of spatial attentional abilities.

Other evidence suggests that ensemble processing requires attention. For instance, in the attentional blink paradigm, encoding the average emotional expression of an ensemble of faces may only occur after attention has been directed to the stimuli (McNair et al., 2017). Jackson-Nielsen et al. (2017) used the inattention blindness paradigm to show that summary statistics changes in color and size may not be registered without attention (Jackson-Nielsen et al., 2017). An attention-demanding secondary task has also been reported to decrease observers' efficiency in making a judgment of average orientation of a set of Gabor patches (Dakin et al., 2009).

In the present study, we address this long-standing question of attentional involvement in processing ensemble statistics by looking at neurophysiological correlates of attentional and non-attentional change detection, namely, event-related potential (ERP) components P3 and the mismatch negativity (MMN). The occurrence of P3 correlates with the detection of a critical target event that the observer is instructed to respond to and that, therefore, is processed under full attention (Polich, 2007). In contrast, the MMN is associated with automatic and non-attentional detection of a change in a repetitive sequence of stimuli (Näätänen, 2000). Through this methodology, we aim to investigate the automaticity of ensemble processing.

The P3 is an ERP component that can be operationally defined as an increased ERP amplitude within a 300 to 500 ms time window after the presentation of a rare target stimulus in an oddball paradigm (Polich, 2012). The P3 may reflect the enhancement of attentional focus, including an early attention-related process due to a change in working memory representations as well as transmission of attention-driven signals from temporal to parietal regions, and may also be related to decision-making processes (Nieuwenhuis, 2005; Polich, 2012). The P3 component shows a dipole pattern with a stronger positive response to deviant stimuli in frontal regions and a more negative response in parietal regions (Polich, 2007). The P3 is usually divided into two subcomponents P3a and P3b (Näätänen et al., 1982). The former is often referred to as the "novelty P3" and reflects the attentional detection of novel, yet task-irrelevant stimuli in the oddball paradigm (Polich, 2003).

The MMN peaks between 160 and 220 ms after the presentation of an unattended oddball stimulus that differs from preceding stimuli (Winkler, 2008). The MMN was initially found in the auditory domain (Näätänen, 2000) but an analogous visual mismatch negativity (vMMN) has since been discovered (Stefanics et al., 2015). The vMMN is thought to be an automatic brain response emerging in the absence of attention or even consciousness since it is elicited by changes of an unattended stimulus even in comatose patients (Morlet & Fischer, 2014). A typical vMMN experimental design for testing the role of attention in stimulus processing involves two tasks, one in which participants are actively involved (e.g., vigilance task, multiple object tracking, Winkler, 2008) and another is a passive oddball task where the occurrence of rare stimuli is unattended and linked to vMMN elicitation. The presence of vMMN can therefore be used as a marker of unattended yet encoded changes in visual input.

The vMMN has been shown to be elicited by changes in a number of visual features such as color, spatial frequency, motion direction, shape, line orientation, location, as well as by conjunctive changes of these features (see Kimura et al., 2011 for a review). The vMMN has also to a limited degree been used in studies on ensembles. Durant et al. (2017) recorded neural responses to a deviant unattended ensemble of randomly oriented lines presented among ordered alternatives and showed that changes in orientation variance can elicit a vMMN. Khvostov et al. (2021) found that unattended background changes of categories defined by certain length-orientation conjunctions of line ensembles elicited a vMMN. While unattended background changes of categories defined by certain length-orientation conjunctions elicited a vMMN response, it could not be detected when categorical differences between changing categories were less distinct. That is also in line with Winkler et al. (2005) who found vMMN evidence that feature binding (in their case, color and orientation) is carried out automatically. The vMMN, therefore, may be sensitive to changes in unattended ensemble statistics, which hints at their automatic nature. However, some types of changes of unattended ensembles in Khvostov et al. (2021) did not evoke a vMMN, which rather suggest that not all ensemble features can be processed automatically.

In the current study, we ask whether changes in ensemble average features can be encoded automatically or whether encoding of ensemble averages depends on the allocation of

attention. In two experiments, we recorded EEG to attended and unattended oddball changes in the ensemble average orientations. In Experiment 1, participants were asked to attend and respond to ensemble mean orientation changes. In Experiment 2, we presented similar streams of orientation ensembles but diverted attention from those streams with a central attentional task requiring participants to track changes at the fixation cross and ignore ensemble changes. We predict different ERP patterns depending on the role of attention in ensemble processing. If ensemble change detection relies on attention, we expect that attended and detectable changes (Experiment 1) should result in a P3 component, while the equivalent unattended changes (Experiment 2) should not be automatically processed and therefore not elicit a vMMN. On the contrary, if ensemble changes can be detected without attention, we expect the presence of a vMMN in Experiment 2.

2. General Materials and Methods

2.1 Participants

Two independent samples of participants took part in Experiments 1 and 2. Whereas the expected effect size for the P3 component is usually high (Cohen's $d_z = 1.86$, Kappenman et al., 2021), our sample size justification for both experiments was based on more conservative estimates from our previous work on the vMMN (Khvostov et al., 2021). We estimated that we would need to recruit at least 18 participants to detect the component ($d_z = 0.8$, $\alpha = 0.05$, power = 0.9, G*Power 3 software (Faul., 2007)). As we expected data loss from potential technical issues (e.g., noisy electrodes, extensive eye blinks), in each experiment we recruited 24 Higher School of Economics (Moscow, Russia) students who participated as volunteers (Experiment 1: 14 female and 10 male, mean age of 25 years old, range 18 to 32 years old; Experiment 2: 16 female and 8 male, mean age of 23 years old, range 18 to 32 years old). One recording from Experiment 1 was corrupted (misplacement of event triggers) and therefore not analyzed, leaving us with a final sample of $N = 23$ in Experiment 1 and $N = 24$ in Experiment 2.

All participants reported normal or corrected to normal vision and no prior history of neurological diseases. All gave written informed consent. The protocol was approved by the

Research Ethics Committee of the Psychology Department, Higher School of Economics and followed the Declaration of Helsinki guidelines.

2.2 Stimuli

In both experiments, stimuli were a set of 64 oriented lines drawn from two different orientation distributions. One distribution was constructed by transforming 64 equally spaced numbers between 0.01 to 0.99 using the logit function: $\log(a/1-a)$. The number list was shifted and scaled to fit into the 25 to 75 degrees range, where 0 is horizontal and 90 is vertical (adapted from Epstein & Emmanouil, 2021). The resulting distribution was skewed which allowed us to use its mirror reflection to get a new distribution with the same range (25 to 75 degrees) and standard deviation (10.41 degrees) but a different mean (40 vs. 60 degrees). This was important to diminish potential confounds from stimulus changes other than mean orientation that might affect the occurrence of vMMN and P3. Both the positively skewed (mean 40 degrees) and negatively skewed (mean 60 degrees) distributions of orientations can be seen in Figure 1.

As we used skewed orientation distributions, orientations drawn from one tail of such distributions are inherently rarer than those from the other tail. This could cause a potential problem, such that any randomly placed item with a rare orientation would be surrounded by multiple items with frequent orientations. This would create highly salient local patches (when individual lines pop out from their surround, Nothdurft, 1993) whose processing can be confounded with ensemble processing. To minimize this possibility, we pregenerated two million random spatial combinations of 64 line orientations from each distribution embedded in an 8x8-cell grid (size 8.6°x8.6°) and selected a subset of combinations where no neighboring lines differed in orientation by more than 25 degrees from each other (half of the distribution range). Ensembles of each trial were drawn from these remaining sets.

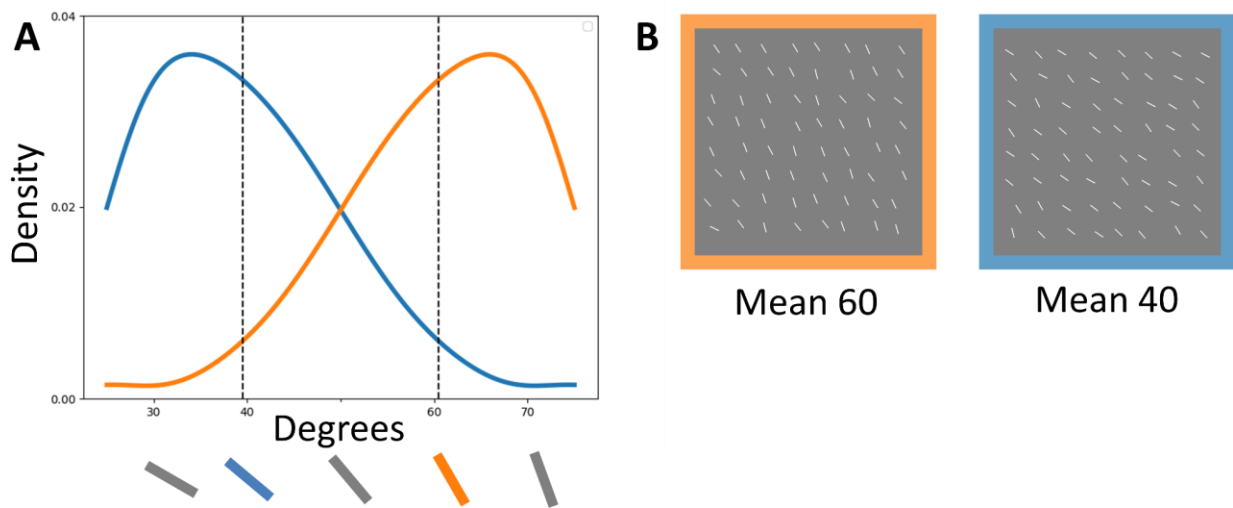


Figure 1. A. A probability density plot representing the orientation distributions from which stimuli were drawn. The blue graph corresponds to the distribution with a mean of 40 degrees. The orange plot corresponds to the distribution with a mean of 60 degrees. **B.** Example ensemble stimuli drawn from these distributions. The ensembles consisted of 64 lines randomly spaced on an 8 x 8 grid. The blue frame corresponds to stimuli drawn from the orientation distribution with a mean of 40 degrees. The orange frame corresponds to the stimuli drawn from the distribution with a mean of 60 degrees. The colored frames are shown here for illustrative purposes and were not presented to participants.

2.3 EEG Recording and Analysis

EEG was recorded using the ActiChamp amplifier with 64-channel active AgCl electrodes (actiChamp Plus, Brain Products GmbH, Gilching, Germany) with both mastoids as reference electrodes. Recording was performed at the 1000 Hz sample rate and re-referenced to the grand average. Off-line filters with a high cut-off at 0.1 Hz and low cut-off at 40 Hz were applied. EEG preprocessing was performed using the BrainVision Analyzer software (BrainVision Analyzer, Brain Products GmbH, Gilching, Germany). Ocular artifacts were corrected using the ocular correction ICA algorithm.

2.4 Procedure

In both experiments, participants completed 1,000 trials divided into 10 blocks of 100 trials each. An intertrial interval was 900 to 1100 ms during which an empty gray screen with a fixation dot at the screen center was presented. Each block was followed by a pause screen. Five

blocks had a standard ensemble mean of 40 degrees and oddball deviant ensemble mean of 60 degrees, and five blocks had standards of mean 60 and deviants of mean 40 degrees (for an example, see Figure 2A). The block order was randomized.

On each trial, a central fixation cross (exp. 1: width 0.16°, height: 0.33°; exp 2: width 0.16°, height: 0.33° or width 0.33°, height: 0.16°) and 64 lines (line length 1°) were presented for 200 ms on a gray background (Figure 2). The lines drawn from the positive and negative skew distributions served as either standard or deviant conditions, depending on the block. The standard trials were presented more frequently (90 standard trials in each block), while the deviant trials were presented less frequently (10 deviant trials in each block). At the beginning of each block, 10 standard trials were always presented. At least three standard trials were presented between deviant trials.

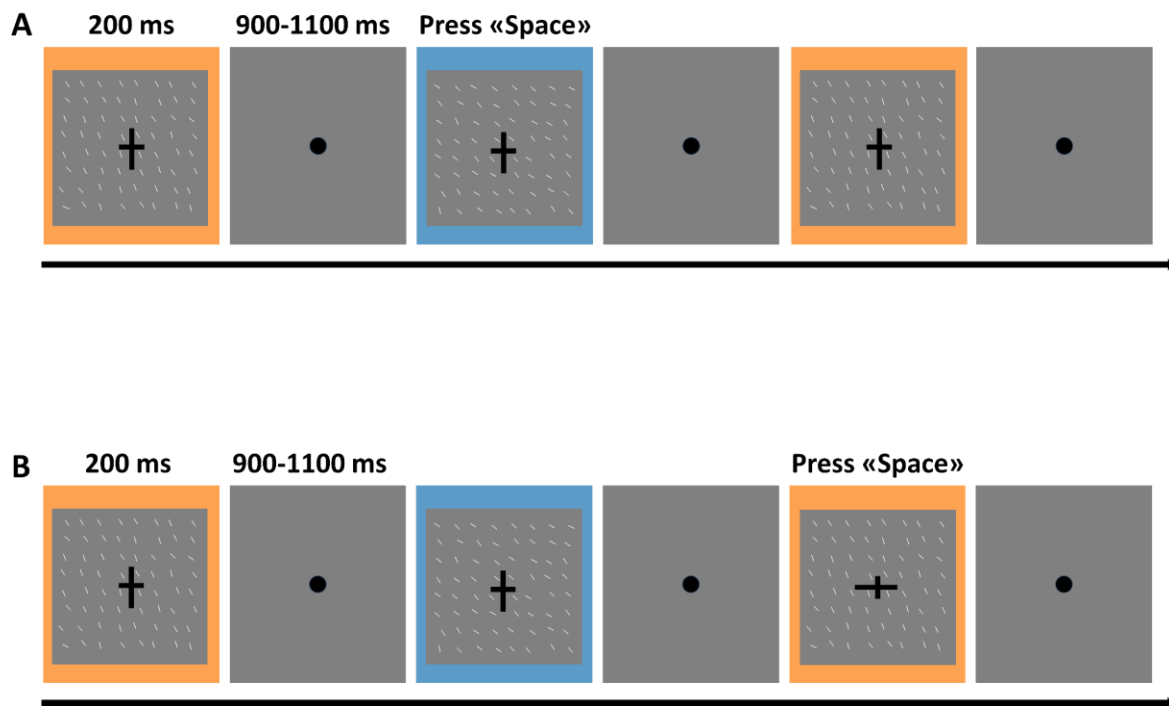


Figure 2. Block schematic of Experiments 1 and 2. Trials were 200 ms with an inter-trial interval of 900-1100 ms. Colored frames (added here for illustrative purposes) correspond to the average orientation of the background stimuli (blue: 40 degrees; orange: 60 degrees). **(A) An example trial sequence of Experiment 1.** Participants were asked to respond to changes of the mean orientation of the white lines, here shown in the second depicted trial marked “press space”. **(B) An example trial sequence of Experiment 2.** Participants were asked to respond to changes of the central cross, here shown in the third depicted trial marked “press space”. The relative sizes of the ensemble, fixation cross and dot are changed for illustrative purposes.

3. Experiment 1: Attended Ensemble Statistics

3.1 Procedure

Participants were asked to explicitly detect the change of the mean orientation of the background line ensemble by pressing the space key (Figure 2A).

3.2 EEG Analysis

The data was epoched from -200 ms to 600 ms from trial onset. Only trials with correct answers (pressing space within the inter-trial interval following deviant trials, and not pressing space within the inter-trial interval following standard trials) were included in the analysis of EEG data as deviants. Standards were selected as trials preceding the deviants. The average number of included trials after preprocessing was 92 for standard epochs (SD = 7) and 86 for deviant epochs (SD = 17). Mean amplitudes for standard and deviant stimuli in frontal, central and parietal electrodes (Fz, F3, F4, F5, F6, F7, F8, AFz, Cz, Pz) were compared using repeated measures ANOVA in a time window spanning 300 to 500 ms after trial onset. Additionally, we ran a Bayesian repeated measures ANOVA within the same time window and electrode regions to calculate Bayes Factor ($\log BF_{10}$, Cauchy distribution as a prior for the fixed effects with a scale parameter of 0.5 was used). The time window and electrode sites were based on previously known characteristics of the P3 component (Kappenman et al., 2021).

3.3 Results

Behavioral Results

A signal-detection discriminability measure, d' was calculated as a $z(\text{HR}) - z(\text{FA})$ for each participant, where HR (hit rate) is a proportion of correctly detected deviant stimuli and FA (false alarm rate) is a proportion of standard stimuli that participants responded to with a button press. The average d' was 3.1 (SD = 1.0) which is consistent with high sensitivity to attended changes in average orientations.

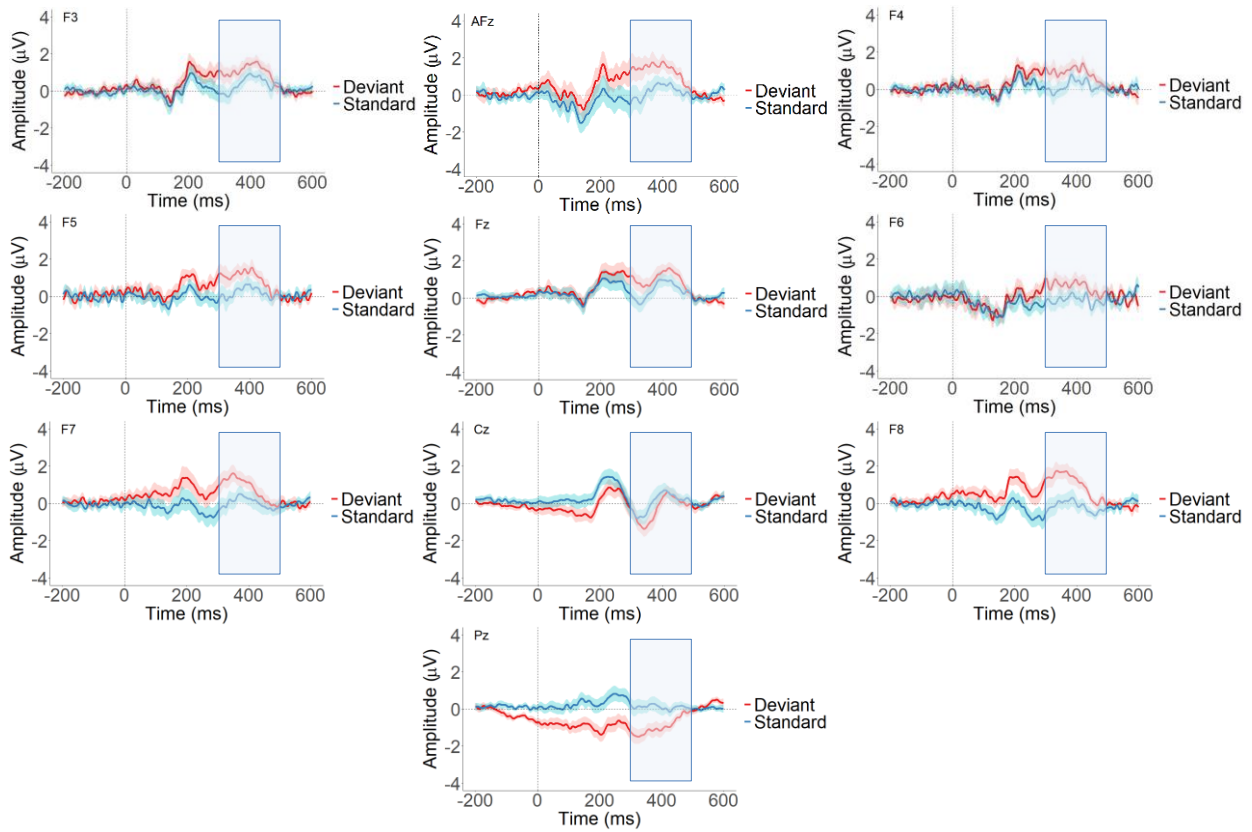


Figure 3. Grand average waveforms in electrodes showing a significant P3 elicited by changes in attended ensemble means in Experiment 1. Shaded areas around lines show ± 1 standard error of the mean with subtracted inter-subject variability. The time window of 300 to 500 ms used in the P3 analysis is highlighted.

EEG Results

The mean amplitude in the time window for standard and deviant waveforms in frontal, central, and parietal electrodes (Fz, F3, F4, F5, F6, F7, F8, AFz, Cz, Pz) was compared using repeated-measures ANOVA (figure 3). A Greenhouse-Geisser correction of degrees of freedom was applied in cases where Mauchly's test indicated a violation of the assumption of sphericity. The analysis revealed significant differences in mean amplitude between standard and deviant stimuli ($F(1,22) = 19.51$, $p < 0.001$, $\eta^2_p = 0.47$, $\log BF_{10} = 4.753$). An effect of electrode site was revealed ($F(2.81,61.84) = 7.29$, $p < 0.001$, $\eta^2_p = 0.25$, $\log BF_{10} = 7.379$), as well as an interaction between stimulus type and electrode ($F(3.47,76.43) = 17.81$, $p < 0.001$, $\eta^2_p = 0.45$, $\log BF_{10} = 12.094$). Post-hoc comparisons revealed significant differences between standards and deviants

in sites AFz ($t = 4.728$, $p_{\text{holm}} < 0.001$), F3 ($t = 4.229$, $p_{\text{holm}} = 0.0003$), F4 ($t = 4.446$, $p_{\text{holm}} < 0.001$), F5 ($t = 4.636$, $p_{\text{holm}} < 0.001$), F6 ($t = 4.974$, $p_{\text{holm}} < 0.001$), F7 ($t = 4.423$, $p_{\text{holm}} < 0.001$), Fz ($t = 2.938$, $p_{\text{holm}} = 0.008$), and Pz ($t = -5.940$, $p_{\text{holm}} < 0.001$).

The scalp potential distribution of the difference wave between activity in response to standards and deviants (figure 4) showed the expected dipole signal distribution between frontal and posterior electrode sites in a time window of 300 to 500 ms. Those spatial and temporal characteristics are common for P3 (Polich, 2007).

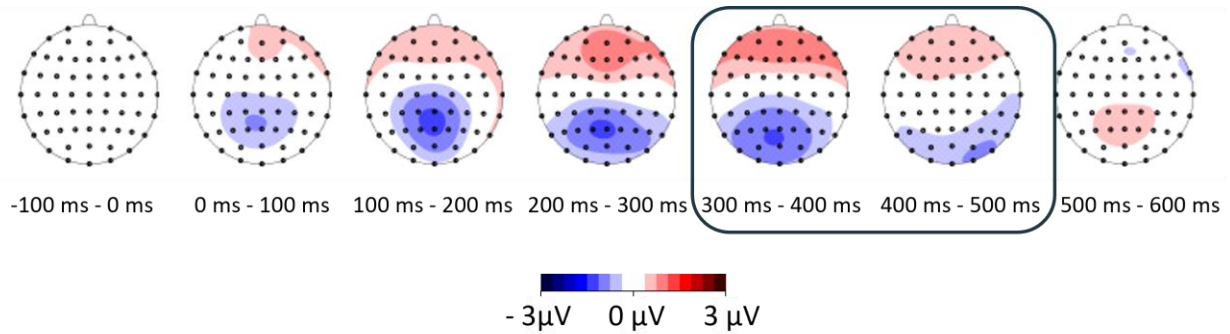


Figure 4. Scalp potential distribution of a difference wave between standard and deviant waveforms in Experiment 1. An expected dipole pattern (more positive in frontal regions and negative in posterior) is observed in the P3 time window (300 to 500 ms, highlighted).

4. Experiment 2: Unattended Ensemble Statistics

4.1 Procedure

In this experiment, participants were instructed to ignore the background stimuli and detect changes in the orientation of the fixation cross. The cross had two possible orientations: vertical (long vertical line and short horizontal line) or horizontal (long horizontal line and short vertical line). The vertical cross was presented on 70% of trials, while the horizontal cross was presented on 30% of trials. Participants were instructed to press the space key immediately after the orientation of the fixation cross changed (Figure 2B). The time window for response was limited to the inter-trial interval (from 900 to 1100 ms). The fixation cross did not change its shape on deviant ensemble trials or the standard ensemble trial preceding a deviant trial.

4.2 EEG Analysis

EEG data was epoched into segments of 700 ms, thereof 200 ms baseline prior to trial onset and 500 ms after trial onset. Standard epochs corresponded to standard trials immediately preceding deviant trials. Deviant epochs corresponded to deviant trials with a change of ensemble orientation mean. The average number per participant of each epoch type after artifact rejection was 98 for standards ($SD = 3.4$) and 99 for deviants ($SD = 3.1$).

We looked for a vMMN by comparing mean amplitudes for deviant ensemble epochs and standard ensemble epochs preceding deviants in a 160-220 ms time window after trial onset in posterior electrodes (O1, O2, Oz, PO7, PO8) using repeated measures ANOVA. Time window and regions of interest for analysis were a priori selected based on the existing literature on vMMN (Czigler, 2007) and our previous research (Khvostov et al., 2021). Bayes Factor (BF) was additionally calculated via a Bayesian repeated measures ANOVA within the same time window and electrode regions.

Additionally, the mean amplitudes for standard and deviant responses in frontal, central and parietal electrodes (Fz, F3, F4, F5, F6, F7, F8, AFz, Cz, Pz) was compared using repeated measures ANOVA in a later P3 (P3a and P3b) time window of 300 to 500 ms, to test whether the participants could be aware of the appearance of infrequent deviant ensemble background stimuli, and Bayes Factor was calculated.

4.3 Results

Behavioral Results

To estimate participants' engagement in the central attentional task, the signal-detection theory sensitivity index d' was calculated the same way as in Experiment 1, but now for the detection of the central fixation cross. The average d' across participants was 4.3 ($SD = 0.8$). Based on such good performance, we conclude that our participants were attentionally engaged in the central task, which is important for the interpretation of the potential vMMN as a neural response to unattended stimuli (oriented ensembles).

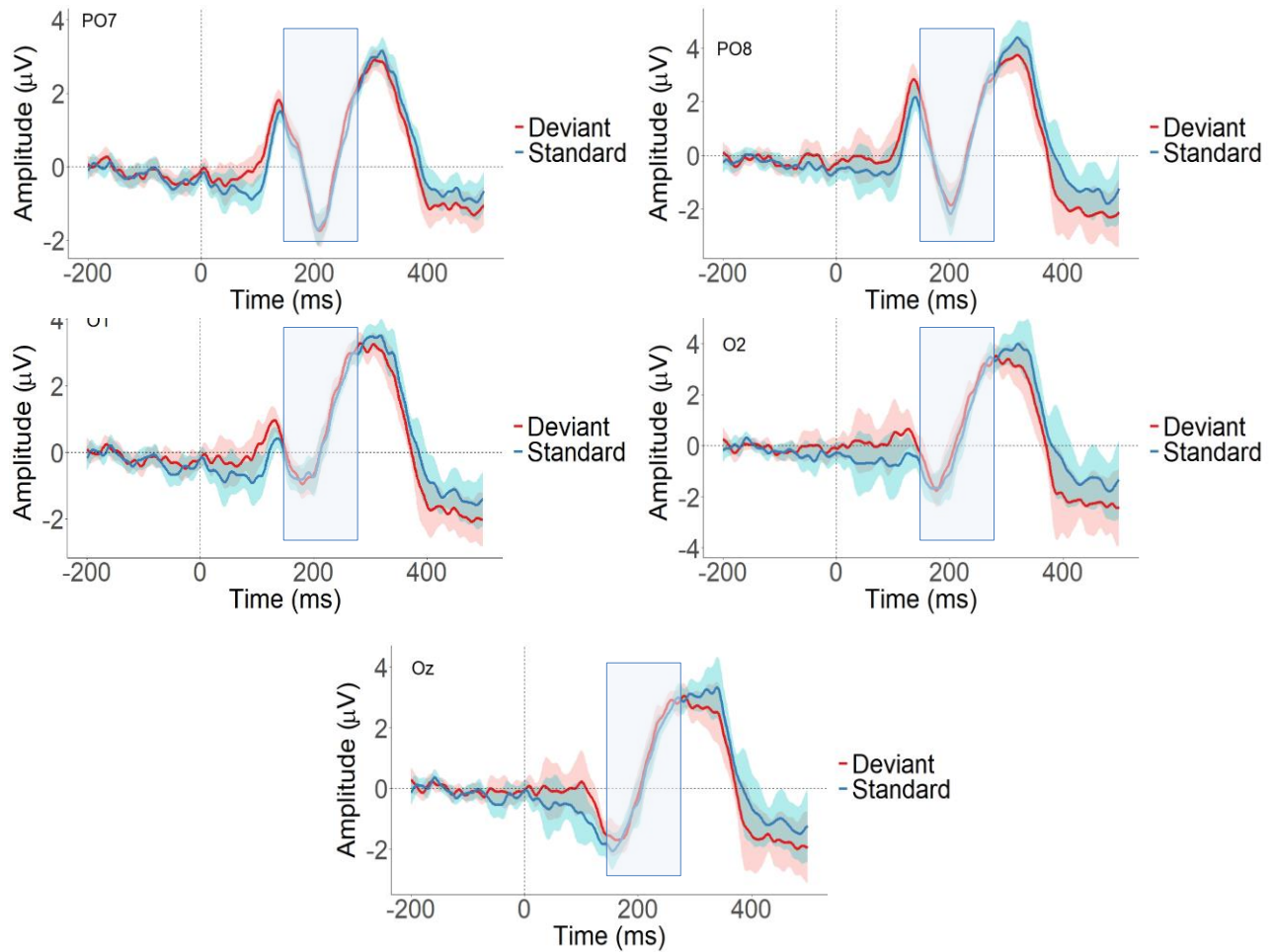


Figure 5. Grand average waveforms in posterior electrodes (PO7, PO8, O1, O2, Oz) in experiment 2. Shaded areas around lines show ± 1 standard error of the mean with subtracted inter-subject variability. The time window of 160 to 260 ms used in the vMMN analysis is highlighted.

EEG Results

We compared mean ERP amplitudes to unattended background ensemble mean changes (deviants) and their preceding standard background ensembles in the vMMN time window of 160 to 260 ms using repeated measures ANOVA (electrodes: O1, O2, Oz, PO7, PO8). A Greenhouse-Geisser correction of degrees of freedom was applied as needed. No effect of stimulus type were found ($F(1, 23) = 0.359$, $p = 0.555$, $\eta^2 = 0.007$, $\log BF_{10} = -0.505$) indicating comparable responses to both standard and deviant stimuli within the specified time window. No effect of electrode was found ($F(1.46, 32.573) = 1.250$, $p = 0.289$, $\log BF_{10} = -0.748$) and no interaction between electrode site and stimulus type ($F(1.302, 29.955) = 0.868$, $p = 0.487$, \log

$BF_{10} = -0.935$). Grand average ERP for standards and deviants in posterior electrode sites and scalp distributions of difference waves are displayed in figures 5 and 6.

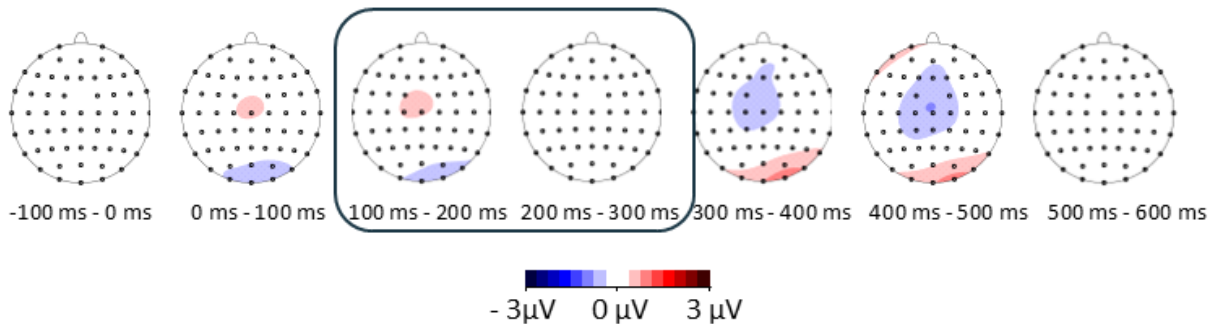


Figure 6. Scalp potential distribution of a difference wave between standard and deviant waveforms in Experiment 2. Time window selected for analysis was 160 to 260 ms. The color scale is kept identical to that of figure 4 from Experiment 1.

Additional ANOVA in Fz, F3, F4, F5, F6, F7, F8, AFz, Cz, Pz electrode sites in the 300 to 500 ms P3 time window (for comparison to Experiment 1) revealed no effect of stimulus type ($F(1, 23) = 2.075$, $p = 0.163$, $\eta^2p = 0.083$, $\log BF_{10} = -0.938$), no effect of electrode (with Greenhouse-Geisser correction: $F(3.234, 74.384) = 25.445$, $p = 0.08$, $\eta^2p = 0.091$, $\log BF_{10} = -0.408$) and no interaction (with Greenhouse-Geisser correction: $F(2.259, 51.951, 23) = 1.465$, $p = 0.240$, $\eta^2p = 0.060$, $\log BF_{10} = -1.274$.)

Overall, these results suggest that the appearance of deviant ensemble stimuli is not processed when unattended.

5. Discussion

We used ERP markers of automatic and attentional visual processing (vMMN and P3, respectively) to test whether the visual system is capable of registering an unattended change of an ensemble mean or if this process requires the deployment of attention. Participants could easily explicitly discriminate between two ensemble means when ensembles were in the focus of their attention. This was supported both by high response accuracy and the presence of the P3 component, considered to be generated in response to attentional allocation and change of neural representation of stimuli (Polich, 2003). On the contrary, we found no evidence of a vMMN in response to a change in ensemble mean when task-irrelevant ensemble stimuli were

presented in the background and attention was engaged with a central task; the lack of attention to ensemble changes was further supported by no detectable P3.

The vMMN is expected to be triggered by unattended – yet registered – deviations from a previously encoded visual input (Kremláček et al., 2016). It is known to be elicited by changes in the orientation of a single unattended line while people are engaged in a central attentional task (File & Czigler, 2019). The vMMN is a marker of pre-attentive automatic processing that cannot be explained by simple differential refractoriness of neuronal populations and requires the forming of memory representation of stimulus regularities (Czigler et al., 2002). We interpret the lack of a vMMN for rare yet unattended ensemble changes in the current study – which were easily detected when attended – as an inability to build an expectation that the mean ensemble feature remains constant. Our results instead indicate a crucial role of attention in the discrimination of orientation statistics.

Do our findings directly contradict the existing pool of literature, claiming the possibility of implicit encoding of summary statistics? Not necessary. We propose that this discrepancy arises from varying definitions of implicit perception. In previous studies, implicit perception often refers to the absence of an explicit ensemble task, even though ensemble stimuli might still be attentively processed. For example, Hansmann-Roth et al. (2021) found evidence for implicit ensemble perception in this sense. However, their participants nonetheless engaged in a visual search for an oddball target that could only be defined in relation to the properties of distractors. This may have required active attentional suppression of ensemble (distractor) stimuli. Similarly, Khayat et al. (2023) showed that participants' responses were biased towards the mean of a previously shown task-irrelevant ensemble. However, these ensembles were passively viewed while people's attention was not engaged elsewhere, so participants were free to actively attend to the ensembles if they so chose. These studies suggest that ensemble statistics can be extracted automatically in the sense that such information is extracted without an instruction to perform an ensemble task. However, they do not rule out the necessity of attention because the elements of the ensemble may at least sometimes be attentively processed. In contrast to these studies, we directly varied attentional setting in the ensemble task, making the ensembles either attended and task-relevant (Experiment 1) or not attended and completely task-irrelevant (Experiment 2).

Under such circumstances, when ensemble stimuli are likely completely ignored, we see no indication of the processing of ensemble means.

Durant et al. (2017) nonetheless demonstrated that the vMMN is elicited by unattended changes in ensemble orientation variance, suggesting the detection of variance outside the focus of attention. However, this vMMN was elicited for very large differences in variance. In our previous study, we also observed a vMMN in conditions where the mean difference between ensembles was extremely high, even though the mean manipulation was not the focus of the study. Similarly, Alvarez and Oliva (2009) demonstrated that participants could discriminate between ensembles with a 45-degree difference under reduced attention. Taken together, this suggests that automatic detection of ensemble means may indeed occur, but only when the differences are large and coarse. Smaller differences, while being well discriminable in the explicit ensemble task, when attention is involved, may not reach the threshold for automatic processing.

6. Conclusions

Overall, our findings indicate that the extraction of ensemble statistics, such as the mean orientation from a visual scene, likely requires attentional resources and cannot be performed independently of attention.

7. Open Science

Data is deposited and can be obtained from <https://osf.io/hgr7f/>

References

- Ahissar, M., & Hochstein, S. (2004). The reverse hierarchy theory of visual perceptual learning. *Trends in Cognitive Sciences*, 8(10), 457–464. <https://doi.org/10.1016/j.tics.2004.08.011>
- Alvarez, G. A. (2011). Representing multiple objects as an ensemble enhances visual cognition. *Trends in Cognitive Sciences*, 15(3), 122–131. <https://doi.org/10.1016/j.tics.2011.01.003>
- Alvarez, G. A., & Oliva, A. (2008). The representation of simple ensemble visual features outside the focus of attention: Research article. *Psychological Science*, 19(4), 392–398. <https://doi.org/10.1111/j.1467-9280.2008.02098.x>
- Alvarez, G. A., & Oliva, A. (2009). Spatial ensemble statistics are efficient codes that can be represented with reduced attention. *Proceedings of the National Academy of Sciences*, 106(18), 7345–7350. <https://doi.org/10.1073/pnas.0808981106>
- Baek, J., & Chong, S. C. (2020). Ensemble perception and focused attention: Two different modes of visual processing to cope with limited capacity. *Psychonomic Bulletin & Review*, 27(4), 602–606. <https://doi.org/10.3758/s13423-020-01718-7>
- Cant, J. S., Sun, S. Z., & Xu, Y. (2015). Distinct cognitive mechanisms involved in the processing of single objects and object ensembles. *Journal of Vision*, 15(4), 12. <https://doi.org/10.1167/15.4.12>
- Chen, Z., Zhuang, R., Wang, X., Ren, Y., & Abrams, R. A. (2021). Ensemble perception without attention depends upon attentional control settings. *Attention, Perception & Psychophysics*, 83(3), 1240–1250. <https://doi.org/10.3758/s13414-020-02067-2>
- Chetverikov, A., Campana, G., & Kristjánsson, Á. (2016). Building ensemble representations: How the shape of preceding distractor distributions affects visual search. *Cognition*, 153, 196–210. <https://doi.org/10.1016/j.cognition.2016.04.018>
- Corbett, J. E., & Oriet, C. (2011). The whole is indeed more than the sum of its parts: Perceptual averaging in the absence of individual item representation. *Acta Psychologica*, 138(2), 289–301. <https://doi.org/10.1016/j.actpsy.2011.08.002>
- Corbett, J. E., Utochkin, I., & Hochstein, S. (2023). *The Pervasiveness of Ensemble Perception: Not Just Your Average Review* (1st ed.). Cambridge University Press. <https://doi.org/10.1017/9781009222716>
- Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and brain sciences*, 24(1), 87–114.
- Czigler, I., Balázs, L., & Winkler, I. (2002). Memory-based detection of task-irrelevant visual changes. *Psychophysiology*, 39(6), 869–873. <https://doi.org/10.1017/S0048577202020218>

- Dakin, S. C., Bex, P. J., Cass, J. R., & Watt, R. J. (2009). Dissociable effects of attention and crowding on orientation averaging. *Journal of Vision*, 9(11), 28–28. <https://doi.org/10.1167/9.11.28>
- Durant, S., Sulykos, I., & Czigler, I. (2017). Automatic detection of orientation variance. *Neuroscience Letters*, 658, 43–47. <https://doi.org/10.1016/j.neulet.2017.08.027>
- Epstein, M. L., & Emmanouil, T. A. (2021). Ensemble Statistics Can Be Available before Individual Item Properties: Electroencephalography Evidence Using the Oddball Paradigm. *Journal of Cognitive Neuroscience*, 33(6), 1056–1068. https://doi.org/10.1162/jocn_a_01704
- Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods*, 39(2), 175–191. <https://doi.org/10.3758/bf03193146>
- File, D., & Czigler, I. (2019). Automatic detection of violations of statistical regularities in the periphery is affected by the focus of spatial attention: A visual mismatch negativity study. *European Journal of Neuroscience*, 49(10), 1348–1356. <https://doi.org/10.1111/ejn.14306>
- Jeffreys, H. (1998). *The theory of probability*. OUP Oxford.
- Haberman, J., & Whitney, D. (2009). Seeing the mean: Ensemble coding for sets of faces. *Journal of Experimental Psychology: Human Perception and Performance*, 35(3), 718–734. <https://doi.org/10.1037/a0013899>
- Han, L., Yamanashi Leib, A., Chen, Z., & Whitney, D. (2021). Holistic ensemble perception. *Attention, Perception, & Psychophysics*, 83(3), 998–1013. <https://doi.org/10.3758/s13414-020-02173-1>
- Hansmann-Roth, S., Kristjánsson, Á., Whitney, D., & Chetverikov, A. (2021). Dissociating implicit and explicit ensemble representations reveals the limits of visual perception and the richness of behavior. *Scientific Reports*, 11(1), 1–12. <https://doi.org/10.1038/s41598-021-83358-y>
- Hochstein, S., Khayat, N., & Pavlovskaya, M. (2023). Comparing explicit and implicit ensemble perception. *Journal of Vision*, 23(9), 4886. <https://doi.org/10.1167/jov.23.9.4886>
- Jackson-Nielsen, M., Cohen, M. A., & Pitts, M. A. (2017). Perception of ensemble statistics requires attention. *Consciousness and Cognition*, 48, 149–160. <https://doi.org/10.1016/j.concog.2016.11.007>
- Kappenman, E. S., Farrens, J. L., Zhang, W., Stewart, A. X., & Luck, S. J. (2021). ERP CORE: An open resource for human event-related potential research. *NeuroImage*, 225, 117465.
- Kacin, M., Cha, O., & Gauthier, I. (2022). The relation between ensemble coding of length and orientation does not depend on spatial attention. *Vision*, 7(1), 3.
- Khayat, N., & Hochstein, S. (2018). Perceiving set mean and range: Automaticity and precision. *Journal of Vision*, 18(9), 23. <https://doi.org/10.1167/18.9.23>

- Khayat, N., Pavlovskaya, M., & Hochstein, S. (2023). Comparing explicit and implicit ensemble perception: 3 stimulus variables and 3 presentation modes. *Attention, Perception, & Psychophysics*, 1–21. <https://doi.org/10.3758/s13414-023-02784-4>
- Khvostov, V. A., Lukashevich, A. O., & Utochkin, I. S. (2021). Spatially intermixed objects of different categories are parsed automatically. *Scientific Reports*, 11(1), 1–8. <https://doi.org/10.1038/s41598-020-79828-4>
- Khvostov, V. A., & Utochkin, I. S. (2019). Independent and parallel visual processing of ensemble statistics: Evidence from dual tasks. *Journal of Vision*, 19(9), 3–3.
- Kim, M., & Chong, S. C. (2020). The visual system does not compute a single mean but summarizes a distribution. *Journal of Experimental Psychology: Human Perception and Performance*, 46(9), 1013–1028. <https://doi.org/10.1037/xhp0000804>
- Kimura, M., Schröger, E., & Czigler, I. (2011). Visual mismatch negativity and its importance in visual cognitive sciences. *NeuroReport*. <https://doi.org/10.1097/WNR.0b013e32834973ba>
- Kremláček, J., Kreegipuu, K., Tales, A., Astikainen, P., Pöldver, N., Näätänen, R., & Stefanics, G. (2016). Visual mismatch negativity (vMMN): A review and meta-analysis of studies in psychiatric and neurological disorders. *Cortex*, 80, 76–112. <https://doi.org/10.1016/j.cortex.2016.03.017>
- Leib, A. Y., Kosovicheva, A., & Whitney, D. (2016). Fast ensemble representations for abstract visual impressions. *Nature Communications*, 7(1), Article 1. <https://doi.org/10.1038/ncomms13186>
- McNair, N. A., Goodbourn, P. T., Shone, L. T., & Harris, I. M. (2017). Summary statistics in the attentional blink. *Attention, Perception, & Psychophysics*, 79(1), 100–116. <https://doi.org/10.3758/s13414-016-1216-2>
- Morlet, D., & Fischer, C. (2014). MMN and Novelty P3 in Coma and Other Altered States of Consciousness: A Review. *Brain Topography*, 27(4), 467–479. <https://doi.org/10.1007/s10548-013-0335-5>
- Näätänen, R., Simpson, M., & Loveless, N. E. (1982). Stimulus deviance and evoked potentials. *Biological psychology*, 14(1-2), 53–98. [https://doi.org/10.1016/0301-0511\(82\)90017-5](https://doi.org/10.1016/0301-0511(82)90017-5)
- Näätänen, R. (2000). Mismatch negativity (MMN): Perspectives for application. *International Journal of Psychophysiology*, 37(1), 3–10. [https://doi.org/10.1016/S0167-8760\(00\)00091-X](https://doi.org/10.1016/S0167-8760(00)00091-X)
- Nieuwenhuis, S., Aston-Jones, G., & Cohen, J. D. (2005). Decision making, the P3, and the locus coeruleus-norepinephrine system. *Psychological bulletin*, 131(4), 510–532. <https://doi.org/10.1037/0033-2909.131.4.510>
- Nothdurft, H.-C. (1993). Saliency effects across dimensions in visual search. *Vision Research*, 33(5-6), 839–844. [https://doi.org/10.1016/0042-6989\(93\)90202-8](https://doi.org/10.1016/0042-6989(93)90202-8)

- Polich, J. (Ed.). (2003). *Detection of change: event-related potential and fMRI findings* (p. 200). Norwell, MA:: Kluwer Academic Publishers.
- Polich, J. (2007). Updating P300: An Integrative Theory of P3a and P3b. *Clinical Neurophysiology : Official Journal of the International Federation of Clinical Neurophysiology*, 118(10), 2128. <https://doi.org/10.1016/j.clinph.2007.04.019>
- Polich, J. (2012). Neuropsychology of P300. *The Oxford Handbook of Event-Related Potential Components*, 641, 159–188.
- Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision*, 3(3), 179–197. <https://doi.org/10.1163/156856888x00122>
- Sekimoto, T., & Motoyoshi, I. (2022). Ensemble perception without phenomenal awareness of elements. *Scientific Reports*, 12(1), Article 1. <https://doi.org/10.1038/s41598-022-15850-y>
- Stefanics, G., Astikainen, P., & Czigler, I. (2015). Visual mismatch negativity (vMMN): A prediction error signal in the visual modality. In *Frontiers in Human Neuroscience* (Vol. 8, Issue JAN). <https://doi.org/10.3389/fnhum.2014.01074>
- Talipski, L. A., Goodhew, S. C., & Edwards, M. (2022). No effect of spatial attention on the processing of a motion ensemble: Evidence from Posner cueing. *Attention, Perception, & Psychophysics*, 84, 1845–1857.
- Utochkin, I. S., & Vostrikov, K. O. (2017). The numerosity and mean size of multiple objects are perceived independently and in parallel. *PLOS ONE*, 12(9), e0185452. <https://doi.org/10.1371/journal.pone.0185452>
- Whitney, D., & Yamanashi Leib, A. (2018). Ensemble Perception. *Ssrn*, August 2017, 1–25. <https://doi.org/10.1146/annurev-psych-010416-044232>
- Winkler, I. (2008). Interpreting the Mismatch Negativity. *Journal of Psychophysiology*. <https://econtent.hogrefe.com/doi/10.1027/0269-8803.21.34.147>
- Winkler, I., Czigler, I., Sussman, E., Horváth, J., & Balázs, L. (2005). Preattentive binding of auditory and visual stimulus features. *Journal of Cognitive Neuroscience*, 17(2), 320–339. <https://doi.org/10.1162/0898929053124866>