

Attention Reverses the Effect of Prediction in Silencing Sensory Signals

Peter Kok¹, Dobromir Rahnev^{1,2}, Janneke F. M. Jehee¹, Hakwan C. Lau^{1,2} and Floris P. de Lange¹

¹Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, 6500 HB Nijmegen, the Netherlands and

²Department of Psychology, Columbia University, New York, NY 10027, USA

Address correspondence to Peter Kok, Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, PO Box 9101, 6500 HB Nijmegen, the Netherlands. Email: p.kok@donders.ru.nl.

Predictive coding models suggest that predicted sensory signals are attenuated (silencing of prediction error). These models, though influential, are challenged by the fact that prediction sometimes seems to enhance rather than reduce sensory signals, as in the case of attentional cueing experiments. One possible explanation is that in these experiments, prediction (i.e., stimulus probability) is confounded with attention (i.e., task relevance), which is known to boost rather than reduce sensory signal. However, recent theoretical work on predictive coding inspires an alternative hypothesis and suggests that attention and prediction operate synergistically to improve the precision of perceptual inference. This model posits that attention leads to heightened weighting of sensory evidence, thereby reversing the sensory silencing by prediction. Here, we factorially manipulated attention and prediction in a functional magnetic resonance imaging study and distinguished between these 2 hypotheses. Our results support a predictive coding model wherein attention reverses the sensory attenuation of predicted signals.

Keywords: expectation, fMRI, predictive coding, spatial attention

Introduction

Over the past decades, it has become increasingly clear that perception is not determined solely by sensory input but is strongly influenced by prior knowledge. Predictive coding models of perception suggest that predicted sensory signals are attenuated, through inhibition of those sensory inputs that are consistent with current top-down predictions (i.e., silencing of prediction error; Rao and Ballard 1999; Friston 2005; Jehee and Ballard 2009). These models account for several extraclassical receptive field effects (Rao and Ballard 1999; Spratling 2010) and are in line with empirical findings that predicted stimuli evoke reduced neural responses (Alink et al. 2010; den Ouden et al. 2010; Todorovic et al. 2011). However, despite their computational and empirical appeal, predictive coding models have been challenged early on (Koch and Poggio 1999) by the fact that prediction sometimes seems to enhance rather than reduce sensory signals (Doherty et al. 2005; Chaumon et al. 2008), as in the case of attentional cueing experiments (Mangun and Hillyard 1991; Anllo-Vento 1995).

One possible explanation is that in these experiments, prediction (i.e., whether a stimulus is likely to be presented) is confounded with attention (i.e., whether a stimulus is behaviorally relevant; for a discussion, see Summerfield and Egner 2009), which is known to boost rather than reduce neural activity in sensory regions (Corbetta et al. 1990; Brefczynski and DeYoe 1999; Gandhi et al. 1999; Martínez et al. 1999; Somers et al. 1999; Boynton 2009; Reynolds and Heeger 2009). In fact, attention is often investigated by

manipulating subjects' expectations about upcoming stimuli (Posner 1980), and the terms "attention" and "expectation" are sometimes used interchangeably (Kastner et al. 1999; Corbetta and Shulman 2002). Thus, attentional cueing may enhance sensory activity, only because the negative impact of prediction is outweighed in magnitude by the positive impact of attentional mechanisms. In this account, prediction and attention have opposing main effects (Fig. 1A).

However, recent Bayesian models of perception inspire an alternative hypothesis. They propose that attention may enhance the precision of perceptual inference (Rao 2005; Friston 2009). Under this account, attention and prediction operate synergistically to optimize perceptual inference. More specifically, attention boosts the precision of predictions, leading to heightened weighting of sensory evidence (or, equivalently, prediction error), thereby reversing the effect of sensory silencing by prediction alone (for a schematic depiction of this effect, see Fig. 2). This reflects the fact that in Bayesian formulations of perceptual inference, prediction errors are weighted according to their precision, equivalent to weighting residuals by the inverse variance (precision) of the measurement. In other words, prediction errors are weighted according to how informative they are, and the information carried by a prediction error depends on both the current predictions and the reliability of the new data. In this sense, spatial attention can be thought of as "highlighting" a region of space, thereby increasing the precision of information coming from this region. A model implementing this mechanism has been shown to successfully simulate electrophysiological and psychophysical correlates of the Posner spatial cueing paradigm (Feldman and Friston 2010).

Mechanistically, precision is assumed to be encoded by the postsynaptic gain of neurons representing sensory data (prediction error), and attention increases precision by boosting this synaptic gain, in line with established accounts of the mechanisms of attention (Treue and Martínez Trujillo 1999; McAdams and Maunsell 2000; Reynolds and Heeger 2009). Accordingly, neuroimaging studies have provided evidence for attentional enhancement of sensory signals in human visual cortex (Corbetta et al. 1990; Kastner et al. 1998; Gandhi et al. 1999; Martínez et al. 1999).

This account predicts a specific interaction between prediction and attention: In the absence of attention, prediction attenuates sensory signals, but attention reverses this effect by boosting the precision of predictions (Figs 1B and 2). This interaction reflects the fact that attentional boosting of prediction errors rests upon the presence of predictions (and subsequent errors).

In summary, prediction and attention may operate as separable and antagonistic processes, the effects of which are purely additive. Conversely, attention may be an integral part of

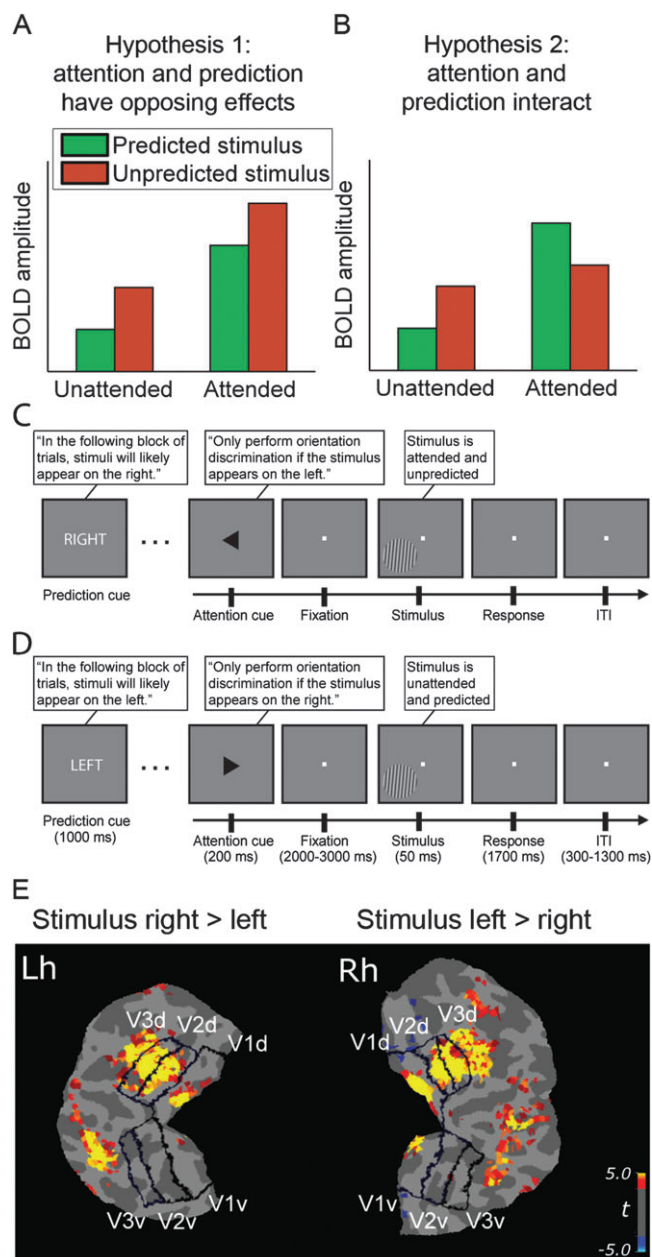


Figure 1. Hypotheses, paradigm, and retinotopy. We adjudicated between 2 hypotheses relating to the role of attention and prediction. (A) Hypothesis 1: attention and prediction have opposing effects, and the excitatory effect of attention outweighs the inhibitory effect of prediction, thereby leading to an increased response to predicted attended stimuli compared with unpredicted unattended stimuli (corresponding to “valid” and “invalid” in spatial cueing terminology, respectively). Under this hypothesis, attention and prediction have purely additive effects. (B) Hypothesis 2: attention and prediction work together synergistically to improve the precision of predictions, leading to stronger weighting of prediction error for predicted attended stimuli. Like hypothesis 1, this hypothesis also predicts an increased response to predicted attended stimuli compared with unpredicted unattended stimuli but due to an interaction of attention and prediction, instead of opposing main effects. (C,D) Paradigm. Each block started with a prediction cue that indicated the likely location of stimuli in the subsequent block (cue validity was 75%). Each trial started with an attention cue, which informed subjects about which side was task relevant. After a variable delay, a stimulus appeared on either the left or the right side. When the stimulus appeared on the task-relevant (attended) side, subjects had to indicate the orientation of the grating. Note that the attention cue had no predictive value with regard to stimulus location. (C) Example of an attended unpredicted trial. Subject predicts a stimulus on the right side but is instructed to attend to the left side. The stimulus appears on the left (attended, unpredicted) side. (D) Example of an unattended predicted trial. Subject predicts a stimulus on the left side but attends to the

optimal prediction and therefore depend on the emergence of predictions and their errors. Formally, the key feature that distinguishes these 2 hypotheses is the presence of an interaction between prediction and attention. In this study, we factorially manipulated spatial attention and prediction and probed their respective effects on the neural responses evoked by visual stimuli using functional magnetic resonance imaging (fMRI), to adjudicate between these 2 hypotheses.

Subjects performed an orientation identification task on gratings that could appear either on the left or on the right of fixation. Spatial attention and prediction were manipulated by 2 separate (and independent) cues, presented at fixation (Fig. 1C,D). The attention cue indicated which visual hemifield was task relevant; subjects were instructed only to perform the orientation identification task if the stimulus appeared in the indicated hemifield. Attentional orienting was encouraged further by presenting the grating stimulus briefly (50 ms) and embedded in noise. Importantly, the attention cue contained no information on the likelihood of a stimulus appearing in the indicated hemifield. Likelihood was indicated by a separate prediction cue, which appeared before each block of 8 trials, consisting of either the word “left” (indicating a 75% likelihood of stimuli appearing on the left), “right” (indicating a 75% likelihood of stimuli appearing on the right), or “neutral” (indicating a 50% likelihood of stimuli appearing on either side). The prediction cue was independent of the attention cue and therefore contained no information on the task relevance of stimuli.

Our data reveal an interaction between prediction and attention in early visual cortex: Predicted stimuli engendered reduced activity compared with unpredicted stimuli when they were unattended and task irrelevant, but this pattern reversed when the stimuli were attended and task relevant. Therefore, these results support a predictive coding model wherein attention and prediction operate synergistically to improve the precision of perceptual inference (Friston 2009). This contributes to resolving the controversy in the literature regarding the effects of prediction on neural responses (Rauss et al. 2011).

Materials and Methods

Subjects

Twenty-two healthy right-handed individuals (15 females, age 24 ± 3.4 , mean \pm standard deviation [SD]) with normal or corrected-to-normal vision gave written informed consent to participate in this study, in accordance with the institutional guidelines of the local ethics committee (Commissie Mensgebonden Onderzoek region Arnhem-Nijmegen, the Netherlands). Data from 3 subjects were excluded due to excessive head movement (more than 3 mm within an experimental session). Two subjects were excluded since they reported after the experiment that they had not understood the task, in particular the meaning of the cues.

Experimental Design

Subjects engaged in a grating orientation identification task. The task was divided into 2 sessions, with each session consisting of 42 blocks of 8 trials, yielding a total of 672 trials per subject. Each block started with

right side. The stimulus appears on the left (unattended, predicted) side. (E) Retinotopic mapping and stimulus activation are overlaid on flattened occipital cortices of a representative subject. Activation maps indicate stimulus-responsive voxels, black lines represent borders of V1, V2, and V3. Stimuli activated dorsal parts of contralateral V1, V2, and V3.

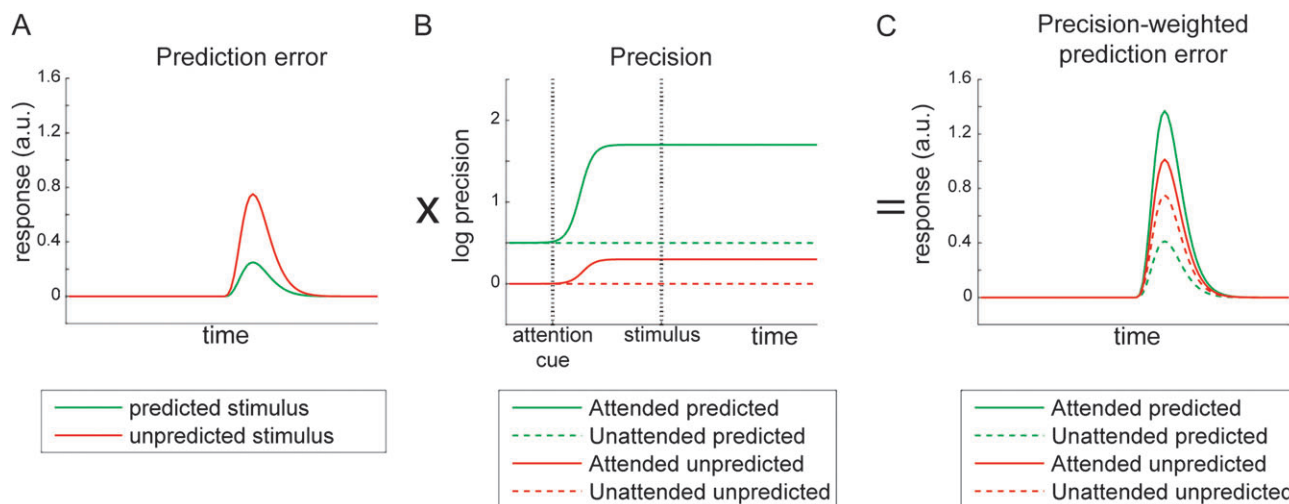


Figure 2. Schematic overview of the joint effects of attention (precision) and prediction on prediction error. (A) Depending on the context, a physically identical stimulus can match an observer's predictions to different degrees. The prediction error is the mismatch between predicted and actual input. Here, we sketch hypothetical prediction error responses to physically identical stimuli, preceded by either a valid (green) or an invalid (red) prediction cue. The amplitudes of the responses were chosen arbitrarily. (B) According to recent models of perceptual inference (Rao 2005; Friston 2009; Feldman and Friston 2010), attention increases the precision of predictions. Statistically, this means that subsequent prediction errors will be weighted more strongly, since optimal inference involves weighting data in proportion to their precision. Mechanistically, this might be implemented by increasing the postsynaptic gain of neurons reporting unexplained sensory data (prediction error). This enhancement of precision by attention occurs in relation to current predictions, reflected here by the fact that attention hardly increases precision when no stimulus is predicted to occur. Mechanistically, this might mean that attention increases the postsynaptic gain of only those prediction error neurons that receive their inputs from neurons that are "primed" by the current predictions. The order of magnitude of the precision values displayed here was based on figures in Feldman and Friston (2010), the exact values were chosen arbitrarily, and their evolution over time was simplified. (C) Prediction errors are weighted by their precision, calculated here as a simple multiplication of prediction error (panel A) and precision (panel B). The fact that attention enhances precision in relation to current predictions leads to an interactive effect of prediction and attention on the amplitude of the prediction error response.

a prediction cue that indicated the likely location of stimuli in the subsequent block of trials (Fig. 1C,D). The cue was presented centrally for 1000 ms and consisted of either the word left (indicating a 75% likelihood of stimuli appearing on the left), right (indicating a 75% likelihood of stimuli appearing on the right), or neutral (a "no prediction" control condition, with a 50% likelihood of stimuli appearing on either side). The different block types were pseudorandomly interleaved (repetitions of the same block type were prevented).

Each trial started with a centrally presented attention cue (for 200 ms), consisting of a small triangle that indicated the hemifield subjects had to attend to, while maintaining fixation. Following a variable delay (2000–3000 ms), a circular grating stimulus was briefly flashed (50 ms) in either the left or the right visual field. Subjects were instructed to respond only when the stimulus appeared in the attended hemifield, that is, the hemifield indicated by the attention cue. Stimuli on the unattended side did not require any response and could be ignored. Importantly, the attention cue contained no information about the likely location of a subsequent stimulus; the prediction cue (see above) was the only probabilistic cue. Instead, the attention cue told subjects which visual hemifield was task relevant. Therefore, the attention cue had 100% validity in terms of where attention needed to be deployed for successful task performance (unlike classical Posner paradigms, in which subjects also need to perform their task if the stimulus appears in the "unattended" field, see Posner 1980). The response interval (1700 ms) was followed by a variable intertrial interval (300–1300 ms), resulting in an interstimulus interval of 4–6 s, sufficiently long to prevent low-level adaptation to a brief (50 ms) stimulus (Nelson 1991; Boynton and Finney 2003). The interval between cue and stimulus and the intertrial interval were jittered to optimize the efficiency of our event-related design and to be able to dissociate the responses to cues and stimuli (Dale and Buckner 1997). Subjects were instructed to maintain fixation on a centrally presented fixation point throughout the trial. There was a rest period after each 7 blocks and between the 2 sessions.

The stimulus consisted of a circular luminance-defined sinusoidal grating with a spatial frequency of 3.33 cycles/degree and fixed phase. Grating contrast ($6.3 \pm 1.1\%$, mean \pm SD) was set on the basis of

individual performance in a practice session outside of the MRI scanner environment and a short practice run inside the scanner. The grating subtended a visual angle of 3° , it was presented 1° below and 2° to the left or right of fixation and could be oriented either horizontally (95° orientation) or vertically (5° orientation). In order to drive subjects to allocate all attentional resources to the side indicated by the attention cue, the grating was presented briefly (50 ms) and embedded in uniform noise (60% contrast, noise was only present during stimulus presentation). Subjects were instructed to quickly and accurately judge the grating orientation when it appeared in the attended hemifield and report it by pressing 1 of 2 keys of a button box with their right hand.

Visual stimuli were generated using MATLAB 7 (MathWorks, Natick, MA) in conjunction with the Psychophysics Toolbox (Brainard 1997) and displayed on a rear-projection screen using an EIKI projector (1024 \times 768 resolution, 60 Hz refresh rate).

Eye Movement Recording

To verify that subjects maintained fixation on the central fixation point throughout the trial, we monitored subjects' eye movements using an infrared eye tracking system in the scanner (Sensomotoric Instruments, Berlin, Germany). We recorded eye movement data for 8 subjects (of the 17 subjects included in the final analyzed sample), which we checked for systematic differences in eye movements between conditions; we analyzed the difference between mean pupil positions 500 ms after and 500 ms before stimulus onset for each trial in analyses of variance (ANOVAs), including stimulus location, attention, and prediction as factors. We used separate ANOVAs to analyze vertical and horizontal deviations from fixation. In terms of behavioral results in the main experiment, these subjects were representative for the group, with no significant differences in terms of accuracy (2-sample *t*-test, $t_{15} = 0.08$, $P = 0.93$) and reaction times ($t_{15} = 0.47$, $P = 0.65$) between subjects with and without eye tracking.

fMRI Acquisition Parameters

Functional images were acquired using a 3-T Trio MRI system (Siemens, Erlangen, Germany), with a T_2^* -weighted gradient-echo echo-planar imaging sequence (time repetition [TR]/time echo [TE] = 1950/30 ms,

31 slices, voxel size $3 \times 3 \times 3$ mm, interslice gap 20%). Anatomical images were acquired with a T_1 -weighted magnetization prepared rapid gradient-echo (MP-RAGE) sequence, using a GRAPPA acceleration factor of 2 (TR/TE = 2300/3.03 ms, voxel size $1 \times 1 \times 1$ mm).

fMRI Data Analysis

We used SPM5 (<http://www.fil.ion.ucl.ac.uk/spm>, Wellcome Trust Centre for Neuroimaging, London, UK) for image preprocessing and analysis. The first 4 volumes of each subject's data set were discarded to allow for T_1 equilibration. All functional images were spatially realigned to the mean image, yielding head movement parameters, which were used as nuisance regressors in the general linear model (GLM), and temporally aligned to the first slice of each volume. The structural image was coregistered with the functional volumes. For retinotopic analyses, structural and functional images were not normalized to Montreal Neurological Institute (MNI) space, and functional images were spatially smoothed with a full-width at half-maximum (FWHM) of 4 mm. To enable whole-brain analysis at the group level, structural images were spatially normalized to a MNI T_1 template. Coregistration allowed the same transformation matrix to be used to spatially normalize the functional volumes to MNI space. Finally, the functional volumes were spatially smoothed with an isotropic Gaussian kernel with a FWHM of 8 mm.

Statistical analysis was performed in 2 stages. In the first stage, the data of each subject were modeled using an event-related approach, within the framework of the GLM. Regressors representing the different attention cues and stimuli were constructed by convolving cue and stimulus onsets with a canonical hemodynamic response function and its temporal derivative (Friston et al. 1998). Cues and stimuli appearing in trials in which subjects failed to respond to a relevant stimulus, or responded to an irrelevant one, were included as regressors of no interest, as were head motion parameters and their first-order derivatives (Lund et al. 2005). Finally, the data were high-pass filtered (cutoff 128 s) to remove low-frequency signal drifts.

The resulting parameter estimates for cue and stimulus regressors comprised the data for the second-level analysis. Effects of interest were specified using linear contrasts. For whole-brain analysis, statistical inference was performed using a corrected cluster threshold of $P < 0.05$, on the basis of a threshold of $P < 0.001$ at the voxel level.

Retinotopic Analysis

We performed retinotopic mapping to identify the boundaries of retinotopic areas in early visual cortex using well-established methods (Sereno et al. 1995; DeYoe et al. 1996; Engel et al. 1997). Subjects viewed a wedge, consisting of a flashing checkerboard pattern (3 Hz), first rotating clockwise for 9 cycles and then anticlockwise for another 9 cycles (at a rotation speed of 23.4 s/cycle). *FreeSurfer* (<http://surfer.nmr.mgh.harvard.edu/>) was used to generate inflated representations of the cortical surface from each subject's T_1 -weighted structural image and to analyze the functional data of the retinotopic mapping session. Fourier-based methods were used to obtain polar angle maps of the cortical surface, on the basis of which the borders of visual areas (dorsal and ventral V1, V2, and V3 in both hemispheres) could be defined for each subject (Sereno et al. 1995). These retinotopic maps were used to create regions of interest (ROIs) using MarsBaR (<http://marsbar.sourceforge.net/>).

Within each retinotopic ROI, we identified responsive and unresponsive voxels by selecting voxels according to their response to the grating stimulus (using the contrast "stimulus left > stimulus right" for ROIs in the right hemisphere and "stimulus right > stimulus left" for ROIs in the left hemisphere, for a similar approach, see, e.g., Bueti et al. 2010). Responsive voxels were defined as those above the 80th percentile of t values for the relevant contrast, with a threshold of $t > 1.65$ (approximately $P < 0.05$), while unresponsive voxels were those below the 20th percentile of absolute t values, with a threshold of $|t| < 0.5$. This yielded 2 ROIs for each visual area, one containing responsive voxels and one containing unresponsive voxels. MarsBaR was used to extract parameter estimates for cue and stimulus regressors from each ROI, for each subject. Prior to group-level analyses, data were collapsed across hemispheres.

In order to investigate whether effects of prediction changed over the course of a prediction block, as a result of potential bottom-up learning effects or switch costs, we created models of each subject's fMRI data in which each condition was represented by 2 separate regressors; one for trials occurring in the first half of a block and one for trials in the second half of a block. As a potentially more sensitive analysis, we also created models which included "time" (i.e., position of a trial within a prediction block, 1–8) as a linear parametric modulator. Also in these models, each condition was represented by 2 regressors, one modeling the blood oxygen level-dependent (BOLD) response evoked by the appropriate stimuli and the other its parametric modulation by time.

Dynamic Causal Modeling

In order to investigate the effects of prediction and attention on the dynamics within and between early visual areas, an effective connectivity analysis using Dynamical Causal Modeling (DCM; Friston et al. 2003) was performed. In DCM, the states of multiple interacting brain regions are modeled at the hidden (i.e., not directly observed with fMRI) neuronal level and combined with a hemodynamic forward model. A Bayesian estimation scheme is used to estimate a combined neuronal and hemodynamic parameter set, such that the modeled BOLD signals are maximally similar to the measured BOLD signals (Friston et al. 2003). The neuronal parameter set consists of 3 subtypes: driving inputs (i.e., the direct influence of stimuli on activity in a region), the fixed connectivity between regions (in the absence of experimental modulations), and the modulation of connectivity by experimental factors. Our system of interest consisted of early visual areas V1, V2, and V3, individually defined for each subject as the stimulus-responsive voxels in V1, V2, and V3, respectively, separately for the left and right hemisphere. The DCM analysis was therefore performed on the same voxels as the main analysis. Time series for each ROI were extracted as the principal eigenvariate across all voxels within the ROI, separately for the first and the second session of the experiment. The time series were adjusted for movement parameters and other regressors of no interest, retaining just the effects of interest evoked by the experimental manipulations. For each subject, separate models were made for both hemispheres and both experimental sessions. The driving input to the system was determined by a regressor containing the onsets of all predicted and unpredicted stimuli contralateral to the hemisphere being modeled. We defined 2 modulatory influences: 1) prediction in the absence of attention and 2) prediction in the presence of attention, determined by regressors containing the onsets of all stimuli appearing in the unattended and the attended hemifield, respectively. To capture the effects of prediction, these regressors were parametrically modulated by the likelihood of the stimulus (25%, 50%, or 75%). In our models, all 3 regions were reciprocally connected to each other, and each region had an inhibitory self-connection (modeling the intrinsic decay of neuronal activity; Friston et al. 2003). Five different models were tested, differing in which connections the 2 factors were allowed to modulate (for a graphical depiction of the models, see Supplementary Fig. S3). In model A (the "full model"), both factors were allowed to modulate all connections between regions and each region's self-connection. Self-connections model the decay of neuronal activity and are therefore suitable to model effects related to prediction error; a larger prediction error is assumed to take longer to resolve. In model B, both factors were allowed to modulate only interregional connections. In the other 3 models, both factors were allowed to modulate V1's self-connection, and the forward connections (from V1 to V2 and V3 and from V2 to V3) were modulated by (C) neither factor, (D) both factors, and (E) only by "prediction in the presence of attention." The particular choice of the 3 models with modulation of forward connectivity had a principled motivation. If attention indeed results in increased precision of predictions (and therefore increased weighting of prediction errors), prediction error units in V2 and V3 should become more sensitive to bottom-up sensory information. In terms of our DCM, this would be expressed as an attention-dependent modulation of forward connectivity when, and only when, subjects predicted stimuli in these sensory channels. Each of these models was specified for both hemispheres and both experimental sessions, yielding $5 \times 2 \times 2$ models for each subject.

After Bayesian estimation of all parameters of the DCMs, we summed the negative free-energy approximation to the log model evidence (Stephan et al. 2009) for the 2 experimental sessions, yielding one value per hemisphere for each of the 5 different models per subject. These free-energy values were subsequently entered into a random-effects Bayesian Model Selection (BMS) at the group level (Stephan et al. 2009), separately for both hemispheres. This analysis yields Dirichlet distribution parameters, describing the probabilities for each model considered. These conditional model probabilities can then be used to calculate exceedance probabilities: The probability that a given model is more likely than any other model considered, given the data. Since our main prediction pertained to the effects of prediction and attention on self-connections, we used random-effects family level inference (Penny et al. 2010) to compare the family of models including these connections (models A and C-E) to the model not including them (model B). Subsequently, we applied Bayesian Model Averaging (Penny et al. 2010) to the winning family of models, in order to obtain evidence-weighted parameter estimates for connectivity modulations, per subject, session, and hemisphere. These parameter estimates were averaged over sessions and tested for significance at the group level through 1-sample *t*-tests, separately for the 2 hemispheres.

Results

Behavioral Results

Subjects successfully responded to virtually all stimuli appearing in the attended visual field and ignored virtually all stimuli appearing in the unattended field (response rates were 95.5% and 2.6%, respectively), indicating that the attention cue effectively manipulated the task relevance of stimuli. These percentages were not affected by the prediction cue (all $P > 0.10$, Supplementary Table S1), suggesting that we successfully isolated prediction and task relevance behaviorally. While the prediction cue did not significantly alter accuracy on the orientation identification task (87.3% correct, Supplementary Table S1), subjects did respond faster to predicted stimuli (mean reaction time [RT] = 806 ms) than to unpredicted stimuli (mean RT = 850 ms; $t_{16} = 3.164$, $P = 0.006$) and stimuli preceded by a neutral cue (mean RT = 839 ms; $t_{16} = 4.220$, $P < 0.001$; Supplementary Table S1).

A sensitive fixed-effects analysis revealed no significant differences in eye movements between conditions (stimulus location, attention, and prediction) in terms of either horizontal ($F_{11,4666} = 0.552$, $P = 0.869$) or vertical ($F_{11,4666} = 0.605$, $P = 0.826$) eye movements.

fMRI Results

Analyses of stimulus-evoked activity in early visual cortex (Fig. 1E) revealed a clear interactive effect of prediction and attention on stimulus processing (Fig. 3A; compare Fig. 1B), as opposed to 2 opposing main effects (Fig. 1A). Stimuli that appeared on the unattended side evoked a reduced response when they were predicted (75% likely) compared with when they were unpredicted (25% likely; $t_{16} = 2.40$, $P = 0.029$) in primary visual cortex (V1), as predicted by predictive coding models. This difference was not observed (significantly) in V2 ($t_{16} = 1.55$, $P = 0.139$) and V3 ($t_{16} = 1.01$, $P = 0.327$). On the other hand, stimuli that appeared on the attended side evoked a larger response in the early visual areas when they were predicted compared with when they were unpredicted (V1: $t_{16} = 2.25$, $P = 0.039$; V2: $t_{16} = 1.96$, $P = 0.067$; V3: $t_{16} = 2.67$, $P = 0.017$). This reversal of prediction-related suppression by attention is consistent with a synergistic boosting of precision by attention and prediction. To test whether these effects were

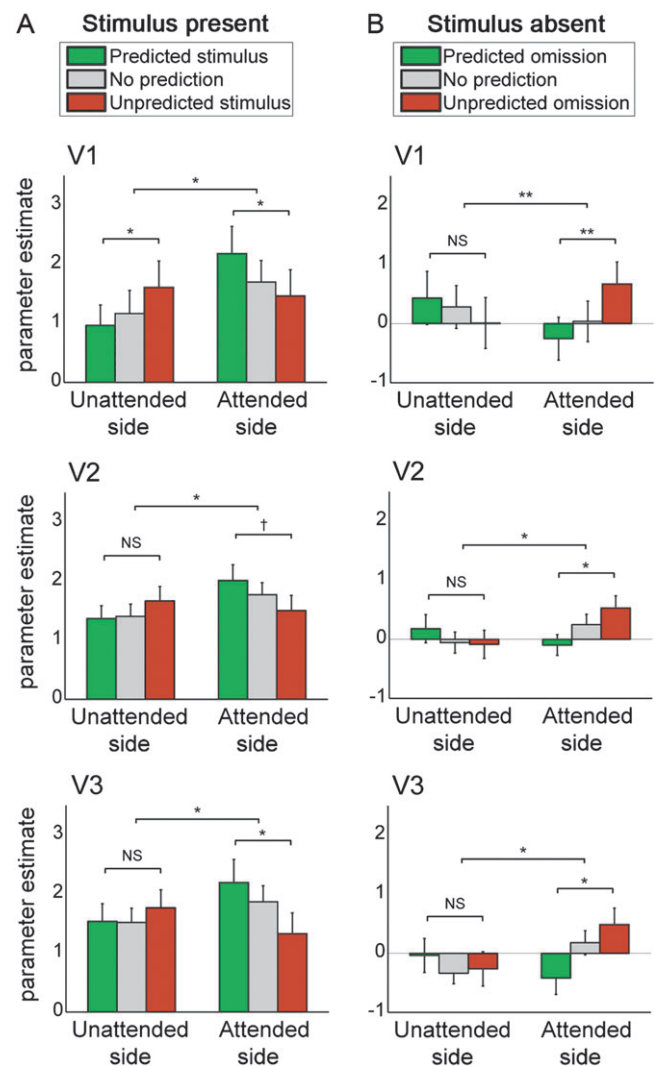


Figure 3. Stimulus-related activity in early visual cortex. (A) In the unattended visual field, predicted stimuli evoked a reduced response in V1 compared with unpredicted stimuli. In the attended field, predicted stimuli evoked a larger response in V1, V2, and V3 than unpredicted stimuli. (B) In visual cortex corresponding to the visual field where no stimulus appeared, that is, ipsilateral to the stimulus, unpredicted omission of a stimulus in the attended visual field evoked a larger response in V1, V2, and V3 than predicted omission of a stimulus. Error bars indicate standard error of the mean (* $P < 0.05$, ** $P < 0.01$, †0.067).

specific to neural regions involved in encoding the stimulus or reflect general activity modulations, we performed the same analyses for stimulus-unresponsive voxels in V1–V3 (for details, see Materials and Methods). No significant effects of attention and prediction were observed in stimulus-unresponsive regions (all $P > 0.10$, Supplementary Fig. S4A). We also tested whether these effects were stable or changed over the course of a prediction block, as a result of potential bottom-up learning effects or switch costs. We did not find any evidence for changes of the effect of prediction over the course of a block (Supplementary Figs S1 and S2), that is, the effects of prediction within a block were stable.

Since predictive coding theories state that the response in sensory cortex is largely determined by the violation of predictions, it may be expected that the failure of a predicted stimulus to appear would similarly evoke a response (prediction error) in the relevant sensory cortex, even though no

physical stimulus is presented (den Ouden et al. 2009). Additionally, if attention indeed increases the precision of predictions, this effect may be particularly prominent when the visual location at which the stimulus was predicted to appear was attended as well. Therefore, we investigated responses in early visual cortex ipsilateral to the stimulus, that is, corresponding to the visual field location where no stimulus appeared (Fig. 3*B*) as a function of attention and prediction. Indeed, unpredicted omission of a stimulus in the attended visual field evoked a larger response in visual cortex contralateral to the attended hemifield than a predicted omission, in all 3 visual areas (V1: $t_{16} = 2.99$, $P = 0.009$; V2: $t_{16} = 2.79$, $P = 0.013$; V3: $t_{16} = 2.90$, $P = 0.010$), while there were no significant differences between predicted and unpredicted omissions in the unattended visual field (V1: $t_{16} = 1.52$, $P = 0.148$; V2: $t_{16} = 1.25$, $P = 0.229$; V3: $t_{16} = 0.96$, $P = 0.353$). Again, the specificity of these effects was tested by performing the same analyses for stimulus-unresponsive voxels: No effects of prediction on the omission of stimuli were found in either the attended or the unattended hemifield (all $P > 0.10$, Supplementary Fig. S4*B*).

Predictive coding accounts of prediction and attention (Feldman and Friston 2010) posit that predictions operate by silencing the prediction error within a cortical unit. Attention reverses this effect by boosting the precision of prediction errors, heightening the impact the region has on downstream regions. To test this hypothesis more directly, we carried out an analysis of interregional directed interactions using DCM (Friston et al. 2003) (for details, see Materials and Methods). We defined several models, differing in which connections were modulated by attention and prediction (see Supplementary Fig. S3 and Materials and Methods). For both hemispheres, the most appropriate model, as determined by BMS (Stephan et al. 2009) and family level inference (Penny et al. 2010), contained modulatory influences of prediction on the self-connection of V1, while the model in which prediction did not modulate this connection (model B) performed the worst (Supplementary Tables S2 and S3). In the absence of attention, prediction modulated V1's self-connection negatively (resulting in faster decay of stimulus-evoked activity), while prediction in the presence of attention positively modulated this connection (Supplementary Table S4 and Fig. 4). Model selection results were somewhat inconclusive with regard to the modulation of feedforward connections (Supplementary Table S2). We therefore applied Bayesian Model Averaging (Penny et al. 2010) to collapse all the models that included the modulation of V1's self-connection (i.e., the "winning" family, see Supplementary Table S3) together. As expected, the modulation of V1's self-connection by prediction in both the absence and the presence of attention was significant in both hemispheres, in these averaged models. Additionally, feedforward connections (from V1 to V2 and V3 and from V2 to V3) were significantly strengthened by prediction in the presence of attention, in both hemispheres (Supplementary Table S4). Note that these connectivity results are consistent with the finding that the effects in the attended hemifield were present in V1 through V3, while the effects of prediction in the unattended hemifield were significant only in V1 (Fig. 3*A,B*). In sum, DCM provided strong evidence for an interactive effect of prediction and attention on the self-connection of V1 and indications of an increase of feedforward connectivity by prediction in the presence of attention (Fig. 4).

In order to disentangle effects related to stimulus processing from cue-induced baseline shifts (which may only reflect a general state of preparedness rather than processing of prediction error per se), we also looked at BOLD signals time locked to the attention cues. We were able to do this because we modeled cues and targets separately in our event-related design. The experiment was designed in order to be able to separately estimate neural responses to cue and stimulus by introducing a variable delay interval between attention cues and stimuli (Fig. 1*C,D*, see Materials and Methods). We quantified the effect of the attention cue on prestimulus activity in early visual cortex by comparing neural activity between attended and unattended hemifields (Fig. 5*A*). Indeed, we observed prestimulus attentional enhancement effects, in line with previous studies on visual spatial attention (Luck et al. 1997; Kastner et al. 1999; Silver et al. 2007; Murray 2008). These effects were spatially specific: Voxels in early visual cortex that were unresponsive to the grating stimuli showed no prestimulus increase in activity by attention (all $P > 0.10$; Fig. 5*A*). Interestingly, this prestimulus attentional modulation was strongest when attention and prediction were incongruent (V1–V3, all $P < 0.03$; Fig. 5*B*), possibly reflecting increased reorienting of attention (Yantis et al. 2002) at the time of the attention cue, when the prediction cue had indicated the opposite hemifield compared with when the prediction cue indicated the same hemifield as the attention cue. The fact that prestimulus activity revealed a different pattern compared with stimulus-evoked activity further suggests that our primary result of the interaction between attention and prediction was specific to processing of prediction error itself, which is time locked to the onset of stimulus presentation.

While our report focuses on the synergistic relationship of prediction and attention in early sensory cortex, for the sake of completeness, we additionally performed more exploratory whole-brain analyses. The results of these analyses can be found in Supplementary Tables S5 and S6 and Figures S5 and S6.

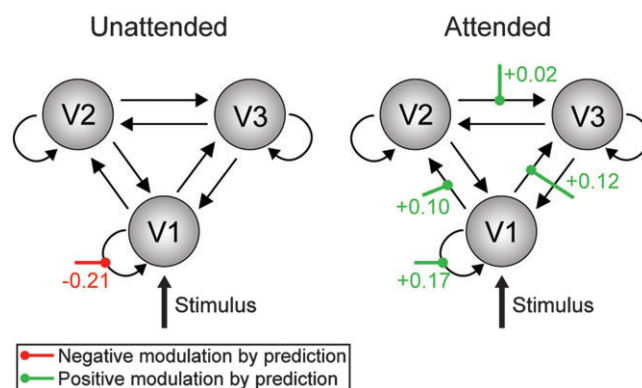


Figure 4. Modulation of neural dynamics in early visual cortex by prediction and attention. In the absence of attention, prediction negatively modulates V1's self-connection, causing stimulus-evoked activity to decay faster, possibly reflecting faster resolution of prediction error. However, in the presence of attention (right panel), prediction positively modulates V1's self-connection, causing activity to decay more slowly, consistent with a stronger weighting of prediction error. In addition, prediction and attention together strengthen the forward connections from V1 to V2 and V3 and from V2 to V3. Parameter values shown here were averaged over the left and right hemispheres; see for all parameters values and their statistical significance, Supplementary Table S4.

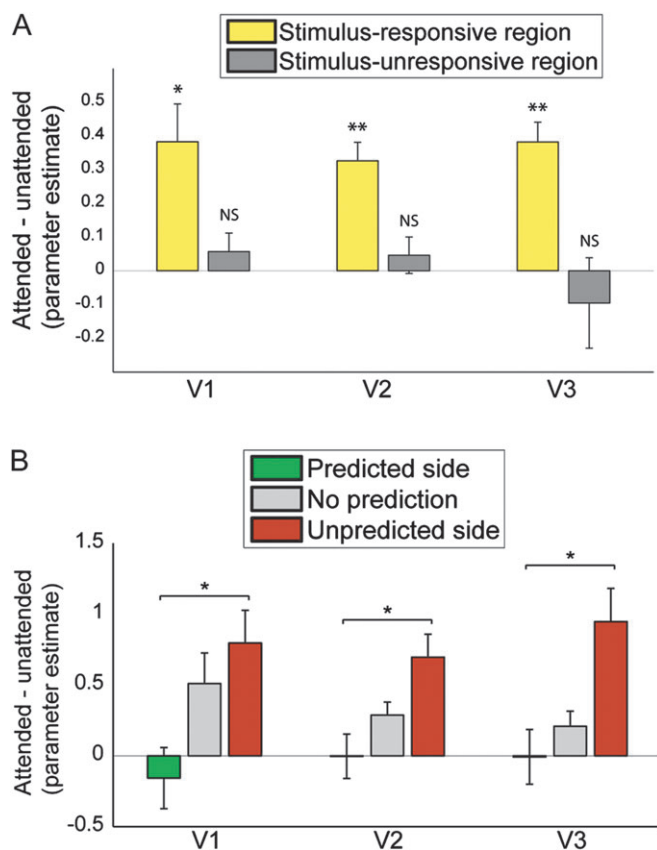


Figure 5. Prestimulus attentional enhancement in early visual cortex. (A) Attention led to a prestimulus increase in activity in contralateral visual cortex, only in stimulus-responsive voxels. Error bars indicate standard error of the mean (SEM) (* $P < 0.01$, ** $P < 0.001$). (B) An interaction between attention and prediction in attention cue-related activity: The increase in activity in early visual cortex contralateral to the attended visual field was strongest when the attention cue was incongruent with the prediction cue. Error bars indicate SEM (* $P < 0.05$).

Discussion

Predictive coding models of perception (Rao and Ballard 1999; Friston 2005, 2009; Jehee and Ballard 2009) have become highly influential in understanding how the brain deals with sensory inputs. However, these models have been challenged by reports of increased neural responses to predicted stimuli that are task relevant (Koch and Poggio 1999). Here, we provide empirical support for a predictive coding model wherein attention boosts the precision of predictions (Rao 2005; Friston 2009), leading to heightened weighting of sensory evidence (prediction error) and a reversal of the silencing effect of prediction.

Attention Reverses the Silencing Effect of Predictions on Sensory Signals

When stimuli were unattended, the neural response to predicted stimuli was reduced in early visual cortex (Fig. 3A) compared with unpredicted stimuli. This reduction in activity was significant in V1 but not in V2 and V3, suggesting that this silencing can take place at the earliest stage of the cortical hierarchy, precluding further forward propagation of the predicted sensory input through the cortical hierarchy. The reduction of activity for unattended predicted stimuli is in line with previous studies reporting reduced neural responses

to task-irrelevant predicted stimuli (den Ouden et al. 2009; Alink et al. 2010; den Ouden et al. 2010) and consistent with theoretical models of predictive coding (Rao and Ballard 1999; Friston 2005; Jehee and Ballard 2009). Predictive coding theories propose that (top-down) predictions and (bottom-up) sensory data are coded for by separate neuronal populations within cortical areas. Sensory data consistent with current predictions are inhibited, leading to remaining activity in sensory units (prediction error) representing discrepancies between current predictions and actual inputs, that is, that part of the sensory input not accounted for by current predictions. Accordingly, in these models, a good match between predicted and actual input (predicted stimuli) results in less neural activity in sensory units than a mismatch (unpredicted stimuli). Results of the effective connectivity analysis (DCM) suggest that prediction (in the absence of attention) negatively modulated V1's internal dynamics (Fig. 4), causing stimulus-evoked activity to decay faster. This may reflect faster resolution of prediction error within V1, consistent with recent computational models of predictive coding (Spatring 2008, 2010).

A different pattern of results was observed for attended stimuli. When stimuli were attended, the neural response in early visual cortex was larger in amplitude (Fig. 3A) for predicted compared with unpredicted stimuli. This is inconsistent with an explanation in terms of opposing main effects of prediction and attention (Fig. 1A), but it is consistent with a synergistic interaction between the 2 (Fig. 1B). An enhancing effect of prediction may at first glance seem incompatible with predictive coding theories (Koch and Poggio 1999), but recent modeling studies have shown that mechanisms of predictive coding and attention (specifically, biased competition; Desimone and Duncan 1995; Reynolds et al. 1999) can comfortably coexist within the same computational model (Feldman and Friston 2010) and can be cast as mathematically equivalent under certain assumptions (Spatring 2008). Recent Bayesian models of perception have proposed that attention may reflect the precision of perceptual inference (Rao 2005; Friston 2009; Hesselmann et al. 2010). Under this account, attention modulates the synaptic gain of neurons representing sensory data (or, equivalently, prediction error), causing prediction errors to be weighted according to the precision of the prediction. A model implementing this mechanism successfully simulated electrophysiological and psychophysical correlates of the Posner spatial cueing paradigm (Feldman and Friston 2010).

Considering attention as a mechanism of modulating the gain of neurons representing sensory data not explained by predictive feedback (prediction error) is in good accordance with our results. When attention and prediction were congruent, that is, when subjects predicted the stimulus to appear at the spatial location they were instructed to attend to, attention increased the precision of the prediction and weighted the sensory data more strongly, resulting in an increased response to predicted stimuli when they were attended compared with when they were unattended (Fig. 3A). When attention and prediction were incongruent, however, subjects did not predict the stimulus to appear at the spatial location they were attending to, and attention could therefore not boost the precision of the sensory prediction. This notion is also corroborated by the fact that the response to unpredicted stimuli was not different depending on whether

they were attended or unattended (Fig. 3*A* compare red bars). So it seems that the enhancement of sensory data by attention was contingent upon the prediction that a stimulus would appear at the attended location, leading to a larger response for attended stimuli when they were predicted than when they were unpredicted, despite the larger mismatch between sensory inputs and prediction in the latter case. Mechanistically, this contingency might reflect that attention increases the postsynaptic gain of only those prediction error neurons that receive their inputs from neurons that are “primed” by the current predictions.

If the increased response to predicted stimuli in the attended visual field was indeed established through the same mechanism that underlies the reduced response to predicted stimuli in the unattended field (i.e., prediction error), we would expect the same intra- and/or interregional connections to be modulated in both cases. Indeed, we found that prediction modulates the internal dynamics within V1 (silencing of prediction error), and this internal modulation is reversed by attention (Fig. 4), causing activity to decay more slowly, consistent with a stronger weighting of prediction error as a result of increased precision. In addition, DCM results suggested that attention strengthened the impact of the sensory input, by enhancing the forward drive of information from V1 to V2 and V3 and from V2 to V3. This causes the excitatory effect of prediction and attention to be propagated up the hierarchy, while the inhibitory effect of prediction in absence of attention is not propagated and confined to the primary visual cortex (see Fig. 3*A*). Similar feedforward effects of attention have previously been observed using electroencephalography (Zhang and Luck 2009). This appears an optimally efficient coding scheme for both amplification of relevant information and silencing of predicted irrelevant information.

Interestingly, the framework outlined above also provides an explanation for the increased activity in response to the unpredicted omission of a stimulus in the attended visual field (Fig. 3*B*). Here, attention and prediction together led to a strong and precise prediction of a stimulus appearing at this visual field location, and the violation of this prediction resulted in prediction error activity, even when no stimulus was present (den Ouden et al. 2009; Todorovic et al. 2011). The fact that this prediction error response was present in all 3 early visual areas (Fig. 3*B*) is consistent with the strengthening of forward connections by attention and prediction together, as suggested by DCM.

Can our findings be explained by attentional modulations alone? For instance, could it be that the increased response to unpredicted unattended stimuli compared with predicted unattended stimuli is the result of increased stimulus-driven attention to unpredicted stimuli? This explanation is hard to reconcile with the increased neural response to the unpredicted omission of a stimulus in the attended field, since stimulus-driven attention in this case would be engaged by the stimulus appearing in the unattended field, and not by the absence of a stimulus in the attended visual field. Additionally, in our experiment, the visual onset of both the predicted and the unpredicted unattended stimuli was irrelevant, and there was no difference between the stimuli in terms of bottom-up salience (which would cause bottom-up attention to be attracted more strongly).

Another potential alternative explanation of the reduced response to predicted unattended stimuli is that the unattended visual field was suppressed more strongly when a distracting,

irrelevant stimulus was likely to appear there than when such a distractor was unlikely. By this account, our results could be explained by 2 top-down attentional mechanisms: enhancement of relevant signals and suppression of distractors (Gazzaley et al. 2005). While this may partly explain our obtained results, it seems not fully consistent with it. First, distractor suppression cannot explain the increased response evoked by the unpredicted omission of a stimulus in the attended field, since 1) suppression should not occur for the task-relevant visual location and 2) in this case, there is no stimulus to suppress at this location. Also, whereas predictions in the presence of attention affected activity in and feedforward drive to all early visual regions (V1–V3), the prediction-modulation for unattended stimuli was only significant in V1. In contrast to this, the effects of top-down attention usually increase when progressing up the cortical hierarchy (Kastner et al. 1998; Bles et al. 2006; Buffalo et al. 2010).

In sum, while an explanation in terms of stimulus-driven or top-down attentional mechanisms does not seem fully consistent with our results, we cannot rule out such an interpretation on the basis of the current data alone.

Finally, since predicted stimuli, by definition, occurred more often than unpredicted stimuli, one may wonder whether our activity differences could be simply due to low-level sensory adaptation. This is particularly relevant since prediction was manipulated in blocks (of 8 trials) and not on a trial-by-trial basis. However, this explanation appears highly unlikely. First, the interstimulus interval was relatively long (4–6 s), and sensory adaptation effects in visual cortex appear absent at the time intervals used in the present study (Nelson 1991; Boynton and Finney 2003). More crucially, the fact that prediction had opposite effects on the amplitude of the neural response to a stimulus depending on attention is not compatible with an explanation of our results in terms of general low-level adaptation. However, future work would be necessary to conclusively rule out such an explanation, by varying both attention and prediction on a trial-by-trial basis.

Our results suggest that unexplained sensory signals (prediction errors) are weighted by their precision, a notion that is well in line with studies on the role of uncertainty in decision-related neural signals (Behrens et al. 2007; Kiani and Shadlen 2009). For example, Kiani and Shadlen (2009) showed that the neural activity in decision-related lateral intraparietal cortex (LIP) neurons increases with certainty about perceptual choice. In another recent study, Behrens et al. (2007) showed that prediction errors are weighted more strongly when they are more informative about future reward likelihood (i.e., in a volatile environment), as reflected by both behavior and activity in the anterior cingulate cortex (ACC). The current study deals with activity modulations in the early sensory regions that are presumably accumulated by downstream decision-related regions like LIP and ACC. Therefore, while our results relate to precision in the sensory signal rather than the decision signal, they are consistent with these studies (e.g., attended and predicted stimuli lead to high sensory precision, putatively leading to larger decision-related activity in areas like LIP, as evidenced by Kiani and Shadlen 2009).

Conclusion

Predictive coding has become a highly influential theory of perceptual inference in the last decade (Rao and Ballard 1999;

Friston 2005, 2009) but has been challenged by the observation that prediction enhances rather than reduces neural responses to task-relevant stimuli (Koch and Poggio 1999; for a review, see Rauss et al. 2011). For example, 2 recent studies found opposite effects of predictability of a visual stimulus on neural activity in early visual areas (Doherty et al. 2005; Alink et al. 2010). Notably, the stimulus was task relevant in the former but irrelevant in the latter study. Recent theoretical work on predictive coding offers a resolution of this problem, by suggesting that prediction and attention work together synergistically to improve the precision of perceptual inference (Friston 2009; Feldman and Friston 2010). Our results provide empirical support for this framework by showing that attention reverses the sensory silencing effects of prediction and may thereby explain the seemingly contradictory findings in the literature regarding the effects of prediction on neural activity (Rauss et al. 2011).

Supplementary Material

Supplementary material can be found at: <http://www.cercor.oxfordjournals.org/>

Funding

Netherlands Organization for Scientific Research (NWO VENI 451-09-001 awarded to F.P.d.L.).

Notes

We thank Jascha Swisher for assistance with the retinotopic analysis, Hanneke den Ouden for helpful discussions on DCM, and Paul Gaalman for MRI support. *Conflict of Interest*: None declared.

References

- Alink A, Schwiedrzik CM, Kohler A, Singer W, Muckli L. 2010. Stimulus predictability reduces responses in primary visual cortex. *J Neurosci*. 30(8):2960–2966.
- Anllo-Vento L. 1995. Shifting attention in visual space: the effects of peripheral cueing on brain cortical potentials. *Int J Neurosci*. 80:353–370.
- Behrens TE, Woolrich MW, Walton ME, Rushworth MF. 2007. Learning the value of information in an uncertain world. *Nat Neurosci*. 10(9):1214–1221.
- Bles M, Schwarzbach J, De Weerd P, Goebel R, Jansma BM. 2006. Receptive field size-dependent attention effects in simultaneously presented stimulus displays. *Neuroimage*. 30:506–511.
- Boynton GM. 2009. A framework for describing the effects of attention on visual responses. *Vision Res*. 49:1129–1143.
- Boynton GM, Finney EM. 2003. Orientation-specific adaptation in human visual cortex. *J Neurosci*. 23(25):8781–8787.
- Brainard DH. 1997. The psychophysics toolbox. *Spat Vis*. 10:433–436.
- Brefczynski JA, DeYoe EA. 1999. A physiological correlate of the ‘spotlight’ of visual attention. *Nat Neurosci*. 2(4):370–374.
- Bueti D, Bahrami B, Walsh V, Rees G. 2010. Encoding of temporal probabilities in the human brain. *J Neurosci*. 30(12):4343–4352.
- Buffalo EA, Fries P, Landman R, Liang H, Desimone R. 2010. A backward progression of attentional effects in the ventral stream. *Proc Natl Acad Sci U S A*. 107(1):361–365.
- Chaumon M, Drouet V, Tallon-Baudry C. 2008. Unconscious associative memory affects visual processing before 100 ms. *J Vis*. 8(3):1–10.
- Corbetta M, Miezin FM, Dobmeyer S, Shulman GL, Petersen SE. 1990. Attentional modulation of neural processing of shape, color, and velocity in humans. *Science*. 248(4962):1556–1559.
- Corbetta M, Shulman GL. 2002. Control of goal-directed and stimulus-driven attention in the brain. *Nat Rev Neurosci*. 3(3):201–215.
- Dale AM, Buckner RL. 1997. Selective averaging of rapidly presented individual trials using fMRI. *Hum Brain Mapp*. 5:329–340.
- den Ouden HEM, Daunizeau J, Roiser J, Friston KJ, Stephan KE. 2010. Striatal prediction error modulates cortical coupling. *J Neurosci*. 30(9):3210–3219.
- den Ouden HEM, Friston KJ, Daw ND, McIntosh AR, Stephan KE. 2009. A dual role for prediction error in associative learning. *Cereb Cortex*. 19:1175–1185.
- Desimone R, Duncan J. 1995. Neural mechanisms of selective visual attention. *Annu Rev Neurosci*. 18:193–222.
- DeYoe EA, Carman GJ, Bandettini P, Glickman S, Wieser J, Cox R, Miller D, Neitz J. 1996. Mapping striate and extrastriate visual areas in human cerebral cortex. *Proc Natl Acad Sci U S A*. 93:2382–2386.
- Doherty JR, Rao A, Mesulam MM, Nobre AC. 2005. Synergistic effect of combined temporal and spatial expectations on visual attention. *J Neurosci*. 25(36):8259–8266.
- Engel SA, Glover GH, Wandell BA. 1997. Retinotopic organization in human visual cortex and the spatial precision of functional MRI. *Cereb Cortex*. 7:181–192.
- Feldman H, Friston KJ. 2010. Attention, uncertainty, and free-energy. *Front Hum Neurosci*. 4:215.
- Friston KJ. 2005. A theory of cortical responses. *Philos Trans R Soc Lond B Biol Sci*. 360(1456):815–836.
- Friston KJ. 2009. The free-energy principle: a rough guide to the brain? *Trends Cogn Sci*. 13(7):293–301.
- Friston KJ, Fletcher P, Josephs O, Holmes A, Rugg MD, Turner R. 1998. Event-related fMRI: characterizing differential responses. *Neuroimage*. 7(1):30–40.
- Friston KJ, Harrison L, Penny W. 2003. Dynamic causal modelling. *Neuroimage*. 19(4):1273–1302.
- Gandhi SP, Heeger DJ, Boynton GM. 1999. Spatial attention affects brain activity in human primary visual cortex. *Proc Natl Acad Sci U S A*. 96:3314–3319.
- Gazzaley A, Cooney JW, McEvoy K, Knight RT, D’Esposito M. 2005. Top-down enhancement and suppression of the magnitude and speed of neural activity. *J Cogn Neurosci*. 17:507–517.
- Hesselmans G, Sadaghiani S, Friston KJ, Kleinschmidt A. 2010. Predictive coding or evidence accumulation? False inference and neuronal fluctuations. *PLoS One*. 5:e9926.
- Jehee JF, Ballard DH. 2009. Predictive feedback can account for biphasic responses in the lateral geniculate nucleus. *PLoS Comput Biol*. 5:e1000373.
- Kastner S, De Weerd P, Desimone R, Ungerleider LG. 1998. Mechanisms of directed attention in the human extrastriate cortex as revealed by functional MRI. *Science*. 282(108):108–111.
- Kastner S, Pinsk MA, De Weerd P, Desimone R, Ungerleider LG. 1999. Increased activity in human visual cortex during directed attention in the absence of visual stimulation. *Neuron*. 22(4):751–761.
- Kiani R, Shadlen MN. 2009. Representation of confidence associated with a decision by neurons in the parietal cortex. *Science*. 324:759–764.
- Koch C, Poggio T. 1999. Predicting the visual world: silence is golden. *Nat Neurosci*. 2(1):9–10.
- Luck SJ, Chelazzi L, Hillyard SA, Desimone R. 1997. Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. *J Neurophysiol*. 77(1):24–42.
- Lund TE, Norgaard MD, Rostrup E, Rowe JB, Paulson OB. 2005. Motion or activity: their role in intra- and inter-subject variation in fMRI. *Neuroimage*. 26(3):960–964.
- Mangun GR, Hillyard SA. 1991. Modulations of sensory-evoked brain potentials indicate changes in perceptual processing during visual-spatial priming. *J Exp Psychol*. 17(4):1057–1074.
- Martínez A, Anllo-Vento L, Sereno MI, Frank LR, Buxton RB, Dubowitz DJ, Wong EC, Hinrichs H, Heinze HJ, Hillyard SA. 1999. Involvement of striate and extrastriate visual cortical areas in spatial attention. *Nat Neurosci*. 2(4):364–369.
- McAdams CJ, Maunsell JHR. 2000. Attention to both space and feature modulates neuronal responses in macaque area V4. *J Neurophysiol*. 83:1751–1755.
- Murray SO. 2008. The effects of spatial attention in early human visual cortex are stimulus independent. *J Vis*. 8:1–11.

- Nelson SB. 1991. Temporal interactions in the cat visual system. I. Orientation-selective suppression in the visual cortex. *J Neurosci.* 11(2):344-356.
- Penny WD, Stephan KE, Daunizeau J, Rosa MJ, Friston KJ, Schofield TM, Leff AP. 2010. Comparing families of dynamic causal models. *PLoS Comput Biol.* 6(3):e1000709.
- Posner MI. 1980. Orienting of attention. *Q J Exp Psychol.* 32(1):3-25.
- Rao RPN. 2005. Bayesian inference and attentional modulation in the visual cortex. *Neuroreport.* 16:1843-1848.
- Rao RPN, Ballard DH. 1999. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat Neurosci.* 2(1):79-87.
- Rauss K, Schwartz S, Pourtois G. 2011. Top-down effects on early visual processing in humans: a predictive coding framework. *Neurosci Biobehav Rev.* 35:1237-1253.
- Reynolds JH, Chelazzi L, Desimone R. 1999. Competitive mechanisms subserve attention in macaque areas V2 and V4. *J Neurosci.* 19(5):1736-1753.
- Reynolds JH, Heeger DJ. 2009. The normalization model of attention. *Neuron.* 61(2):168-185.
- Sereno MI, Dale AM, Reppas JB, Kwong KK, Belliveau JW, Brady TJ, Rosen BR, Tootell RB. 1995. Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science.* 268:889-893.
- Silver MA, Ress D, Heeger DJ. 2007. Neural correlates of sustained spatial attention in human early visual cortex. *J Neurophysiol.* 97(1):229-237.
- Somers DC, Dale AM, Seiffert AE, Tootell RBH. 1999. Functional MRI reveals spatially specific attentional modulation in human primary visual cortex. *Proc Natl Acad Sci U S A.* 96:1663-1668.
- Spratling MW. 2008. Reconciling predictive coding and biased competition models of cortical function. *Front Comput Neurosci.* 2:4.
- Spratling MW. 2010. Predictive coding as a model of response properties in cortical area V1. *J Neurosci.* 30(9):3531-3543.
- Stephan KE, Penny WD, Daunizeau J, Moran RJ, Friston KJ. 2009. Bayesian model selection for group studies. *Neuroimage.* 49:1004-1017.
- Summerfield C, Egner T. 2009. Expectation (and attention) in visual cognition. *Trends Cogn Sci.* 13(9):403-409.
- Todorovic A, Van Ede F, Maris E, De Lange FP. 2011. Prior expectation mediates neural adaptation to repeated sounds in the auditory cortex: an MEG study. *J Neurosci.* 31(25):9118-9123.
- Treue S, Martínez Trujillo JC. 1999. Feature-based attention influences motion processing gain in macaque visual cortex. *Nature.* 399(6736):575-579.
- Yantis S, Schwarzbach J, Serences JT, Carlson RL, Steinmetz MA, Pekar JJ, Courtney SM. 2002. Transient neural activity in human parietal cortex during spatial attention shifts. *Nat Neurosci.* 5(10):995-1002.
- Zhang W, Luck SJ. 2009. Feature-based attention modulates feedforward visual processing. *Nat Neurosci.* 12(1):24-25.