Interrelation of attention and prediction in visual processing

Citation for published version (APA):

Marzecová, A., Widmann, A., SanMiguel, I., Kotz, S. A., & Schröger, E. (2017). Interrelation of attention and prediction in visual processing: Effects of task-relevance and stimulus probability. *Biological Psychology*, 125, 76-90. <https://doi.org/10.1016/j.biopsycho.2017.02.009>

Document status and date:

Published: 01/04/2017

DOI:

[10.1016/j.biopsycho.2017.02.009](https://doi.org/10.1016/j.biopsycho.2017.02.009)

Document Version:

Publisher's PDF, also known as Version of record

Document license:

Taverne

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Interrelation of attention and prediction in visual processing: Effects of task-relevance and stimulus probability

Anna Marzecová^{a,b,*}, Andreas Widmann^a, Iria SanMiguel^{c,d,e}, Sonja A. Kotz^{b,f},
Erich Schröger^a

^a Institute of Psychology, University of Leipzig, Neumarkt 9-19, 04109 Leipzig, Germany

^b Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstraße 1A, 04103 Leipzig, Germany

^c Brainlab—Cognitive Neuroscience Research Group, Dept. of Clinical Psychology and Psychobiology, University of Barcelona, Barcelona, Spain

^d Institute of Neurosciences, University of Barcelona, Barcelona, Spain

^e Institut de Recerca Pediàtrica Hospital Sant Joan de Déu, Barcelona, Spain

^f Faculty of Psychology and Neuroscience, Department of Neuropsychology & Psychopharmacology, Maastricht University, Maastricht, Universiteitssingel 40, 6229 ER Maastricht, The Netherlands

ARTICLE INFO

Article history:

Received 27 January 2016

Received in revised form 23 February 2017

Accepted 27 February 2017

Available online 9 March 2017

Keywords:

Prediction

Attention

Visual perception

Predictive coding

ERP

Sensory gain

vMMN

ABSTRACT

The potentially interactive influence of attention and prediction was investigated by measuring event-related potentials (ERPs) in a spatial cueing task with attention (task-relevant) and prediction (probabilistic) cues. We identified distinct processing stages of this interactive influence. Firstly, in line with the attentional gain hypothesis, a larger amplitude response of the contralateral N1, and Nd1 for attended gratings was observed. Secondly, conforming to the attenuation-by-prediction hypothesis, a smaller negativity in the time window directly following the peak of the N1 component for predicted compared to unpredicted gratings was observed. In line with the hypothesis that attention and prediction interface, unpredicted/unattended stimuli elicited a larger negativity at central-parietal sites, presumably reflecting an increased prediction error signal. Thirdly, larger P3 responses to unpredicted stimuli pointed to the updating of an internal model. Attention and prediction can be considered as differentiated mechanisms that may interact at different processing stages to optimise perception.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

The brain needs to process multiple incoming sensory stimuli selectively. To do so effectively, it may exploit statistical regularities in the environment to predict what happens next, and thereby minimise surprise. Consequently, several influential theories of perception consider that the brain forms predictions about incoming sensory information based on Bayesian inferential principles (Friston, 2005, 2009; Lee & Mumford, 2003). Predictions, also referred to as perceptual expectations, are defined as top-down signals that facilitate perception by utilising information about prior probability (Schröger, Marzecová, & SanMiguel, 2015; Summerfield & de Lange, 2014). Predictions are compared to incoming sensory input, and an observed difference is expressed in the prediction

error, a signal that is passed upwards in the sensory hierarchy to update and refine the current model. Within this predictive coding framework, perception can be described as an inferential process of minimising prediction errors through the integration of bottom-up sensory input with the top-down modulation effected by prediction. Attention, defined as a mechanism that prioritises processing of sensory information that is relevant for current goals (Summerfield & Egner, 2009), constitutes another source of top-down influence. While the neural responses to attended vs. unattended, and predicted vs. unpredicted stimuli have been widely studied, the interactive influence of these two factors has rarely been considered, and studies conducted so far show a seemingly inconsistent pattern of findings (for reviews see Schröger et al., 2015; Summerfield & de Lange, 2014).

Attending to goal-relevant stimuli facilitates their detectability and the behavioural response to them. On the neuronal level, the attentional effect is reflected in increased and less variable neuronal responses to attended stimuli (Carrasco, 2011). The effects of spatial attention on event-related potentials (ERP) have mainly been studied in spatial cueing tasks, in which cues direct spatial atten-

* Corresponding author at: BioCog—Cognitive incl. Biological Psychology, Neumarkt 9-19, 04109 Leipzig, Germany.

E-mail addresses: anna.marzecova@uni-leipzig.de, anna.marzecova@gmail.com (A. Marzecová).

tion to the most likely location of an upcoming lateralised target in either a sustained or a transient manner. In these paradigms, attended stimuli usually evoke enhanced early responses, namely the parieto-occipital P1—a positive deflection with a latency of 80–130 ms – and the N1—a negative deflection with a latency of 120–200 ms (for reviews see Herrmann & Knight, 2001; Hillyard & Anillo-Vento, 1998; Luck, Woodman, & Vogel, 2000). The effects of attention have been observed mostly for electrode-sites contralateral to the stimulus, and the enhancements have been attributed to the mechanism of sensory gain (Hillyard, Vogel, & Luck, 1998). However, an enhancement of the ipsilateral P1 component has also been reported. This enhancement presumably reflects the inhibition of potentially competing information processing in the task-irrelevant visual field (for a review see Klimesch, 2011, 2012). In studies in which the effects of attention have been tested by means of transient probabilistic cues (for ERP studies see e.g., Eimer, 1994; Hillyard, Luck, & Mangun, 1994; Mangun & Hillyard, 1991), an additional negative component, specific to trial-by-trial manipulations of attention, has been reported at midline parietal electrodes (“Nd1”, Eimer, 1994; Schröger & Eimer, 1993). Nd1 is thought to reflect spatially selective and modality unspecific activity within the posterior parietal cortex occurring with transient – visual and auditory – spatial attention (for reviews see Eimer, 1998; Näätänen, Alho, & Schröger, 2002). In spatial cueing studies, effects of attentional orienting are observed already in the pre-stimulus period (e.g., Dale, Simpson, Foxe, Luks, & Worden, 2008; Simpson et al., 2011; Yamaguchi, Tsuchiya, & Kobayashi, 1995). Attention cues elicit lateralised activity that is manifested in several cue-related ERP components (Harter, Miller, Price, Lalonde, & Keyes, 1989; Nobre, Sebestyen, & Miniussi, 2000; van der Lubbe, Neggers, Verleger, & Kenemans, 2006; van der Lubbe & Utzerath, 2013; Velzen & Eimer, 2003). The earliest of these is a posterior component between 150 and 350 ms, known as the early directing attention negativity (EDAN). The EDAN likely reflects attentional selection of the relevant parts of the cue (Talsma et al., 2007; Talsma, Mulckhuysen, Slagter, & Theeuwes, 2007; van der Lubbe et al., 2006; Velzen & Eimer, 2003). The EDAN is usually followed by the anterior directing attention negativity (ADAN, ~400 ms) and the late directing attention positivity (LDAP, ~500–650 ms) at posterior sites.

In probabilistic spatial cueing tasks, attention is directed to the expected and more probable location or feature predicted by the cue, and the attention-related enhancement of the N1 component is usually observed (e.g., Doherty, Rao, Mesulam, & Nobre, 2005; Eimer, 1997). However, an opposite pattern, namely a larger amplitude response to invalidly cued events, has also been observed (‘an inverse N1 effect’; Eimer, 1993, Experiment 1a). This pattern has been attributed to the confounding effects of prediction (Eimer, 1993; Lange, 2013). Predictions are formed based on stimulus probability and are induced in a spatial cueing task through probabilistic manipulations of attention (Summerfield & de Lange, 2014; Summerfield & Egner, 2009, 2014). Predictions facilitate the interpretation of sensory data, and lead to more efficient behavioural responses in a similar fashion as attentional selection (see e.g. Kok, Jehee, & de Lange, 2012, for review see Summerfield & de Lange, 2014). However, task-relevance, that drives attentional selection, and signal probability that leads to sensory predictions may constitute potentially orthogonal sources of information flow (Summerfield & Egner, 2009; Summerfield & de Lange, 2014). The underlying neural signatures of prediction appear to differ from the markers of attentional selection. Diverse evidence indicates that while attended as compared to unattended input increases neural responses predicted input usually elicits reduced ERP responses as compared to unpredicted input (for review see Schröger et al., 2015; see also Summerfield & de Lange, 2014; Summerfield & Egner, 2009). For example, in a number of oddball studies, low-probability

unexpected stimuli (i.e., deviants) elicited larger responses than high-probability expected stimuli (i.e., standards). In ERPs, this differential response is indicated by the mismatch negativity (MMN; auditory MMN: Näätänen, Gaillard, & Mäntysalo, 1978; visual MMN: Czigler, Balázs, & Winkler, 2002). Within a predictive coding framework, this effect is explained as a reduction of prediction error responses to sensory input that matches predictions generated by an internal model (see e.g. Garrido, Kilner, Stephan, & Friston, 2009). The phenomenon of repetition suppression (RS; for a review see Grill-Spector, Henson, & Martin, 2006), namely an “experience-related” adaptation, or attenuation, of the N1 component in response to repeated stimuli is considered to partially contribute to the MMN response (see e.g. Horváth et al., 2008; Stefanics, Kremláček, & Czigler, 2014). Furthermore, a suppression of the auditory N1 component for stimuli predicted based on one’s own action (for reviews see Horváth, 2015; Hughes, Desantis, & Waszak, 2012; Schröger et al., 2015) has also been interpreted as a reduction of sensory consequences to a stimulus that is predicted, and thus explained by an internal model. Consequently, this effect has been considered to reflect a reduction of prediction error signal (for review see Schröger et al., 2015). The attenuation effect was also observed for the visual N1 component (see e.g. Gentsch & Schütz-Bosbach, 2011; Roussel, Hughes, & Waszak, 2014). Based on the evidence from the diverse research lines reviewed above, it appears that ERPs in the time window of the N1 component are sensitive to both attention (i.e., increased for attended stimuli) and prediction (i.e., suppressed for predicted stimuli).

Taking a predictive coding perspective, attention and prediction can also be considered as interdependent mechanisms. Recent studies (Brown & Friston, 2013; Feldman & Friston, 2010) have proposed that attention can be understood as a top-down gain control mechanism that optimises the precision of prediction errors through synaptic gain modulation (see also Bastos et al., 2012). The effects of prediction, reflected in the amplitude of the prediction error, would therefore be dependent on the degree of attentional precision. Evidence for an interface between prediction and attention comes from an fMRI study employing a modified spatial cueing task. Kok, Rahnev, Jehee, Lau, and de Lange (2012) manipulated selective endogenous attention and prediction independently through two types of cues: probabilistic (‘prediction’) cues and task-relevant cues. Probabilistic cues presented at the beginning of each block indicated the likely location of visual stimuli in the subsequent block of trials. Across blocks, stimuli appeared either with higher probability (75%) on one side of the fixation (predicted condition) than on the other (unpredicted condition), or with equal probability on either side (no prediction blocks). Task-relevant cues that preceded the stimulus on every trial indicated which side to attend and respond to. The results showed interactive effects of attention and prediction on the Blood Oxygenation Level Dependent (BOLD) response in primary visual cortex (V1). Response suppression to predicted compared to unpredicted input was observed for the unattended (i.e., task-irrelevant) visual stimuli. However, consistent with the hypothesis that the prediction error is scaled by attentional precision, the prediction effect was reversed when stimuli were attended to (i.e., task-relevant), and an enhanced BOLD response to predicted compared to unpredicted gratings was observed. In ERP studies, markers of attentional and predictive processing have largely been studied separately, in attentional cueing and oddball studies, respectively. However, the degree of interdependence between prediction-related and attentional processes is under debate (Kimura, 2012; Stefanics, Kremláček, & Czigler, 2014).

In the current study, we utilised the ERP method to investigate the time course of attentional selection, prediction, and their possible interaction in the context of a spatial cueing task. Analysing the time-course of these possible interactive effects seems important

for understanding whether these two mechanisms are indeed fully interdependent, or function in an independent manner along different stages of a processing hierarchy. To investigate the time-course of these effects, we used a modified variant of Posner's spatial cueing task (Posner, 1980) proposed by Kok, Rahnev et al. (2012), in which attention (operationalised as task-relevance) and prediction (operationalised as the probability of stimulus occurrence) were manipulated by two independent cues. We combined EEG and eye-tracking to control for artefacts elicited by small eye movements towards the to-be-attended location (Kennett, van Velzen, Eimer, & Driver, 2007), and to distinguish effects of different input at the retinal level. In the study by Kok, Rahnev et al. (2012), eye tracking was monitored only for half of the participants, and the trials with eye-movements were not eliminated from analyses, hence a full control for fixation was not ensured.

Firstly, we hypothesised that we would observe dissociable effects of attention and prediction. On the one hand, consistent with previous findings of an increased sensory gain for attended input (for reviews see Herrmann & Knight, 2001; Hillyard & Anillo-Vento, 1998; Luck et al., 2000), we expected the attention effect to result in the enhancement of the N1 component in response to task-relevant visual stimuli compared to task-irrelevant stimuli (*attentional gain hypothesis*). Furthermore, as a result of the transient, trial-by-trial manipulation of attention, we expected to observe an N1 to attended stimuli (Eimer, 1994; Schröger & Eimer, 1993). On the other hand, based on the evidence for the attenuation of the N1 component for predicted stimuli (see paragraph 3); we expected to observe a prediction effect in the N1 time window. We assumed that unpredicted stimuli would elicit a larger negative deflection than predicted ones (*attenuation-by-prediction hypothesis*). We expected that this prediction effect would be followed by an increased positivity in the P3 range. Termed as novelty P3 (or P3a; for reviews see Friedman, Cycowicz, & Gaeta, 2001; Polich, 2007), this component is known to follow the earlier effects of prediction violation and has been related to enhanced processing of surprising or unexpected stimuli (Nieuwenhuis, De Geus, & Aston-Jones, 2011). Secondly, we aimed to uncover the time-course of the possible *interrelation between attention and prediction*. In line with the hypothesis that prediction errors are weighted by attentional precision (see paragraph 4), and based on previous fMRI results (Kok, Rahnev et al., 2012), we expected a prediction effect (i.e., suppression of the N1 component for predicted stimuli) in the task-irrelevant condition, and a reversed pattern (i.e., an increased amplitude for predicted relative to unpredicted stimuli) in the task-relevant condition. Finally, effects of spatial attention (e.g., Dale et al., 2008; Simpson et al., 2011; Yamaguchi et al., 1995), prediction (e.g., Wöstmann, Herrmann, Wilsch, & Obleser, 2015), and their interaction (Todorovic, Schoffelen, van Ede, Maris, & de Lange, 2015) can be observed in pre-stimulus period. Therefore, we also investigated if probabilistic and task-relevant information is integrated in the cue-target interval. To this end, we examined whether the early cue-related ERP components are modulated by cue congruity by comparing ERPs to task-relevant cues in the congruent (i.e., task-relevance and probabilistic cues indicate the same direction) and incongruent (i.e., task-relevance and probabilistic cues indicate opposite directions) condition.

2. Methods

2.1. Participants

Twenty-three participants (16 female, 7 male) took part in the study. Six participants were excluded from further analysis due to EEG artefacts or excessive eye movements. The mean age of the remaining seventeen participants (11 female, 6 male)

was 26.4 years (SD = 4.7, range: 21–35 years). Participants were predominantly right-handed (mean lateralisation quotient: 90%, SD = 23%; range 14–100%; Oldfield, 1971) volunteers with normal or corrected-to-normal vision and no history of psychiatric or neurological impairment, recruited through a database of participants at the Institute of Psychology, University of Leipzig. They either received course credits or were reimbursed for their participation. After being informed about the nature of the study, they gave written informed consent to participate.

2.2. Stimuli and apparatus

Stimuli consisted of luminance-defined sinusoidal gratings with fixed phase and a spatial frequency of 3.33 cpd, embedded in uniform random noise (60% Michelson contrast; see Fig. 1). Their size was 3° and they were positioned on the horizontal meridian either 3.5° to the right or to the left of the fixation cross (from the center of the grating). The gratings were oriented either horizontally (5°) or vertically (95°; 50% each). The contrast of the stimuli was adjusted individually for each participant in a weighted up-down adaptive procedure (Kaernbach, 1991; see below for the description of the procedure). The resulting mean contrast of the gratings was 8.6% (range: 7–10.5%, SD 1.2%, Michelson contrast). MATLAB (The Math-Works) in conjunction with the Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997) was used for creating and presenting the stimuli, and for recording responses. The experiment was presented on a 19" CRT (G90fB, ViewSonic, Walnut, CA) monitor (resolution 1024 × 768, refresh rate of 100 Hz).

2.3. Procedure

Participants were seated comfortably in a dimly lit, electrically shielded chamber, and viewed the display from a distance of 57 cm, with their heads stabilised by a chinrest. Their task was to identify the orientation (horizontal or vertical) of gratings presented either to the left or right of the fixation cross. Each individual trial started with a presentation of task-relevant cue, which was comprised of two greater-than signs (">>") to direct attention to the right visual field (RVF) on 50% of the trials, or two less-than signs ("<<") to direct attention to the left visual field (LVF) on the remaining 50%. The cues were presented in the center of the computer screen for 200 ms. After a variable cue-target interval (600–900 ms), during which a fixation cross ("+" sign) was displayed in the center of the screen, a grating was presented for 50 ms in the LVF or RVF (see Fig. 1 for a schematic representation of a trial timeline). Participants were asked to identify the orientation of the grating (horizontal or vertical) only if it appeared in the visual field indicated by the task-relevant cue (i.e., a response was required if the attention cue pointed to the RVF and the grating appeared in the RVF; while no response was expected if the attention cue pointed to the LVF but the grating appeared in the RVF, and vice-versa). Participants responded by pressing one of two buttons on a response box with the index or middle finger of their right hand, respectively. The matching of the grating orientation and response buttons was counterbalanced across participants. The response window lasted 1650 ms, followed by an inter-trial interval varying between 50 and 350 ms.

The task was organised in short blocks of 16 trials each. Each block started with the presentation of the probabilistic cue, which indicated the most likely spatial location of the gratings. The probabilistic cue was either the word "LINKS" ["LEFT"], indicating that gratings would likely (i.e., with 75% probability) appear in the LVF, or the word "RECHTS" ["RIGHT"] indicating that the gratings would likely (i.e., with 75% probability) appear in the RVF. A neutral ('no-prediction') condition was also included, in which the word "BEIDE" ["BOTH"] indicated that gratings would appear in the LVF or in the

Fig. 1. Schematic representation of the experimental design. Trial timeline.

RVF with an equal probability (50%). Verbal cues were presented in the center of the screen for 1000 ms. The three types of blocks were presented in a pseudo-randomised order to prevent repetitions of the same block type.

In total, the task consisted of 960 trials divided into 60 blocks of 16 trials each. The experimental blocks were administered in 20 triplets of 3 blocks each. In between each triplet, participants were able to rest for variable intervals. Experimental blocks were preceded by a training run, in which participants got familiarised with the task (1 block of each kind).

Stimulus contrast was adjusted in the beginning of the experiment, using an adaptive staircase procedure based on Kaernbach (1991). For this procedure, the trial sequence and timing were kept similar as in the experimental task, but the following aspects were modified: 1) block-wise cues were not presented; 2) trial-by-trial cues were kept, but they did not inform participants about the task-relevance of stimuli (they were composed of two greater-than signs (“>>”) and two less-than signs (“<<”) superimposed on top of each other); 3) noise embedded gratings were presented in both visual fields with an equal probability; 4) a response was required on each trial. As in the experimental task, participants were asked to identify the orientation of the grating. The initial grating contrast was set to 20% (Michelson contrast). The just noticeable difference was set at 90%, thus following the weighted up-down rule; the contrast value decreased with each correct response by one step (1%) and increased with each incorrect response by nine steps (9%). The task continued until the 12th reversal (i.e., an incorrect response followed by a correct response or a correct response followed by an incorrect response). The resulting contrast value was calculated as the median of contrast values from the last eight reversals. The task was repeated twice and the contrast value calculated from the second run was used in the experiment.

2.4. Eye tracking

To ensure that participants maintained central fixation during the trials, we recorded their eye-movements with an infrared eye-tracking system (EyeLink 1000; SR Research) at a sampling rate of 500 Hz. The monocular recordings were controlled by the EyeLink Software. All trials, during which the gaze was not on the fixation cross (area $\pm 1^\circ$ around the center of the screen) at the time of the presentation of the grating (50 ms), were excluded from the subsequent EEG analysis (12.8%).

2.5. Electroencephalogram recording

The electroencephalogram (EEG) was continuously recorded at a sampling rate of 500 Hz, from 59 Ag/AgCl active electrodes using a BrainAmp amplifier and the Vision Recorder 1.20 software (Brain Products). Electrodes were mounted onto an elastic cap (actiCAP, Brain Products) following the extended international 10–20 system (Chatrian, Lettich, & Nelson, 1985). An electrode placed on the tip of the nose served as an online reference, a ground electrode was placed on the forehead, and two electrodes were attached to the earlobes for the purpose of offline re-referencing. The electroocu-

logram (EOG) was recorded using electrodes placed at the outer canthi, and below and above (electrode Fp1) one eye.

2.6. Behavioural data analysis

Effects of probabilistic cues were examined on mean response times (RT) and accuracy by means of one-way ANOVAs with prediction as a within-participant factor (predicted, no prediction, unpredicted). The predicted condition included trials that were validly indicated by the probabilistic cue, i.e., appeared on the side indicated correctly by the prediction cue (75% likelihood) and the unpredicted condition consisted of trials, that appeared on the opposite side to the one indicated by the probabilistic cue (25%). The no-prediction condition included trials from those blocks in which the likelihood of the spatial location was equally distributed between visual fields (indicated with the ‘BOTH’ cue). Effects of task-relevance were assessed in the behavioural data by calculating the percentage of misses in the task-relevant condition (i.e., amount of trials in which stimuli appeared on the side indicated by the task-relevant cue but were not responded to) and the percentage of false alarms (FA) in the task-irrelevant condition (i.e., responses to task-irrelevant stimuli). The per cent misses and FAs were both subjected to one-way ANOVAs to assess the effects of prediction. Only trials with correct responses to grating orientation were included in the analysis of mean RTs. Only those participants whose data were retained for the EEG analysis were included in the behavioural analysis.

2.7. EEG preprocessing and data analysis

EEG pre-processing was carried out using EEGLAB (Delorme & Makeig, 2004). Data were re-referenced offline to the average of the left and right earlobes. Data were filtered with a 0.5 Hz high-pass and subsequently a 45 Hz low-pass windowed sinc finite impulse response (FIR) filter (Kaiser window, Kaiser $\beta = 5.653$, filter order 3624 and 184 for high-pass and low-pass filter, respectively). Vertical electrooculogram (VEOG) and horizontal electrooculogram (HEOG) data were calculated from the EOG data. Subsequently, cue-locked epochs of -200 – 800 ms and stimulus-locked epochs of -100 – 500 ms were defined. Epochs were baseline corrected using the 200-ms, and 100-ms time window before stimulus presentation for cue-locked, and target-locked epochs, respectively. Trials with eye movements were rejected based on the eye-tracking data (see above). Additionally, artefacts were rejected using the sorted averaging method (Rahne, von Specht, & Mühler, 2008). Specifically, trials were sorted according to their root mean square amplitude and only trials that increased signal-to-noise ratio of the individual average per condition were included in the analysis (overall 19.4% of trials were excluded). Trials with misses and FAs were rejected from the analysis; however, trials with incorrect grating orientation identification were kept in the analysis (VanRullen, 2011). Table 1 presents mean number of trials per condition that were used for computing ERP averages.

Table 1
Mean number of trials after artefact rejection used for averaging (SD in brackets).

Task-relevance	Prediction		
	Predicted	No-prediction	Unpredicted
Task-relevant			
Left VF	98 (11)	66 (7)	32 (5)
Right VF	96 (13)	64 (9)	32 (5)
Task-irrelevant			
Left VF	96 (12)	64 (10)	31 (5)
Right VF	96 (16)	65 (8)	33 (5)

2.8. ERP analysis

2.8.1. Cue-related activity

2.8.1.1. Congruity effects: N2 and P3. ERP average waveforms were computed separately for each condition, and for ipsilateral (i.e., left hemisphere electrodes in trials with left-pointing cues and right hemisphere electrodes in trials with right-pointing cues), and contralateral (i.e., left hemisphere electrodes in trials with right-pointing cues and right hemisphere electrodes in trials with left-pointing cues) electrodes. In order to investigate the effects of cue congruity, we contrasted the congruent condition, in which probabilistic and task-relevant cues indicated the same direction, with the incongruent condition, in which the task-relevant cue indicated a direction opposite to the direction of the probabilistic cue. The first maximum peak of difference was identified between 250 and 300 ms at fronto-central channels (see Fig. 2). The mean amplitude in a cluster of 6 fronto-central electrodes ('F1/2i', 'Fz', 'F1/2c', 'FC1/2i', 'FCz', and 'FC1/2c') was subjected to an ANOVA with the within-subject factor of congruity (congruent, neutral, incongruent) and electrode location (ipsilateral: 'F1/2i' and 'FC1/2i', midline: 'Fz' and 'FCz', contralateral: 'F1/2c' and 'FC1/2c'). To analyse the congruity effect in detail, we conducted *t*-tests comparing the congruity conditions.

Further, incongruent cues elicited a central-posterior P3 response between 300 and 450 ms (see Fig. 2). Therefore, mean amplitudes in a cluster of 6 centro-parietal electrodes ('CP1/2i', 'CPz', 'CP1/2c', 'P1/2i', 'Pz', and 'P1/2c') in the time window of 316–436 ms were analysed by means of a 3×3 ANOVA with congruity (congruent, neutral, incongruent) and electrode location (ipsilateral: 'CP1/2i' and 'P1/2i', midline: 'CPz' and 'Pz', contralateral: 'CP1/2c' and 'P1/2c') as factors.

2.8.1.2. Congruity effects: EDAN. We also analysed the EDAN by comparing contralateral vs. ipsilateral activity at the lateral parieto-occipital electrodes PO7/8 in the time-window of 200–280 ms. The effect of congruity on this component was analysed by means of a 3×2 ANOVA with the within-subject factors congruity (congruent, neutral, incongruent) and lateralisation (contralateral 'PO7/8c', ipsilateral 'PO7/8i').

2.8.2. Target-related activity

2.8.2.1. Lateral parieto-occipital cluster: attention effects. We focused our analysis on the visual sensory responses at lateral parieto-occipital sites, contralateral and ipsilateral to the stimulus presentation. The signal from contralateral (i.e., left hemisphere electrodes in response to RVF stimuli and right hemisphere electrodes in response to LVF stimuli) and ipsilateral (i.e., right hemisphere electrodes in response to RVF stimuli and left hemisphere electrodes in response to LVF stimuli) electrodes was averaged for each condition. The P1 and N1 components were assessed by the averaged signal from a cluster of five ipsilateral ('O1/2i', 'PO3/4i', 'PO7/8i', 'P5/6i', and 'P7/8i') and five contralateral ('O1/2c', 'PO3/4c', 'PO7/8c', 'P5/6c', and 'P7/8c') temporal parieto-occipital electrodes (see a depiction of clusters of tested electrodes

Fig. 2. Effects of congruity between prediction and task-relevance. A) ERPs in response to task-relevant cues for three cue congruity conditions (congruent, no-prediction, incongruent) at 7 electrode positions. Grey panels mark the time windows tested (250–300 ms and 316–436 ms after the presentation of task-relevant cue).

B) Scalp topographies in three experimental conditions (congruent, no-prediction, incongruent) and topography of the difference between the incongruent and congruent conditions in the time window of 250–300 ms. The ipsilateral condition is visualised in the left hemisphere, and the contralateral condition is visualised in the right hemisphere.

Task- relevance effects

Fig. 3. Task-relevance effects. A) ERPs in response to the task relevant and the task-irrelevant stimuli displayed in three prediction conditions (predicted, no-prediction, and unpredicted). Grey panels mark the time windows, in which significant effects of task-relevance were observed, namely the ipsilateral P1 (116–156 ms), the contralateral N1/Nd1 (150–200 ms), and the No-Go P3 (244–292 ms) respectively.

B) Scalp topographies of task-relevance effects: the ipsilateral P1, Nd1, and No-Go P3; and topographies of difference waves between task-relevant vs. task-irrelevant conditions. The ipsilateral condition is visualised in the left hemisphere, and the contralateral condition is visualised in the right hemisphere.

in Fig. 3). The time-window of analysis for the P1 component was set around the maximum positive deflection identified on the grand average in the window of 50–160 ms (84–116 ms and 116–156 ms for the contralateral and the ipsilateral P1 respectively). The time-window for the analysis of visual N1 component was set around the maximum negative deflection identified in the grand average (150–200 ms). We compared the task-relevant and task-irrelevant conditions to identify the attention effects. We conducted $2 \times 3 \times 2$ ANOVA with task-relevance (task-relevant, task-irrelevant), prediction (predicted, no prediction, unpredicted), and lateralisation (ipsilateral parieto-occipital cluster, contralateral parieto-occipital cluster) as within-subject factors.

2.8.2.2. Lateral parieto-occipital cluster: prediction effects. Prediction effects were identified by comparing responses to unpredicted stimuli with responses to predicted stimuli. We analysed the prediction effects in the time window of the contralateral N1 component (150–200 ms). Additionally, we conducted an analysis centered on the first peak observed in the difference wave between unpredicted and predicted conditions (174–220 ms). A $2 \times 3 \times 2$ ANOVA with task-relevance (task-relevant, task-irrelevant), prediction (predicted, no prediction, unpredicted), and lateralisation (ipsilateral parieto-occipital cluster, contralateral parieto-occipital cluster) as within-subject factors was conducted, to identify the prediction effect, and its possible interaction with the task-relevance factor (See Fig. 4 for a depiction of clusters of the tested electrodes). To analyse the prediction effect in detail, we ran *t*-tests comparing the prediction conditions.

2.8.2.3. Central-parietal and parieto-occipital cluster: Nd1 attention effect. The ERP waveforms manifested effects related to the task-relevance manipulation characterised by a central-parietal scalp distribution, typical for the Nd1 effect. The Nd1 effect was assessed at four central-parietal and parieto-occipital channels: 'Cz', 'CPz', 'Pz', 'POz' (see Fig. 3 for a depiction of electrode cluster tested). The time-window for the analysis of the task-relevance modulation was defined around the first maximum deflection of the task-relevant vs. task-irrelevant difference wave (150–200 ms). The data were subjected to an ANOVA with task-relevance (task-relevant, task-irrelevant), prediction (predicted, no prediction, unpredicted), and anterior-posterior (central-parietal: 'Cz' and 'CPz', and parieto-occipital: 'Pz' and 'POz') as within-subject factors.

2.8.2.4. Central-parietal cluster: prediction effect. Effects of probabilistic manipulations on the ERP waveforms were not circumscribed only to lateral parieto-occipital sites, typical for visual sensory ERPs, rather, they were characterised by a broader, central-parietal scalp distribution. These effects were assessed at a cluster of central-parietal electrodes (see Fig. 4 for depiction of clusters of electrodes tested). The time-window for the analysis of these prediction effects was centered on the first maximum deflection of the unpredicted vs. predicted difference wave (174–220 ms). The data were subjected to an ANOVA with task-relevance (task-relevant, task-irrelevant), prediction (predicted, no prediction, unpredicted), and electrode location (ipsilateral: 'C1/2i' and 'CP1/2i', midline: 'CZ' and 'CPz', contralateral: 'C1/2c' and 'CP1/2c') as within-subject factors. Post-hoc *t*-tests were performed to compare the prediction conditions.

2.8.2.5. Fronto-central P3 responses. We expected to observe at least two types of positive deflections in response to visual stimuli in the range of the P3 (approx. 200–400 ms); a frontal P3 elicited by task-irrelevant gratings (no-go trials), i.e. No-Go P3 (see e.g., Eimer, 1993; Salisbury, Griggs, Shenton, & McCarley, 2004), and a P3 response to unpredicted stimuli that has been associated with violations of predictions (novelty P3 or P3a; for reviews see Friedman

Table 2
Behavioural results: mean values (and SD in brackets).

Task-relevance	Prediction		
	Predicted	No-prediction	Unpredicted
Task-relevant			
Misses (%)	1.30 (1.74)	1.25 (1.68)	1.91 (2.26)
Incorrect grating identification (%)	4.24 (5.64)	4.19 (5.22)	5.22 (5.10)
Response times (ms)	675 (78)	689 (81)	719 (82)
Task-irrelevant			
False alarms (%)	1.52 (1.44)	1.62 (1.67)	1.40 (2.46)

et al., 2001; Polich, 2007). The window for the analysis of No-Go P3 was set to 48 ms around the maximum difference between task-irrelevant vs. task-relevant stimuli at the fronto-central midline channels (244–292 ms). The window for the analysis of the novelty P3 was identified around the maximum positive deflection of the difference wave between the unpredicted and predicted condition at the fronto-central midline electrodes, which was also encompassed by the time window of 244–292 ms. We conducted analyses of the P3 amplitudes in this time window for the midline electrodes 'Fz', 'FCz' and 'Cz' (see Fig. 3). An ANOVA with within-subject factors of task-relevance (task-relevant, task-irrelevant), prediction (predicted, no prediction, unpredicted) and anterior-posterior ('Fz', 'FCz', 'Cz') was performed (see Figs. 3 and 4 for a depiction of the tested electrode-cluster). A positive deflection in response to task-relevant stimuli was also observed at fronto-central midline channels in a later time window (318–368 ms). An ANOVA with within-subject factors of task-relevance (task-relevant, task-irrelevant), prediction (predicted, no prediction, unpredicted), and anterior-posterior ('Fz', 'FCz', 'Cz') was also performed for the P3 amplitude in this time window.

In all analyses, the Greenhouse-Geisser correction was applied if Mauchly's test detected a violation of sphericity. We report the corrected probability values. Generalized eta-squared (η^2_G) and Pearson *r* are reported as measures of effect size respectively for ANOVAs and *t*-tests. Statistical analyses were performed using the statistical software R (R Core Team, 2015) and the package "ez" (Lawrence, 2013).

3. Results

3.1. Behavioural results

Mean RT was 694 ms (SD = 81 ms). The mean percentage of FA was 1.53% and the mean percentage of misses was 1.38% suggesting that participants were able to follow the task instructions accurately (see Table 2 for mean values for each condition). Neither the percentage of FA, nor misses were significantly modulated by the probabilistic manipulation, $F < 1$ and $F(2,32) = 2.68$, $p = 0.084$, $\eta^2_G = 0.026$, respectively. As indicated by the main effect of prediction on RT; $F(2,32) = 50.4$, $p < 0.001$, $\eta^2_G = 0.053$, RTs were significantly faster for predicted than for either the unpredicted, $M = -44$ ms; $t(16) = -8.61$, $p < 0.001$, $r = 0.91$ or the no-prediction stimuli, $M = -14$ ms; $t(16) = -4.29$, $p < 0.001$, $r = 0.73$. RTs in the no-prediction condition were faster than in the unpredicted condition, $M = -30$ ms; $t(16) = -6.20$, $p < 0.001$, $r = 0.84$. These results indicate that the manipulation of prediction was effective at the behavioural level. The accuracy of grating identification was not modulated by prediction, $F < 1$.

3.2. EEG results

3.2.1. Cue-related activity

3.2.1.1. Congruity effects: N2 and P3. The difference wave subtracting the congruent condition (task-relevance and probabilistic cues

Prediction effects

A Task - relevant condition

Task - irrelevant condition

B Interaction between prediction and task-relevance (174–220 ms)

Novelty P3 (244–292 msec)

Fig. 4. Prediction effects. A) ERPs for three prediction conditions (predicted, no-prediction, unpredicted), displayed separately in the task-relevant and the task-irrelevant condition. Grey panels mark the times window tested: the interaction between attention and prediction (174–220 ms), P3a (244–292 ms), and P3b (318–368 ms). B) Scalp topographies for each 6 conditions of the task showing an interaction between attention and prediction (174–220 ms); and a scalp topography of the novelty P3 (P3a; 244–292 ms) effect. For each effect, cluster of electrodes tested is indicated in the scalp figure with black circles. The ipsilateral condition is visualised in the left hemisphere, and the contralateral condition is visualised in the right hemisphere.

indicate the same direction) from the incongruent condition (i.e., task-relevance and probabilistic cues indicate opposite directions) indicated a congruity effect contralateral to the direction of the task-relevant cue, peaking between 250 and 300 ms (see Fig. 2).¹ In the cluster of fronto-central electrodes analysed, a significant interaction between congruity and electrode location was observed, $F(4,64)=3.54$, $p=0.011$, $\eta^2_G=0.004$. At contralateral electrodes, incongruent trials elicited a larger negativity relative to congruent trials, $M=-0.76 \mu V$, $t(16)=-3.67$, $p=0.002$, $r=0.68$, and relative to neutral trials, $M=-0.49 \mu V$, $t(16)=-2.10$, $p=.052$, $r=0.46$, while the congruent vs. neutral condition did not differ significantly, $M=0.26 \mu V$, $t(16)=1.50$, $p=0.154$. At ipsilateral electrode-sites, a similar pattern was observed with a larger negativity in the incongruent vs. congruent condition, $M=-0.62 \mu V$, $t(16)=-3.31$, $p=0.004$, $r=0.64$, and no significant differences between incongruent vs. neutral, and congruent vs. neutral conditions, $t(16)=-1.39$ and $t(16)=1.77$ respectively, $ps>0.095$. In the incongruent condition only, the amplitudes were more negative at contralateral vs. ipsilateral sites, $M=-0.16 \mu V$, $t(16)=-4.57$, $p<0.001$, $r=0.75$, while the contralateral vs. ipsilateral difference was not significant in the congruent or neutral conditions, $t(16)=-0.61$ and $t(16)=0.11$, $ps>0.548$.

In a later time window, incongruent stimuli also elicited a centro-parietal P3. The congruity $\times$ electrode location ANOVA showed a significant main effect of congruity, $F(2,32)=10.89$, $p<.001$, $\eta^2_G=0.027$, with a larger P3 in the incongruent vs. congruent condition, $M=0.54 \mu V$, $t(16)=2.74$, $p=0.014$, $r=0.57$, and incongruent vs. neutral condition, $M=0.76 \mu V$, $t(16)=5.19$, $p<0.001$, $r=0.79$. No significant difference was observed between the congruent and neutral conditions, $M=0.22 \mu V$, $t(16)=1.40$, $p=0.179$. The interaction between congruity and lateralisation was not significant, $F(4,64)=1.13$, $p=0.350$, $\eta^2_G<0.001$.

3.2.1.2. EDAN. The congruity $\times$ lateralisation ANOVA in the time-window of the EDAN component (200–280 ms) showed a significant interaction, $F(2,32)=13.34$, $p<0.001$, $\eta^2_G=0.004$. In the incongruent condition, the mean amplitude was more negative at the contralateral than at the ipsilateral electrodes, $M=-0.23 \mu V$, $t(16)=-2.32$, $p=0.003$, $r=0.50$. However, in the congruent condition, the mean amplitude was more positive at the contralateral than at the ipsilateral electrodes, $M=0.36 \mu V$, $t(16)=2.67$, $p=0.017$, $r=0.55$, indicating the opposite pattern, or a reversed EDAN (EDANr, see Velzen & Eimer, 2003). In the neutral condition, the contralateral vs. ipsilateral difference was not significant, $M=0.02 \mu V$, $t(16)=0.16$, $p=0.876$.

3.2.2. Target-related activity

3.2.2.1. Lateral parieto-occipital cluster: attention effects. In the time-window quantifying the contralateral P1 (84–116 ms), task-relevance $\times$ prediction $\times$ lateralisation ANOVA showed a main effect of lateralisation, $F(1,16)=30.08$, $p<0.001$, $\eta^2_G=0.160$, with a more positive P1 amplitude for contralateral ($M=0.13 \mu V$) than for ipsilateral channels ($-1.37 \mu V$). Neither the main effects of task-relevance and prediction, nor any interactions reached significance, all $Fs<3.73$, all $ps>0.071$.

In the time-window of the ipsilateral P1 (116–156 ms), a significant main effect of lateralisation was observed, $F(1,16)=31.58$, $p<0.001$, $\eta^2_G=0.218$, with more positive amplitudes for ipsilat-

eral ($M=0.78 \mu V$) than for contralateral ($M=-1.95 \mu V$) channels. Furthermore, a significant interaction between task relevance and lateralisation was observed, $F(1,16)=9.02$, $p=0.008$, $\eta^2_G=0.002$ (see Fig. 3). In the ipsilateral cluster of electrodes, the mean amplitude was significantly more positive in the task-relevant ($M=1.12 \mu V$) than in the task-irrelevant condition ($M=0.45 \mu V$), $F(1,16)=5.66$, $p=0.030$, $\eta^2_G=0.025$, while the effect was not significant in the contralateral cluster, $F<0.74$. No other main effects or interactions reached significance, all $Fs<3.0$, $ps>0.107$.

In the time window of the contralateral N1 (150–200 ms), the task-relevance $\times$ prediction $\times$ lateralisation ANOVA on parieto-occipital electrodes showed a significant main effect of task-relevance, $F(1,16)=8.25$, $p=0.011$, $\eta^2_G=0.042$. The mean amplitude was significantly more negative for task-relevant ($M=-2.56 \mu V$) vs. task-irrelevant ($M=-1.46 \mu V$) stimuli. This effect did not reliably interact with the lateralisation factor, $F<0.01$, $p=0.959$.

3.2.2.2. Lateral parieto-occipital cluster: prediction effects. In the time window of the contralateral N1 (150–200 ms), a significant interaction between prediction and lateralisation was also observed on the lateral parieto-occipital electrodes, $F(2,32)=7.05$, $p=0.003$, $\eta^2_G=0.001$. A significant effect of prediction in the contralateral electrode cluster, $F(2,32)=4.70$, $p=0.016$, $\eta^2_G=0.007$, showed an enhanced negativity in the unpredicted condition as compared to the predicted condition, $M=-0.54 \mu V$, $t(16)=-2.70$, $p=0.016$, $r=0.56$ and as compared to the no-prediction condition, $M=-0.43 \mu V$, $t(16)=-2.30$, $p=0.035$, $r=0.50$ (predicted vs. no-prediction condition, $M=0.11 \mu V$, $t=0.64$, $p=0.532$). The prediction effect was not significant in the ipsilateral cluster, $F=1.34$, $p=0.275$.

In the later time window, directly following the peak of the contralateral N1 component (174–220 ms), the first peak of the unpredicted vs. predicted difference wave was observed (see Fig. 4). The task-relevance $\times$ prediction $\times$ lateralisation ANOVA conducted for this time window also showed a significant interaction between prediction and lateralisation, $F(2,32)=6.26$, $p=0.005$, $\eta^2_G=0.001$. In the contralateral cluster, the main effect of prediction, $F(2,32)=6.77$, $p=0.003$, $\eta^2_G=0.010$ revealed an enhanced negativity for unpredicted as compared to predicted stimuli, $M=-0.64 \mu V$, $t(16)=-3.29$, $p=0.005$, $r=0.64$, or stimuli from the no-prediction blocks, $M=-0.44 \mu V$, $t(16)=-2.53$, $p=0.022$, $r=0.53$. Predicted and no-prediction conditions did not differ significantly, $M=0.19 \mu V$, $t=1.20$, $p=0.249$. The prediction effect was not significant in the ipsilateral cluster, $F=1.17$, $p=0.324$. The effect of task-relevance was not significant, nor was the interaction between task-relevance and prediction, all $Fs<1.90$, all $ps>0.166$.

3.2.2.3. Central-parietal and parietal-occipital cluster: Nd1 attention effect. The difference wave depicting the N1 effect of task-relevant and task-irrelevant conditions revealed an enhanced negativity for task-relevant gratings over the midline centro-parietal and parieto-occipital scalp in the time window of 150–200 ms (see Fig. 3). The task-relevance $\times$ prediction $\times$ anterior-posterior ANOVA showed a significant interaction between the task-relevance and anterior-posterior factors, $F(1,16)=10.67$, $p=0.005$, $\eta^2_G=0.004$. Post-hoc t -tests comparing task-relevant vs. task-irrelevant condition showed that a significant task-relevance effect was present both at the centro-parietal, $M=-2.90 \mu V$, $t(16)=-5.17$, $p<0.001$, $r=0.79$, and parieto-occipital electrode clusters, $M=-2.12 \mu V$, $t(16)=-3.93$, $p=0.001$, $r=0.70$. The effect of prediction and the interaction between task-relevance and prediction did not reach significance, $Fs<1.63$, $ps>0.060$.

3.2.2.4. Central-parietal cluster: prediction effect. The analysis of the prediction effect was centered around the first maximum difference between unpredicted and predicted conditions (174–220 ms),

¹ We observed a lateralisation effect in the time-window of the N1; namely, the ipsilateral N1 was larger than the contralateral N1 (See Fig. 2; for a report of ipsilateral N1 attention effects see also Jones and Forster, 2014). However, no significant congruity effects were observed. This effect could reflect an early latency "recursive" effect of attention, whereby cue-related activity is modulated by attention (see e.g. Talsma, Sikkens, & Theeuwes, 2011).

which was characterised by a broader central negativity (see Fig. 4). In this time window, the task-relevance $\times$ prediction $\times$ electrode location ANOVA revealed a significant effect of task-relevance, $F(1,16) = 12.41$, $p = 0.003$, $\eta^2_G = 0.071$, showing an enhanced negativity for task-relevant stimuli. Importantly, a significant interaction between task relevance and prediction was observed, $F(2,32) = 3.60$, $p = 0.039$, $\eta^2_G = 0.003$. This interaction seemed driven by a negativity elicited by unpredicted stimuli at centro-parietal channels that was apparent only in the task-irrelevant condition (see Fig. 4). In the task-irrelevant condition, unpredicted stimuli were characterised by a more negative amplitude than predicted stimuli, $M = -1.04 \mu V$, $t(16) = -2.39$, $p = 0.029$, $r = 0.51$, while no significant differences were observed between the unpredicted vs. no-prediction conditions and the predicted vs. no-prediction conditions, $t(16) = -1.78$ and $t(16) = 1.62$ respectively, $ps > 0.094$. No significant differences between prediction conditions were observed in the task-relevant condition, all $ts < -1.23$, all $ps > 0.24$. The effect of task-relevance (task-relevant vs. task-irrelevant condition) was significantly larger in the predicted condition relative to unpredicted condition, $M = -0.98 \mu V$, $t(16) = -2.47$, $p = .025$, $r = .52$, while there was no significant difference in the magnitude of the task-relevance effect in the predicted vs. no-prediction and no-prediction vs. unpredicted conditions, $t(16) = -1.64$ and $t(16) = -0.94$ respectively, $ps > 0.121$.

3.2.2.5. Fronto-central P3 responses. Task-irrelevant gratings elicited a late fronto-central positivity, which was identified as a No-Go P3 (see Fig. 3). The task-relevance $\times$ prediction $\times$ anterior-posterior ANOVA confirmed a significant main effect of task-relevance, $F(1,16) = 8.73$, $p = 0.009$, $\eta^2_G = 0.123$, in the time window corresponding to the maximum difference between task-relevant vs. task-irrelevant stimuli (244–292 ms). The task-relevance effect interacted with the anterior-posterior factor, $F(2,32) = 8.67$, $p = 0.003$, $\eta^2_G = 0.005$; therefore we compared the task-irrelevant vs. task relevant condition at each electrode. The No-Go P3 effect was the largest at the fronto-central 'FCz' electrode, $M = 3.71 \mu V$, $t(16) = 3.28$, $p = 0.005$, $r = 0.63$, but was also significant at 'Cz', $M = 3.07 \mu V$, $t(16) = 2.82$, $p = 0.012$, $r = 0.54$ and 'Fz', $M = 2.29 \mu V$, $t(16) = 2.56$, $p = 0.021$, $r = 0.58$.

In the time window corresponding to No-Go P3 (244–292 ms), a maximum positive difference between unpredicted and predicted stimuli was observed (see Fig. 4). A statistically significant novelty P3 effect was confirmed by a main effect of prediction, $F(2,32) = 5.70$, $p = 0.008$, $\eta^2_G = 0.007$. Unpredicted gratings elicited a larger P3 than predicted gratings, $M = 0.85 \mu V$, $t(16) = 2.93$, $p = 0.010$, $r = 0.59$, while predicted vs. no-prediction, and unpredicted vs. no-prediction conditions did not differ significantly, $t(16) = -2.01$ and $t(16) = 1.71$ respectively, $ps > 0.062$. The effect of prediction was not significantly modulated by the task-relevance or anterior-posterior factor, $F_s < 2.89$, $ps > 0.061$.

In a later time window (318–368 ms), a significant effect of task-relevance was also observed, $F(1,16) = 7.70$, $p = 0.013$, $\eta^2_G = 0.109$, with a larger P3 for task irrelevant ($M = 7.41 \mu V$) relative to task-relevant ($M = 4.93 \mu V$) conditions. The main effect of prediction was also significant in this time window, $F(2,32) = 4.68$, $p = 0.016$, $\eta^2_G = 0.006$ with a larger amplitude for the unpredicted ($M = 6.48 \mu V$) than the predicted ($M = 5.81 \mu V$) condition. The main effects did not interact with each other or with the anterior-posterior factor, all $F_s < 2.74$, $ps > 0.111$.²

² In this time window (318–368 ms), task-relevant stimuli also seemed to have elicited a larger P3 at posterior channels (target P3; or P3b; Friedman et al., 2001). A main effect of task relevance was not significant in the parieto-occipital cluster ('Pz' and 'POz' electrode), $F(1,16) = 0.87$, $p = 0.36$, however a main effect of prediction

4. Discussion

In the current study we aimed to characterise the time course of a possible interplay between attention and prediction by measuring event-related potentials in response to visual stimuli. Attention was manipulated by trial-by-trial task-relevant cues, while prediction was induced by the use of block-wise probabilistic cuing. In line with the *attentional gain hypothesis*, we expected to observe an enhanced N1 and Nd1 amplitude in response to task-relevant stimuli. In line with the *attenuation-by-prediction hypothesis*, we assumed that the N1 would be attenuated in response to predicted vs. unpredicted stimuli, followed by an increased P3 response to unpredicted stimuli. Furthermore, based on the hypothesis that *attention and prediction may interface*, it was expected that the attenuation by prediction would be present for task-irrelevant stimuli only, whereas the prediction effect would be reversed for task-relevant stimuli, resulting in an enhanced N1 for predicted vs. unpredicted stimuli. We also hypothesised to observe effects of cue congruity in the pre-stimulus period, which would suggest that information provided by the two types of cues is integrated prior to the stimulus onset.

The results were in accord with the *attentional gain hypothesis*, as shown by an increased parieto-occipital N1 for task-relevant compared to task-irrelevant gratings, accompanied by a more centrally distributed Nd1 effect. The pattern of results also spoke for the *attenuation-by-prediction hypothesis*. When compared to predicted gratings, unpredicted gratings elicited enhanced contralateral parieto-occipital visual N1 responses, followed by a centrally distributed novelty P3 response. The results also confirmed that attention and prediction effects are interrelated, however, the interactive pattern was not as hypothesised. Specifically, in a time window directly following the peak of the N1, and displaying a central scalp distribution, we found attenuation-by-prediction effects that were circumscribed to task-irrelevant gratings. Simultaneously, we found larger attentional gain effects for predicted gratings compared to unpredicted gratings.

An interactive influence of predictive and task-relevant cues was also found in the cue-target interval. Responses to cues conveying incongruent task-relevance and probabilistic information were characterised by a combined N2 and EDAN, followed by a P3 response, while a reverse EDAN effect was present for cues conveying congruent task-relevance and probabilistic information.

All in all, the results demonstrate the presence of both independent and interdependent effects of attention and prediction at different stages along the time course of visual stimulus processing. All these effects are discussed in turn in the following sections.

4.1. Probabilistic and task-relevance information is integrated in the pre-stimulus period

We compared responses to task-relevant cues that were either congruent (i.e., directed attention to the same spatial location) or incongruent (i.e., directed attention to the opposite spatial location) with responses to block-wise probabilistic cues. Incongruent task-relevant cues elicited an increased negativity in the N2 range (250–300 ms) at contralateral fronto-central channels, but also at the lateral parieto-occipital channels, reflected in the EDAN component (Velzen & Eimer, 2003; van der Lubbe et al., 2006). The EDAN was present in the incongruent condition, while a reversed EDAN was observed in the congruent condition (EDANr; Velzen & Eimer, 2003). This pattern of results suggests that EDAN may be sensitive to a predictive context provided by the cues (see also

was significant, $F(2,32) = 4.78$, $p = 0.015$, $\eta^2_G = 0.005$, with a larger P3b amplitude for unpredicted stimuli at these electrodes.

Lasaponara, Chica, Lecce, Lupiáñez, & Doricchi, 2011). The contralateral N2 negativity in response to cues possibly reflects the orienting of attention to cued locations, and the deployment of attention to task-relevant events signalled by the cue (Talsma et al., 2007; Velzen & Eimer, 2003; van der Lubbe et al., 2006). Interestingly, the response to incongruent cues was characterised by an enhanced N2, which may reflect the re-orienting from a spatial location indicated by a block-wise probabilistic cue when the opposite direction is signalled by the task-relevant cue. The enhanced N2 in response to incongruent input may thus be interpreted as signalling the need to adapt to new information provided by the task-relevant cue.

The fronto-central N2 complex often shows when conflict arises during performance monitoring (for review see Folstein & Van Petten, 2008). This N2 modulation was followed by a centro-parietal P3. The P3 response has been associated with context-updating (Polich, 2007) or updating of probabilistic representations (Chennu et al., 2013; Feldman & Friston, 2010). Thus, in the pre-stimulus period, incongruent task-relevant cues may have triggered error-related cognitive control operations as evident in the N2 that may have lead to the adjustment of an internal model displayed in the centro-parietal P3 (cf. Kotz, Stockert, & Schwartz, 2014; SanMiguel, Saupe, & Schröger, 2013). These cue congruity effects suggest that task-relevance information provided by the cue may have been integrated with the pre-existing predictive information on every trial.

4.2. Attentional gain hypothesis: attention increases sensory gain to target stimuli

An attentional gain modulation was found in an enhanced parieto-occipital contralateral N1 (150–190 ms). The negative difference in this time window was most pronounced in the centro-parietal cluster, bilaterally distributed around the midline electrodes. This effect likely reflects an Nd1 difference, which is linked to transient modulations within the posterior parietal cortex, and seems to reflect a location expectation gradient within the dorsal ‘where’ pathway (Eimer, 1998; Tata & Ward, 2005). This pattern of results corresponds to previous findings demonstrating attentional gain (for reviews see e.g., Eimer, 1998; Hillyard et al., 1998; Luck et al., 2000), and is consistent with a predictive coding model, in which attention optimises perception by increasing gain of the target stimuli (Feldman & Friston, 2010).

Moreover, we observed a modulation of the P1 component, namely the late, ipsilateral P1 (P130; Clark & Hillyard, 1996). The functional significance of the P1 is much debated as some evidence suggests it reflects attentional gain (see e.g., Hillyard et al., 1998; Mishra, Martinez, Schroeder, & Hillyard, 2012), while other results suggest that P1 reflects top-down inhibition of task-relevant and task-irrelevant structures, functionally related to pre-stimulus alpha oscillations (Klimesch 2011, 2012). Interestingly, contrary to previous studies (see e.g. van der Lubbe & Woestenburg, 2000), a modulation of the contralateral P1, thought to reflect an increase in the signal-to-noise ratio in task-relevant structures (Klimesch 2011, 2012), was not observed in the current study. In most studies, a P1 attention effect occurs bilaterally (for a review see Klimesch, 2011) and a contralateral P1 enhancement for attended vs. unattended stimuli has been observed in detection and discrimination tasks (e.g., Mangun & Hillyard, 1991). However, as the current results show, the ipsilateral P1 seems to be larger, more pronounced, and delayed in trial-by-trial cueing tasks as compared to sustained attention tasks (cf., Freunberger et al., 2008; Klimesch, 2011). Notably, however, a recent study (Slagter, Prinssen, Reteig, & Mazaheri, 2016) reported an ipsilateral P1 attention effect in a sustained attention task, in which attention was directed exclusively to one visual field throughout the experiment. This effect

was observed in the absence of a contralateral P1 effect and in the absence of a modulation of pre-stimulus alpha activity. The authors concluded that the ipsilateral P1 attention effect may be functionally dissociable from alpha modulations and may reflect a local process of inter-hemispheric inhibition (Slagter et al., 2016).

To conclude, attentional gain was reflected in enhanced N1/Nd1 components, while the ipsilateral P1 attention effect may reflect the inhibition of task-irrelevant structures prior to gain modulation in the N1 time-window.

4.3. Attenuation-by-prediction hypothesis: attenuation of neural responses to predicted stimuli followed by an increased processing of unpredicted stimuli

In line with the attenuation-by-prediction hypothesis, a modulation by prediction was observed in the time window of the contralateral N1 component, with the difference between unpredicted and predicted stimuli peaking in the time window directly following the N1 peak (174–220 ms). Responses to unpredicted stimuli were characterised by a larger negativity than responses to predicted stimuli. In contrast to the task-relevance effect, which was distributed bilaterally, the prediction effect was evident over contralateral channels only.

The observed prediction effect may be considered as analogous to the visual MMN (vMMN) that has been interpreted in line with predictive coding theories as reflecting sensory context-based predictions (Kimura, 2012), and more specifically, as a prediction error signal to input inconsistent with a probabilistic model (Stefanics et al., 2014). More specifically, however, this prediction effect may reflect an adaptation of the N1 in response to more probable and more frequently occurring stimuli (for review see Grill-Spector et al., 2006), which is known to contribute to the MMN response (for reviews see Garrido et al., 2009; May & Tiitinen, 2010; Stefanics et al., 2014). This attenuation has sometimes been claimed to reflect neural refractoriness or “fatigue”, “...a basic physiological phenomenon, without any cognitive functional significance” (Stefanics et al., 2014, p. 3). However, contrary to this view (see e.g. O’Shea, 2015), the N1 adaptation due to stimulus frequency may also be understood as an instance of active perceptual inference. According to the hierarchical predictive coding framework, this adaptation would reflect a reduction of the prediction error signal as a consequence of low-level predictions derived from local transition probabilities (Aukstulewicz & Friston, 2016). The N1 adaptation may also be related to the N1-suppression effect observed for predicted auditory and visual stimuli (for reviews see Horváth, 2015; Hughes et al., 2012; Schröger et al., 2015). The N1-suppression seems to result from a match between the predicted and sensory input and the prediction error is reduced when the prediction and the sensory input match. Although in these auditory studies predictions are usually not manipulated probabilistically, and overall stimulus frequency is controlled for, it is worth noting that – similar to the present visual contralateral N1 suppression effect – the auditory N1-suppression effect appears to occur independent of attention (Saupe, Widmann, Trujillo-Barreto, & Schröger, 2013; Timm, SanMiguel, Saupe, & Schröger, 2013).

As hypothesised, an early prediction effect was followed by modulations at later processing stages. Namely, we observed enhanced P3 responses to unpredicted input. This observation is in line with previous findings that low probability and therefore surprising events trigger larger frontal P3 responses (P3a, Friedman et al., 2001; Verleger, Jaskowski, & Wauschkuhn, 1994). A prediction effect was also observed in a later time-window at both fronto-central and parietal electrode-sites suggesting a P3b modulation (see e.g., Polich, 2007). In the hierarchical predictive coding models, the P3b response is said to reflect a higher-level stage in the prediction error hierarchy (Wacongne et al., 2011) that is linked

to the conscious detection of an expectancy violation (Strauss et al., 2015). The P3 prediction effect in the current study may therefore reflect increased novelty processing that triggers the updating of the internal representations.

4.4. Interrelation of attention and prediction

Consistent with our hypothesis about the interactive influence of attention and prediction, an interaction between prediction and task-relevance was observed in the time window following the peak of the N1 over centro-parietal channels. However, the pattern of this interactive effect was unexpected: unpredicted stimuli elicited an enhanced negativity relative to predicted stimuli in the unattended condition. We assume that the deviance-related negativity in the task-irrelevant condition may index a prediction error that signals a divergence from the prediction, and the need to update the internal model. Responses to deviant stimuli, as reflected in the MMN component, are predominantly studied when attention is diverted away from a stimulus by a secondary task, and a large body of research suggests that predictive processing indexed by the MMN is to a large extent automatic and independent of attentional resources (for a review see Stefanics et al., 2014). However, the interrelatedness of predictive processes that underlie the vMMN and top-down attention, or task-relevance, is a matter of an ongoing debate (see e.g., Kimura & Takeda, 2013; Kimura, Widmann, & Schröger, 2010; Stefanics et al., 2014).

There are at least two plausible explanations for the selective enhancement of the negativity in the task-irrelevant and unpredicted condition. Looking at cue-related effects in the pre-stimulus period, unpredicted and unattended stimuli were preceded by the congruent cue condition, in which both probabilistic and task-relevant cues pointed to the visual field opposite to the one, in which the gratings eventually appeared. The unpredicted and task-irrelevant stimulus, which was preceded by this congruent cue condition was thus bound to elicit the largest surprise, which in turn may have increased processing costs. On the other hand, the unpredicted but attended stimuli were preceded by task-relevant cues that were incongruent with the prediction cue. In this condition, as reflected in the fronto-central N2 and centro-parietal P3 elicited by the incongruent cues, an update of the task-context, or the internal model, may have taken place in the pre-stimulus period. A second, alternative interpretation is that in the incongruent cue conditions, conflict-related processing may have caused cognitive distraction (see Section 4.1), which may have subsequently interfered with the processing of the following unpredicted/attended stimulus and/or predicted/unattended stimulus. This possibility should be further investigated.

Finally, differences in the processing of task-irrelevant stimuli seem to have resulted in a larger attentional gain (i.e., difference between processing of task-relevant vs. task-irrelevant stimuli) for predicted compared to unpredicted stimuli. This interactive effect shows that attentional gain may be modulated prediction, as proposed by the predictive coding account discussed below.

4.5. A predictive coding perspective on the interrelation of attention and prediction

The results of the present study partially contrast with the study by Kok, Rahnev et al. (2012) that originally employed the modified spatial cueing paradigm to investigate the interrelation between attention and prediction. Kok, Rahnev et al. (2012) found a suppression of BOLD response in V1 to predicted stimuli in the task-irrelevant condition, whereas an enhancement of BOLD response to predicted stimuli was observed in the task-relevant condition. In the current study, an interaction showed a significant prediction effect in the task-irrelevant condition, while no

differences in the task-relevant condition were observed. It is important to note that Kok et al.'s (2012) analysis of the interaction focused on the early sensory cortex. Our analyses focused on the parieto-occipital P1 and N1 components, whose neural generators have been localised in the lateral extrastriate areas (Clark, Fan, & Hillyard, 1995; Gomez Gonzalez, Clark, Fan, Luck, & Hillyard, 1994); and the Nd1 component with likely sources in the dorsal pathway (Tata & Ward, 2005). Therefore, the possibility that our measurement was not sensitive to detect a potential interaction in the striate cortex cannot be excluded. Rather, observation of the interaction at the centro-parietal midline channels suggests that the interactive effect may be related to processing in higher visual areas, presumably in the dorsal pathway. Therefore, under the predictive coding account (cf. Aukstulewicz & Friston, 2015, 2016), this prediction effect may reflect hierarchically higher predictions based on the structure of the stimulus sequence or stimulus likelihood. Notably, a body of research employing combined EEG and fMRI has suggested that attentional modulations in the primary visual areas result from delayed re-entrant feedback from higher cortical areas (Boehler, Schoenfeld, Heinze, & Hopf, 2008; Martínez et al., 1999, 2001; Noesselt et al., 2002). Given the slow time resolution of fMRI, it thus remains plausible that the interaction effect in the V1 may as well be an attribute of the late re-entrant processing. Furthermore, the cue congruency effects in the pre-stimulus period seem consistent with the pattern of fMRI data in the study by Kok, Rahnev et al. (2012), in which incongruent cues elicited a larger cue-related modulation of BOLD response than congruent cues. Trials in the unpredicted/unattended condition and trials in the predicted/attended condition may have a lower pre-stimulus baseline activity than trials in the predicted/unattended or unpredicted/attended conditions. Stimulus-related responses may thus interact with the preceding pre-stimulus baseline shifts, making it challenging to disentangle the two modulations in the fMRI data. It is worth noting that contrary to findings by Kok, Rahnev et al. (2012), three other studies employing different paradigms and methodologies reported interactions between prediction and attention partially consistent with our findings, showing that predictions selectively influenced responses to unattended, task-irrelevant stimuli (EEG study: Jones, Hughes, and Waszak, 2013; MEG study: Todorovic et al., 2015; fMRI study: St John-Saaltink, Utzerath, Kok, Lau, & de Lange, 2015).

In the prediction error minimisation framework, perceptual inference is concerned not only with the content, but also with the confidence in sensory data, and thus estimates precision (or variability) of prediction errors (for a discussion see e.g., Feldman & Friston, 2010; Hohwy, 2013; Schröger et al., 2015). The interrelatedness of attention and prediction may therefore be understood as the interplay of these two types of perceptual inferences, where the inference of the content of the sensory input encoded by prediction errors is modulated by their expected precision (Hohwy, 2013). If prediction errors are weighted by their precision, the gain should be larger for more predictable input, which is inferred to be more precise. The interaction pattern observed in the current study indeed showed a larger attentional gain for predicted than for unpredicted stimuli.

Nevertheless, it remains plausible that prediction can modulate even earlier stages of attentional gain than those observed in the current study. Such a synergistic pattern was reported in a study by Hsu, Hämäläinen, and Waszak (2014), in which the auditory N1 was enhanced only in the attended and predicted condition, suggesting an increased sensory gain for attended and predicted input at an early stage of auditory processing. Similarly, a recent study on ERP signatures of predictive processing in infants (Kouider et al., 2015) proposed that the modulation of neural responses by prediction unfolds in two stages: an early attentional gain for pre-

dicted input is followed by late surprise responses to unpredicted input. We hypothesise that the specific way in which the interactive influences of attention and prediction unfold may be dependent on the precision of the sensory prediction; namely, on the availability and precision of a perceptual template of the physical features of the predicted stimulus (for a related discussion see Paris, Kim, & Davis, 2016). In the current study, a stimulus could be more or less expected at a certain spatial location based on the probabilistic context. Nevertheless, the precise features (i.e., orientation) of the stimulus could not be predicted based on probabilistic cues, thus a low-level, precise perceptual template of the upcoming target could not be formed. We suggest that synergistic effects of attention and prediction at earlier processing stages (i.e., in unimodal sensory areas) may depend on how precisely a top-down prediction (perceptual template) can be established based on the information provided by the context (cues). This suggestion requires further studies.

Finally, several other important issues remain open. Future studies should investigate whether synergistic influences of attention and prediction are also observed when attention and prediction are manipulated by different, independent features, unlike in the current study, where both manipulations concerned a single feature (i.e., spatial location). Moreover, similarly to previous studies (cf. Kok, Jehee et al., 2012; Kok, Rahnev et al., 2012) the current study manipulated prediction via cue, or instruction; however, effects of a purely probabilistic manipulation should be probed.

4.6. Time-course of the interactive influence of attention and prediction

To conclude, the current findings show that attention and prediction differentially influence electrophysiological signatures of visual processing. Based on the modulations of the electrophysiological responses examined in the current study, we suggest that the time-course of the modulation of perceptual processing by attention and prediction is characterised by distinct stages. In the first stage, in line with the *attentional gain hypothesis*, the modulation N1 and Nd1 components demonstrated an early attentional increase in gain for stimuli appearing at the attended, task-relevant locations. This stage was followed by a modulatory effect of prediction, which was reflected in an attenuated N1 response. Conforming to the *attenuation-by-prediction hypothesis*, this response may be considered as a smaller prediction error to predicted compared to unpredicted stimuli. An interaction between attention and prediction observed at this stage points to the *interrelation of attention and prediction*. An increased centro-parietal negativity in response to unpredicted and task-irrelevant input may be a sign of an increased prediction error in response to unexpected and unattended stimulus; however, alternative explanations based on pre-stimulus effects need to be considered. The later stage is reflected in larger P3 responses to unpredicted input, which may reflect the updating of the internal generative model. We suggest that interactive influences of attention and prediction may unfold differently at each of the stages, depending on how precise is the predictive information provided by the context. Future research should delineate conditions, under which attention and prediction interact at each stage.

Acknowledgments

We would like to thank Dr. Johanna Steinberg for her help with data collection, Dr. Antonio Schettino for his useful suggestions on data analysis, and Samyogita Hardikar for proofreading the manuscript. This work was supported by the Max Planck Society (International Max Planck Research School on Neuroscience

of Communication: Function, Structure, and Plasticity; scholarship to A.M.) and by the Reinhart-Koselleck Grant from the German Research Foundation (DFG) to E.S.

References

- Auksztulewicz, R., & Friston, K. (2015). Attentional enhancement of auditory mismatch responses: A DCM/MEG study. *Cerebral Cortex*, <http://dx.doi.org/10.1093/cercor/bhu323>
- Auksztulewicz, R., & Friston, K. (2016). Repetition suppression and its contextual determinants in predictive coding. *Cortex*, <http://dx.doi.org/10.1016/j.cortex.2015.11.024>
- Bastos, A. M., Usrey, W. M., Adams, R. A., Mangun, G. R., Fries, P., & Friston, K. J. (2012). Canonical microcircuits for predictive coding. *Neuron*, 76(4), 695–711. <http://dx.doi.org/10.1016/j.neuron.2012.10.038>
- Boehler, C. N., Schoenfeld, M. A., Heinze, H. J., & Hopf, J. M. (2008). Rapid recurrent processing gates awareness in primary visual cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 105(25), 8742–8747. <http://dx.doi.org/10.1073/pnas.0801999105>
- Brainard, D. H. (1997). *The psychophysics toolbox*. *Spatial Vision*, 10(4), 433–436.
- Brown, H. R., & Friston, K. J. (2013). The functional anatomy of attention: A DCM study. *Frontiers in Human Neuroscience*, 7 <http://dx.doi.org/10.3389/fnhum.2013.00784>
- Carrasco, M. (2011). Visual attention: The past 25 years. *Vision Research*, 51(13), 1484–1525. <http://dx.doi.org/10.1016/j.visres.2011.04.012>
- Chatrjian, G. E., Lettich, E., & Nelson, P. L. (1985). Ten percent electrode system for topographic studies of spontaneous and evoked EEG activities. *American Journal of EEG Technology*, 25(2), 83–92.
- Chennu, S., Noreika, V., Gueorguiev, D., Blenkmann, A., Kochen, S., Ibáñez, A., & Bekinschtein, T. A. (2013). Expectation and attention in hierarchical auditory prediction. *The Journal of Neuroscience*, 33(27), 11194–11205. <http://dx.doi.org/10.1523/JNEUROSCI.11114-13.2013>
- Clark, V. P., & Hillyard, S. A. (1996). Spatial selective attention affects early extrastriate but not striate components of the visual evoked potential. *Journal of Cognitive Neuroscience*, 8(5), 387–402. <http://dx.doi.org/10.1162/jocn.1996.8.5.387>
- Clark, V. P., Fan, S., & Hillyard, S. A. (1995). Identification of early visual evoked potential generators by retinotopic and topographic analyses? *Human Brain Mapping*, 2(3), 170–187.
- Czigler, I., Balázs, L., & Winkler, I. (2002). Memory-based detection of task-irrelevant visual changes. *Psychophysiology*, 39(6), 869–873. <http://dx.doi.org/10.1017/S0048577202020218>
- Dale, C. L., Simpson, G. V., Foxe, J. J., Luks, T. L., & Worden, M. S. (2008). ERP correlates of anticipatory attention: Spatial and non-spatial specificity and relation to subsequent selective attention. *Experimental Brain Research*, 188(1), 45–62. <http://dx.doi.org/10.1007/s00221-008-1338-4>
- Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134(1), 9–21. <http://dx.doi.org/10.1016/j.jneumeth.2003.10.009>
- Doherty, J. R., Rao, A., Mesulam, M. M., & Nobre, A. C. (2005). Synergistic effect of combined temporal and spatial expectations on visual attention. *The Journal of Neuroscience*, 25(36), 8259–8266. <http://dx.doi.org/10.1523/JNEUROSCI.1821-05.2005>
- Eimer, M. (1993). Spatial cueing, sensory gating and selective response preparation: An ERP study on visuo-spatial orienting. *Electroencephalography and Clinical Neurophysiology*, 88(5), 408–420.
- Eimer, M. (1994). Sensory gating as a mechanism for visuospatial orienting: Electrophysiological evidence from trial-by-trial cuing experiments. *Perception & Psychophysics*, 55(6), 667–675.
- Eimer, M. (1997). Attentional selection and attentional gradients: An alternative method for studying transient visual-spatial attention. *Psychophysiology*, 34(3), 365–376.
- Eimer, M. (1998). Mechanisms of visuospatial attention: Evidence from event-related brain potentials. *Visual Cognition*, 5(1–2), 257–286.
- Feldman, H., & Friston, K. J. (2010). Attention, uncertainty, and free-energy. *Frontiers in Human Neuroscience*, 4, 215. <http://dx.doi.org/10.3389/fnhum.2010.00215>
- Folstein, J. R., & Van Petten, C. (2008). Influence of cognitive control and mismatch on the N2 component of the ERP: A review. *Psychophysiology*, 45, 152–170.
- Freunberger, R., Höller, Y., Griesmayr, B., Gruber, W., Sauseng, P., & Klimesch, W. (2008). Functional similarities between the P1 component and alpha oscillations. *European Journal of Neuroscience*, 27(9), 2330–2340. <http://dx.doi.org/10.1111/j.1460-9568.2008.06190.x>
- Friedman, D., Cycowicz, Y. M., & Gaeta, H. (2001). The novelty P3: An event-related brain potential (ERP) sign of the brain's evaluation of novelty. *Neuroscience and Biobehavioral Reviews*, 25(4), 355–373.
- Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 360(1456), 815–836. <http://dx.doi.org/10.1098/rstb.2005.1622>
- Friston, K. (2009). The free-energy principle: A rough guide to the brain? *Trends in Cognitive Sciences*, 13(7), 293–301. <http://dx.doi.org/10.1016/j.tics.2009.04.005>

- Garrido, M. I., Kilner, J. M., Stephan, K. E., & Friston, K. J. (2009). The mismatch negativity: A review of underlying mechanisms. *Clinical Neurophysiology*, 120(3), 453–463. <http://dx.doi.org/10.1016/j.clinph.2008.11.029>
- Gentsch, A., & Schütz-Bosbach, S. (2011). I did it: Unconscious expectation of sensory consequences modulates the experience of self-agency and its functional signature. *Journal of Cognitive Neuroscience*, 23(12), 3817–3828. <http://dx.doi.org/10.1162/jocn.a.0001>
- Gomez Gonzalez, C. M., Clark, V. P., Fan, S., Luck, S. J., & Hillyard, S. A. (1994). Sources of attention-sensitive visual event-related potentials. *Brain Topography*, 7(1), 41–51.
- Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: Neural models of stimulus-specific effects. *Trends in Cognitive Sciences*, 10(1), 14–23. <http://dx.doi.org/10.1016/j.tics.2005.11.00>
- Harter, M. R., Miller, S. L., Price, N. J., Lalonde, M. E., & Keyes, A. L. (1989). Neural processes involved in directing attention. *Journal of Cognitive Neuroscience*, 1(3), 223–237. <http://dx.doi.org/10.1162/jocn.1989.1.3.223>
- Herrmann, C. S., & Knight, R. T. (2001). Mechanisms of human attention: Event-related potentials and oscillations. *Neuroscience & Biobehavioral Reviews*, 25(6), 465–476.
- Hillyard, S. A., & Anllo-Vento, L. (1998). Event-related brain potentials in the study of visual selective attention. *Proceedings of the National Academy of Sciences of the United States of America*, 95(3), 781–787.
- Hillyard, S. A., Luck, S. J., & Mangun, G. R. (1994). The cuing of attention to visual field locations: Analysis with ERP recordings. *Cognitive Electrophysiology*, 1–25.
- Hillyard, S. A., Vogel, E. K., & Luck, S. J. (1998). Sensory gain control (amplification) as a mechanism of selective attention: Electrophysiological and neuroimaging evidence. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 353(1373), 1257–1270. <http://dx.doi.org/10.1098/rstb.1998.0281>
- Hohwy, J. (2013). *The predictive mind*. Oxford University Press.
- Horváth, J. (2015). Action-related auditory ERP attenuation: Paradigms and hypotheses. *Brain Research*, 1626, 54–65. <http://dx.doi.org/10.1016/j.brainres.2015.03.038>
- Horváth, J., Czigler, I., Jacobsen, T., Maess, B., Schröger, E., & Winkler, I. (2008). MMN or no MMN: No magnitude of deviance effect on the MMN amplitude. *Psychophysiology*, 45(1), 60–69. <http://dx.doi.org/10.1111/j.1469-8986.2007.00599>
- Hsu, Y. F., Hämäläinen, J. A., & Waszak, F. (2014). Both attention and prediction are necessary for adaptive neuronal tuning in sensory processing. *Frontiers in Human Neuroscience*, 8, 152. <http://dx.doi.org/10.3389/fnhum.2014.00152>
- Hughes, G., Desantis, A., & Waszak, F. (2012). Mechanisms of intentional binding and sensory attenuation: The role of temporal prediction, temporal control, identity prediction, and motor prediction. *Psychological Bulletin*, <http://dx.doi.org/10.1037/a0028566>
- Jones, A., & Forster, B. (2014). Neural correlates of endogenous attention, exogenous attention and inhibition of return in touch. *The European Journal of Neuroscience*, 40(2), 2389–2398. <http://dx.doi.org/10.1111/ejn.12583>
- Jones, A., Hughes, G., & Waszak, F. (2013). The interaction between attention and motor prediction. An ERP study. *Neuroimage*, 83, 533–541. <http://dx.doi.org/10.1016/j.neuroimage.2013.07.004>
- Kaernbach, C. (1991). Simple adaptive testing with the weighted up-down method. *Perception & Psychophysics*, 49(3), 227–229.
- Kennett, S., van Velzen, J., Eimer, M., & Driver, J. (2007). Disentangling gaze shifts from preparatory ERP effects during spatial attention. *Psychophysiology*, 44(1), 69–78. <http://dx.doi.org/10.1111/j.1469-8986.2006.00470.x>
- Kimura, M., & Takeda, Y. (2013). Task difficulty affects the predictive process indexed by visual mismatch negativity. *Frontiers in Human Neuroscience*, 7, 267. <http://dx.doi.org/10.3389/fnhum.2013.00267>
- Kimura, M. (2012). Visual mismatch negativity and unintentional temporal-context-based prediction in vision. *International Journal of Psychophysiology*, 83(2), 144–155. <http://dx.doi.org/10.1016/j.ijpsycho.2011.11.010>
- Kimura, M., Widmann, A., & Schröger, E. (2010). Human visual system automatically represents large-scale sequential regularities. *Brain Research*, 1317, 165–179. <http://dx.doi.org/10.1016/j.brainres.2009.12.076>
- Klimesch, W. (2011). Evoked alpha and early access to the knowledge system: The P1 inhibition timing hypothesis. *Brain Research*, 1408, 52–71. <http://dx.doi.org/10.1016/j.brainres.2011.06.003>
- Klimesch, W. (2012). A-band oscillations, attention, and controlled access to stored information. *Trends in Cognitive Sciences*, 16(12), 606–617. <http://dx.doi.org/10.1016/j.tics.2012.10.007>
- Kok, P., Jehee, J. F., & de Lange, F. P. (2012). Less is more: Expectation sharpens representations in the primary visual cortex. *Neuron*, 75(2), 265–270. <http://dx.doi.org/10.1016/j.neuron.2012.04.034>
- Kok, P., Rahnev, D., Jehee, J. F., Lau, H. C., & de Lange, F. P. (2012). Attention reverses the effect of prediction in silencing sensory signals. *Cerebral Cortex*, 22(9), 2197–2206. <http://dx.doi.org/10.1093/cercor/bhr310>
- Kotz, S. A., Stockert, A., & Schwartz, M. (2014). *Cerebellum, temporal predictability, and the updating of a mental model*. *Philosophical Transactions of the Royal Society London B: Biological Sciences*, 369(1658), 20130403.
- Kouider, S., Long, B., Le Stanc, L., Charron, S., Fievet, A. C., Barbosa, L. S., & Gelskov, S. V. (2015). Neural dynamics of prediction and surprise in infants. *Nature Communications*, 6, 8537. <http://dx.doi.org/10.1038/ncomms9537>
- Lange, K. (2013). The ups and downs of temporal orienting: A review of auditory temporal orienting studies and a model associating the heterogeneous findings on the auditory N1 with opposite effects of attention and prediction. *Frontiers in Human Neuroscience*, 7. <http://dx.doi.org/10.3389/fnhum.2013.00263>
- Lasaponara, S., Chica, A. B., Lecce, F., Lupiáñez, J., & Doricchi, F. (2011). ERP evidence for selective drop in attentional costs in uncertain environments: Challenging a purely premotor account of covert orienting of attention. *Neuropsychologia*, 49(9), 2648–2657. <http://dx.doi.org/10.1016/j.neuropsychologia.2011.05.01>
- Lawrence, M. A. (2013). ez: Easy analysis and visualization of factorial experiments. R package version 4.2-2. <http://CRAN.R-project.org/package=ez>
- Lee, T. S., & Mumford, D. (2003). Hierarchical bayesian inference in the visual cortex. *Journal of the Optical Society of America A*, 20(7), 1434. <http://dx.doi.org/10.1364/josaa.20.001434>
- Luck, S. J., Woodman, G. F., & Vogel, E. K. (2000). Event-related potential studies of attention. *Trends in Cognitive Sciences*, 4(11), 432–440. [http://dx.doi.org/10.1016/S1364-6613\(00\)01545-X](http://dx.doi.org/10.1016/S1364-6613(00)01545-X)
- Martínez, A., Anllo-Vento, L., Sereno, M. I., Frank, L. R., Buxton, R. B., Dubowitz, D. J., & Hillyard, S. A. (1999). Involvement of striate and extrastriate visual cortical areas in spatial attention. *Nature Neuroscience*, 2(4), 364–369. <http://dx.doi.org/10.1038/7274>
- Martínez, A., Di Russo, F., Anllo-Vento, L., Sereno, M. I., Buxton, R. B., & Hillyard, S. A. (2001). Putting spatial attention on the map: Timing and localization of stimulus selection processes in striate and extrastriate visual areas. *Vision Research*, 41(10–11), 1437–1457.
- May, P. J., & Tiitinen, H. (2010). Mismatch negativity (MMN), the deviance-elicited auditory deflection, explained. *Psychophysiology*, 47(1), 66–122. <http://dx.doi.org/10.1111/j.1469-8986.2009.00856.x>
- Mishra, J., Martínez, A., Schroeder, C. E., & Hillyard, S. A. (2012). Spatial attention boosts short-latency neural responses in human visual cortex. *Neuroimage*, 59(2), 1968–1978. <http://dx.doi.org/10.1016/j.neuroimage.2011.09.02>
- Näätänen, R., Gaillard, A. W., & Mäntysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta Psychologica*, 42(4), 313–329.
- Näätänen, R., Alho, K., & Schröger, E. (2002). Electrophysiology of attention. In H. Pashler, & H. Wixted (Eds.), *Stevens' handbook of experimental psychology: third edition volume four: Methodology in experimental psychology* (pp. 601–653). New York: John Wiley.
- Nieuwenhuis, S., De Geus, E. J., & Aston-Jones, G. (2011). The anatomical and functional relationship between the P3 and autonomic components of the orienting response. *Psychophysiology*, 48(2), 162–175. <http://dx.doi.org/10.1111/j.1469-8986.2010.01057.x>
- Nobre, A. C., Sebestyen, G. N., & Miniussi, C. (2000). The dynamics of shifting visuospatial attention revealed by event-related potentials. *Neuropsychologia*, 38(7), 964–974.
- Noesselt, T., Hillyard, S. A., Woldorff, M. G., Schoenfeld, A., Hagner, T., Jäncke, L., & Heinze, H. J. (2002). Delayed striate cortical activation during spatial attention. *Neuron*, 35(3), 575–587.
- O'Shea, R. P. (2015). Refractoriness about adaptation. *Frontiers in Human Neuroscience*, 9, 38. <http://dx.doi.org/10.3389/fnhum.2015.00038>
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, 9(1), 97–113.
- Paris, T., Kim, J., & Davis, C. (2016). The processing of attended and predicted sounds in time. *Journal of Cognitive Neuroscience*, 28(1), 158–165.
- Pelli, D. G. (1997). The videotoolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10(4), 437–442.
- Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, 118(10), 2128–2148. <http://dx.doi.org/10.1016/j.clinph.2007.04.019>
- Posner, M. I. (1980). Orienting of attention. *The Quarterly Journal of Experimental Psychology*, 32(1), 3–25.
- R Core Team. (2015). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing [URL <http://www.R-project.org/>].
- Rahne, T., von Specht, H., & Mühler, R. (2008). Sorted averaging-application to auditory event-related responses. *Journal of Neuroscience Methods*, 172(1), 74–78. <http://dx.doi.org/10.1016/j.jneumeth.2008.04.006>
- Roussel, C., Hughes, G., & Waszak, F. (2014). Action prediction modulates both neurophysiological and psychophysical indices of sensory attenuation. *Frontiers in Human Neuroscience*, 8, 115. <http://dx.doi.org/10.3389/fnhum.2014.0011>
- Salisbury, D. F., Griggs, C. B., Shenton, M. E., & McCarley, R. W. (2004). The nogo P300 'anteriorization' effect and response inhibition. *Clinical Neurophysiology*, 115(7), 1550–1558. <http://dx.doi.org/10.1016/j.clinph.2004.01.028>
- SanMiguel, I., Saupe, K., & Schröger, E. (2013). I know what is missing here: Electrophysiological prediction error signals elicited by omissions of predicted what but not when. *Frontiers in Human Neuroscience*, 7, 407.
- Saupe, K., Widmann, A., Trujillo-Barreto, N. J., & Schröger, E. (2013). Sensorial suppression of self-generated sounds and its dependence on attention. *International Journal of Psychophysiology*, 90(3), 300–310. <http://dx.doi.org/10.1016/j.ijpsycho.2013.09.006>
- Schröger, E., & Eimer, M. (1993). Effects of transient spatial attention on auditory event-related potentials. *Neuroreport*, 4(5), 588–590.
- Schröger, E., Marzecová, A., & SanMiguel, I. (2015). Attention and prediction in human audition: A lesson from cognitive psychophysiology. *The European Journal of Neuroscience*, 41(5), 641–664. <http://dx.doi.org/10.1111/ejn.12816>
- Simpson, G. V., Weber, D. L., Dale, C. L., Pantazis, D., Bressler, S. L., Leahy, R. M., & Luks, T. L. (2011). Dynamic activation of frontal, parietal, and sensory regions underlying anticipatory visual spatial attention. *The Journal of Neuroscience*, 31(38)–13889. <http://dx.doi.org/10.1523/JNEUROSCI.1519-10.2011>

- Slagter, H. A., Prinssen, S., Reteig, L. C., & Mazaheri, A. (2016). Facilitation and inhibition in attention: Functional dissociation of pre-stimulus alpha activity, P1, and N1 components. *NeuroImage*, 125, 25–35. <http://dx.doi.org/10.1016/j.neuroimage.2015.09.058>
- St John-Saaltink, E., Utzerath, C., Kok, P., Lau, H. C., & de Lange, F. P. (2015). Expectation suppression in early visual cortex depends on task set. *PLoS One*, 10(6), e0131172. <http://dx.doi.org/10.1371/journal.pone.0131172>
- Stefanics, G., Kremláček, J., & Czigler, I. (2014). Visual mismatch negativity: A predictive coding view. *Frontiers in Human Neuroscience*, 8.
- Strauss, M., Sitt, J. D., King, J. R., Elbaz, M., Azizi, L., Buiatti, M., & Dehaene, S. (2015). Disruption of hierarchical predictive coding during sleep. *Proceedings of the National Academy of Sciences of the United States of America*, 112(11), E1353–E1362. <http://dx.doi.org/10.1073/pnas.1501026112>
- Summerfield, C., & Egner, T. (2009). Expectation (and attention) in visual cognition. *Trends in Cognitive Sciences*, 13(9), 403–409. <http://dx.doi.org/10.1016/j.tics.2009.06.003>
- Summerfield, C., & Egner, T. (2014). Attention and decision-making. In A. C. Nobre, & S. Kastner (Eds.), *The Oxford handbook of attention* (pp. 837–864). Oxford: Oxford University Press.
- Summerfield, C., & de Lange, F. P. (2014). Expectation in perceptual decision making: Neural and computational mechanisms. *Nature Reviews. Neuroscience*, 15(11), 745–756. <http://dx.doi.org/10.1038/nrn3838>
- Talsma, D., Mulckhuyse, M., Slagter, H. A., & Theeuwes, J. (2007). Faster, more intense! The relation between electrophysiological reflections of attentional orienting, sensory gain control, and speed of responding. *Brain Research*, 1178, 92–105. <http://dx.doi.org/10.1016/j.brainres.2007.07.099>
- Talsma, D., Sikkens, J. J., & Theeuwes, J. (2011). Stay tuned: What is special about not shifting attention? *PLoS One*, 6(3), e16829. <http://dx.doi.org/10.1371/journal.pone.0016829>
- Tata, M. S., & Ward, L. M. (2005). Spatial attention modulates activity in a posterior where auditory pathway. *Neuropsychologia*, 43(4), 509–516. <http://dx.doi.org/10.1016/j.neuropsychologia.2004.07.019>
- Timm, J., SanMiguel, I., Saupe, K., & Schröger, E. (2013). The N1-suppression effect for self-initiated sounds is independent of attention. *BMC Neuroscience*, 14, 2. <http://dx.doi.org/10.1186/1471-2202-14-2>
- Todorovic, A., Schoffelen, J. M., van Ede, F., Maris, E., & de Lange, F. P. (2015). Temporal expectation and attention jointly modulate auditory oscillatory activity in the beta band. *PLoS One*, 10(3), e0120288. <http://dx.doi.org/10.1371/journal.pone.0120288>
- van der Lubbe, R., & Utzerath, C. (2013). Lateralized power spectra of the EEG as an index of visuospatial attention. *Advances in Cognitive Psychology*, 9(4), 184–201. <http://dx.doi.org/10.5709/acp-0144-7>
- van der Lubbe, R. H., & Woestenburg, J. C. (2000). Location selection in the visual domain. *Psychophysiology*, 37(5), 662–676. <http://dx.doi.org/10.1111/1469-8986.3750662>
- van der Lubbe, R. H., Neggers, S. F., Verleger, R., & Kenemans, J. L. (2006). Spatiotemporal overlap between brain activation related to saccade preparation and attentional orienting. *Brain Research*, 1072(1), 133–152. <http://dx.doi.org/10.1016/j.brainres.2005.11.087>
- VanRullen, R. (2011). Four common conceptual fallacies in mapping the time course of recognition. *Frontiers in Psychology*, 2, 365. <http://dx.doi.org/10.3389/fpsyg.2011.00365>
- Velzen, J. V., & Eimer, M. (2003). Early posterior ERP components do not reflect the control of attentional shifts toward expected peripheral events. *Psychophysiology*, 40(5), 827–831.
- Verleger, R., Jaskowski, P., & Wauschkuhn, B. (1994). Suspense and surprise: On the relationship between expectancies and P3. *Psychophysiology*, 31(4), 359–369.
- Wöstmann, M., Herrmann, B., Wilsch, A., & Obleser, J. (2015). Neural alpha dynamics in younger and older listeners reflect acoustic challenges and predictive benefits. *The Journal of Neuroscience*, 35(4), 1458–1467.
- Wacongne, C., Labyt, E., van Wassenhove, V., Bekinschtein, T., Naccache, L., & Dehaene, S. (2011). Evidence for a hierarchy of predictions and prediction errors in human cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 108(51), 20754–20759. <http://dx.doi.org/10.1073/pnas.1117807108>
- Yamaguchi, S., Tsuchiya, H., & Kobayashi, S. (1995). Electrophysiologic correlates of visuo-spatial attention shift. *Electroencephalography and Clinical Neurophysiology*, 94(6), 450–461.