

Early effects of previous experience on conscious perception

Jaan Aru^{1,2,5,6}, Renate Rutiku^{2,3,5,†}, Michael Wibral⁴, Wolf Singer^{1,2} and Lucia Melloni^{1,2,*}

¹Ernst Strüngmann Institute (ESI) for Neuroscience in Cooperation with Max Planck Society, Deutschordenstrasse 46, 60528 Frankfurt am Main, Germany; ²Department of Neurophysiology, Max Planck Institute for Brain Research, Max-von-Laue-Strasse 4, 60438 Frankfurt am Main, Germany; ³Institute of Psychology, University of Tartu, Näituse 2, Tartu 50409, Estonia; ⁴MEG Unit, Brain Imaging Center, Goethe University, Heinrich Hoffmann Strasse 10, Frankfurt am Main 60528, Germany and Present address: ⁵Institute of Public Law, University of Tartu, Tallinn 10119, Estonia. ⁶Institute of Computer Science, University of Tartu, Tartu 50409, Estonia.

*Correspondence address. Department of Neurophysiology, Max Planck Institute for Brain Research, Max-von-Laue-Str. 4, 60438 Frankfurt am Main, Germany. E-mail: lucia.melloni@brain.mpg.de

†These authors contributed equally to this work.

Abstract

Constructive theories of brain function such as predictive coding posit that prior knowledge affects our experience of the world quickly and directly. However, it is yet unknown how swiftly prior knowledge impacts the neural processes giving rise to conscious experience. Here we used an experimental paradigm where prior knowledge augmented perception and measured the timing of this effect with magnetoencephalography (MEG). By correlating the perceptual benefits of prior knowledge with the MEG activity, we found that prior knowledge took effect in the time-window 80–95 ms after stimulus onset, thus reflecting an early influence on conscious perception. The sources of this effect were localized to occipital and posterior parietal regions. These results are in line with the predictive coding framework.

Key words: predictive coding; contents of consciousness; MEG; prior knowledge; conscious perception, cognitive penetrability

Introduction

Prior knowledge massively influences the way we experience the world (e.g. Kersten *et al.*, 2004). For instance, objectively gray bananas are perceived as yellowish due to our previous experience with those fruits (Hansen *et al.*, 2006). Recent functional resonance imaging (fMRI) studies have demonstrated that this effect is correlated with activity in the early visual cortex (Bannert and Bartels, 2013; Vandenbroucke *et al.*, 2014) suggesting that prior knowledge may affect conscious perception directly, facilitating the processes leading up to object-level perception instead of acting on post-perceptual processes. Due

to the limited temporal resolution of fMRI, however, it is unknown whether prior knowledge affects conscious perception quickly or later on in the processing stream most likely reflecting decisional processes (Henderson and Hollingworth, 1999; Firestone and Scholl, 2015).

The question whether prior knowledge influences conscious perception early during stimulus processing is central to theories of perception such as predictive coding. Predictive coding is a powerful framework for explaining how sensory processing unfolds in the hierarchical cortical networks (Rao and Ballard, 1999; Friston, 2005; Summerfield and De Lange, 2014). In particular, it

Received: 6 August 2015; Revised: 13 March 2015. Accepted: 8 March 2016

© The Author 2016. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

posits that the main function of the cortex is to actively predict sensory inputs using internal or generative models (Friston, 2005) and to optimize predictions to minimize sensory prediction errors i.e. the mismatch between priors and sensory evidence. Although predictive coding is an attractive theory for explaining neural processing, sensory responses, and even conscious experience (Clark, 2013; Hohwy, 2013; Seth, 2015), fundamental cornerstones of this theory are yet to be explored. For example, if what we perceive is indeed the outcome of a process where pre-existing priors are compared to sensory data (Clark, 2013; Hohwy, 2013; Seth, 2015), prior knowledge may affect conscious perception already early on. Critically, the size of this effect should depend on how precisely prior knowledge predicts or explains incoming sensory evidence (Friston, 2005; Hohwy, 2013).

Recently, several electrophysiological studies have found that prior knowledge affects neural processes already before 100 ms (Chaumon et al., 2008; Chaumon et al., 2009; Gamond et al., 2011), suggesting that prior knowledge can impact perception rather early. Thus, previous studies have provided compelling evidence for early effects of expectations on neural measures. However, as behavioral measures, i.e., perceptual effects, were not investigated in those studies it is unclear whether the neural changes observed as a function of prior knowledge directly affect perception, and in particular conscious perception.

Only a handful of studies have investigated the time course of the effects of prior knowledge on perception. Specifically, Ghuman et al. (2008) demonstrated that early neural synchrony around 230 ms correlates with reaction time benefit of repeated presentation of visual objects. However, this effect may take place after the emergence of conscious experience, which according to some accounts occurs ~200 ms post-stimulus (e.g. Bachmann, 2000). Moreover, Ghuman et al. (2008) used reaction times to assess the facilitation by prior knowledge; however, reaction times are not a direct measure of perceptual effects as they can be dissociated from conscious perception (e.g. Bachmann, 2000). Recently, Sohoglu et al. (2012) described that the effect of prior knowledge on the perceived clarity of speech manifests itself on early (90–130 ms) as well as on several late evoked components. However, only a later component (450–700 ms) directly correlated with subjective clarity. Taken together, previous studies have not demonstrated that prior knowledge has a fast direct effect on conscious perception.

Here, we investigated the time course of the effects of prior knowledge on conscious perception by presenting targets close to the threshold of conscious perception and by asking subjects to report whether they perceived the target on a trial-by-trial basis. Our task and behavioral measures thus aimed at tapping directly into perceptual experience. To investigate the specificity of the effects of prior knowledge, we contrasted the sensory benefit stemming from prior knowledge with those stemming from sensory evidence by manipulating both factors independently in a single experimental paradigm while concurrently acquiring neural activity by means of magnetoencephalography (MEG). To study the effects specific to prior knowledge we correlated the perceptual benefits of prior knowledge with the MEG activity. Predictive coding posits that priors i.e., the availability of prior knowledge, that matches sensory information should lead to weaker neural responses as the latter can be “explained away” by the priors (Friston, 2005; Murray et al., 2004; Summerfield and De Lange, 2014). If those premises hold then a negative correlation between the perceptual benefits of prior knowledge and the neural activity is expected. In line with this prediction, we found a negative correlation which is apparent early in time, between 80 and 95 ms, and that localizes to

occipital and posterior parietal regions. These findings indicate that prior knowledge alters conscious perception early in time in line with theories such as predictive coding that postulate direct effects of priors on sensory processing.

Methods

Subjects

We recorded MEG from 26 subjects (9 males, 17 females). The data of two male participants were excluded due to measurement problems and extensive blink artifacts. The age of the remaining 24 subjects ranged from 21 to 28 years (mean age 24.4 years, standard deviation 2.4 years). All subjects were right-handed, had normal or corrected-to-normal vision, and no self-reported history of neurological or psychiatric disorders. The study was conducted in accordance with the Declaration of Helsinki. Prior to the participation in the study, all subjects gave written informed consent and received 15 Euros per hour for their participation in the study.

Stimuli and procedure

Stimuli and procedure are similar to those used in Aru et al. (2012a). In brief, stimuli consisted of 276 gray-scale images, containing a single person in the foreground as well as diverse backgrounds. In all, 54 catch images were included which had a similar background but no person in the foreground. The visibility of the images was parametrically manipulated by adding random Gaussian noise (Fig. 1A) while keeping the contrast constant for all stimulus degradation levels. Stimuli were edited with custom code using Matlab (R2008b, The MathWorks). The noise level values that yielded decreased visibility were chosen based on our previous study (Aru et al., 2012a) and ranged from 60% to 90% noise in 5% steps. Stimuli were displayed at the center of the screen, spanning 6×4.5 degrees of visual angle in the horizontal and vertical plane, respectively, and were surrounded by a gray background. Stimuli were presented on a translucent screen at a distance of 53 cm onto which they were projected via two front-silvered mirrors inside the MEG room from a liquid crystal display projector (60 Hz refresh rate) located outside the magnetically shielded MEG room.

Each subject completed a threshold experiment to obtain individualized noise levels that were later on used for the main experiment. We briefly (150 ms) presented 60 degraded images in randomized order at two different degradation levels (e.g. 70% and 80% noise). Those values were iteratively adjusted until obtaining an individual threshold yielding recognition performance around 70% accuracy in a discrimination task (male/female judgment) for two neighboring degradation levels (e.g. 80% noise and 75% noise level). To avoid any effect of familiarization, new gray-scale pictures were subsequently used in the main experiment.

The main experiment consisted of 27 experimental blocks. Each block comprised two phases: a familiarization phase to establish prior knowledge of a subset of the images, and a test phase (Fig. 1B). During the familiarization phase, four images without noise were presented twice for 3 seconds each. Subjects were asked to commit those pictures to memory. To assure attention to and encoding of the images, during the first presentation, subjects indicated via a button press first, the gender of the person on the picture (male/female task) and subsequently, guessed their age (older or younger than 30 years). Pictures were presented for a second time without an explicit task and subjects were encouraged to freely explore and memorize them.

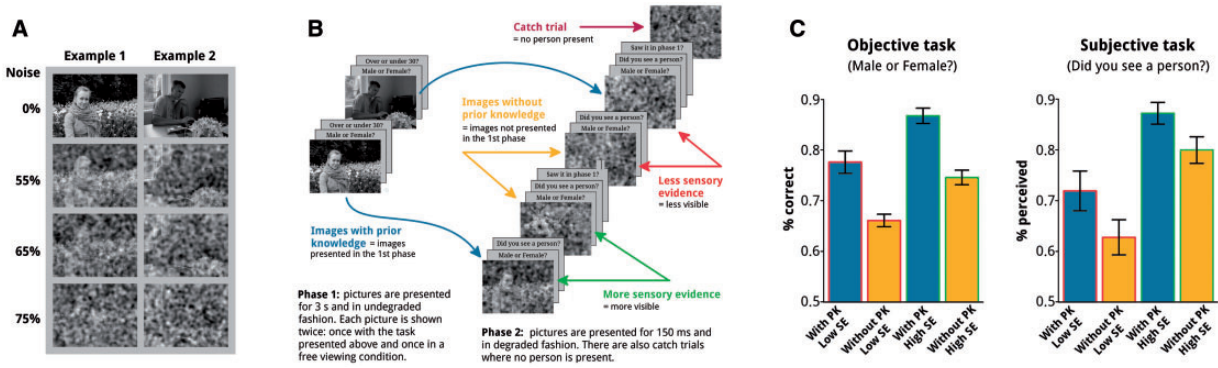


Figure 1. Experimental paradigm and behavioral results. (A) Example images used in the experiment, with different levels of noise. (B) Each block consisted of two phases: in the first phase, half of the images are familiarized. In the second phase, images are degraded and shown briefly. Images from phase 1 are presented together with new images (manipulation of prior knowledge). Familiar and unfamiliar images are also shown at two different degradation levels (i.e. high and low degradation; manipulation of sensory evidence). On a few trials an image without a person is presented (catch trials). Each image is followed by an objective and a subjective question. On some trials a third question (Was this picture presented in phase 1?) is also asked. (C) Behavioral results. Both prior knowledge and sensory evidence enhance perception. Effects are shown for both the objective discrimination data and the subjective reports about visibility.

In the test phase, a total of 20 degraded images were briefly presented (150 ms). Four stimuli only contained background and served as catch trials (see below), the other 16 images contained a person. The visibility of these 16 images depended on two orthogonal factors: the degradation level of the images (4 images presented at high and 4 images presented at low noise), effectively controlling sensory evidence, and the availability of a previous memory trace, which reflects prior knowledge. Prior knowledge was manipulated by presenting either familiarized (4) or new (4) images. Each picture was presented twice. On a given trial, subjects had to indicate first the gender of the person in the picture (male/female judgment, objective task) and subsequently report whether they had indeed perceived a person on the picture (subjective task). To assess the reliability of the subjects' judgment, specifically for the subjective task, we included four catch trials only containing background on the higher degradation level. Subjective reports were indeed reliable as evidenced by the low amount of false alarms in the catch trials (mean = 8.5%, SD = 8.5%). On half of the trials, subjects were asked to indicate whether the picture, now shown in the degraded fashion, had been presented during the familiarization phase (accuracy for this task: mean = 70%, SD = 12%). A block lasted 3–4 min and subjects could take breaks between blocks. A total of 108 trials per condition were presented. Stimuli were presented and responses recorded using Presentation v13 (Neurobehavioral Systems, USA).

To rule out any picture-specific effects in the neural measures of the conditions of interest (i.e. sensory evidence and prior knowledge), pictures containing a person were randomized such that across 4 subjects each picture was assigned to every condition (2 sensory evidence \times 2 prior knowledge) exactly 1 time.

The behavioral data were analyzed with a 2×2 repeated measures ANOVA with factors sensory evidence (more vs. less sensory evidence) and prior knowledge (familiarized vs. not familiarized). The dependent measure was either the proportion of seen person responses in the subjective task or the percentage of correct responses in the male/female task.

Data acquisition

MEG data were acquired with a 275-channel whole-head system (Omega 2005, VSM MedTech Ltd., BC, Canada) at a sampling rate

of 1200 Hz with a hardware antialiasing filter at 300 Hz in a synthetic third-order axial gradiometer configuration (Data Acquisition Software Version 5.4.0, VSM MedTech Ltd., BC, Canada). Head movements were limited using foam pads. Head position was measured before and after each run (i.e. 3 experimental blocks, approximately every 10 min) using three coils placed at the subject's nasion and preauricular points to make sure that the subjects' heads did not drift more than 5 mm from the original position at the beginning of the recording. Runs in which head movements exceeded 5 mm were excluded from further analysis (1 run from 2 subjects and 2 runs from 2 other subjects). We monitored eye movements and blinks during the recordings with two pairs of electrooculogram (EOG) electrodes, one pair placed vertically with one electrode above and other below the left eye, and the other pair placed horizontally, one electrode 1 cm lateral from the outer canthus of the left eye, the other 1 cm lateral from the outer canthus of the right eye. Behavioral responses were recorded using in-house fiber optic light barriers.

Individual high-resolution structural MRIs were acquired on a 3 Tesla Magnetom Allegra scanner (Siemens, Erlangen, Germany) using a T1-weighted magnetization-prepared rapid acquisition gradient-echo (MPRAGE) sequence (160 slices; TR: 2300 ms; TE: 3.93 ms; Flip Angle: 12°; Field of View: 256 mm; voxel size $1 \times 1 \times 1$ mm). To facilitate the alignment of the MEG and MRI data, the position of the nasion and the preauricular points were marked with vitamin E capsules during the MRI scan.

Event-related field analysis

MEG data were analyzed using the FieldTrip MATLAB Toolbox (Oostenveld et al., 2011) and custom code. The continuous data were first bandpass filtered between 0.1 and 30 Hz with a 4th order zero phase Butterworth filter, and subsequently segmented into trials from -300 ms to 700 ms relative to the onset of the target images. EOG recordings were visually inspected for eye movements and blinks and all trials contaminated by artifacts were discarded from further analysis. Trials containing SQUID jumps were also discarded. The remaining, artifact-free trials, were averaged according to the four experimental conditions

and baseline corrected over a 100-ms window prior to the stimulus onset.

Global field power (GFP), a measure of overall neural response strength (Murray et al., 2008), was used to quantify the event-related field (ERF) effects. The GFP analysis was chosen to reduce the dimensionality of the sensor-level data in order to answer the specific question about the timing of the effects of prior knowledge. GFP is equivalent to the spatial standard deviation of the magnetic field and is calculated as the square root of the mean of the squared value recorded at each sensor. GFP allows investigation of the response strength differences between conditions without a priori selection of electrodes. Before the statistical analysis, GFP values were down-sampled to 200 Hz to increase statistical power (by reducing the amount of statistical tests) while not sacrificing the temporal resolution (i.e. 200 Hz = 5 ms resolution).

In order to investigate the timing at which prior knowledge exerts an effect on conscious perception, we correlated the behavioral responses, i.e., *perceptual gain* due to prior knowledge (difference in the proportion of “seen” responses between conditions with and without prior knowledge) with the neural responses, i.e., *neural gain* (GFP differences for conditions with and without prior knowledge) across subjects. Spearman correlation was performed timepoint-by-timepoint from 50 to 500 ms, and corrected for multiple comparisons using False Discovery Rate (FDR) at $q = 0.05$ (Benjamini and Hochberg, 1995). Perceptual and neural gain were assessed from the same sets of trials (i.e. only the clean trials after artifact rejection).

In addition, we investigated amplitude differences across time (50–500 ms) in the GFP through a nonparametric cluster based two-way repeated measures ANOVA (Helbling et al., in preparation) with factors prior knowledge and sensory evidence. This approach allows for an equivalent test on the neural and the behavioral data. Statistical significance was determined with the cluster randomization method computed over 50–500 ms of the GFP (Maris and Oostenveld, 2007). The empirical distribution under the null hypothesis was obtained by 10 000 random permutations of the data. Single time points were considered significant and included into their respective cluster if they exceeded a P-value of 0.05. Clusters with a $P < 0.05$ and lasting longer than 10 ms were considered significant. To shed light on the sensors that contributed to the significant GFP differences observed in the ANOVA, we averaged over the significant time windows (335–350 ms, 400–425 ms, and 430–460 ms, highlighted areas on Fig. 2B) and ran a t-test per sensor between the condition with and without prior knowledge. Cluster permutation test (with 10 000 random permutations and a cluster entry threshold of $P < 0.05$) was used to correct for multiple comparisons across sensors (Fig. 2C). On Fig. 2D we show the sensors that had significant correlation with the perceptual gain. As no sensors survive correction for multiple comparisons, we plotted for illustrative purposes the sensors at the uncorrected $P < 0.05$ threshold.

Source reconstruction

Source activity was reconstructed with the MATLAB package SPM8 (<http://www.fil.ion.ucl.ac.uk/spm/>; version 10 May 2010). First, individual forward models were created using a 8196 vertex cortical template mesh by scaling SPM’s generic boundary element (BEM) model to the individual subject’s head, based on their MRIs. Second, source reconstruction was performed using the Multiple Sparse Priors (MSP) algorithm in the group inversion mode (Friston et al., 2008). This entails separate source

reconstruction for each experimental condition via a group inversion step where the condition-specific ERFs of all subjects are inverted together to ensure consistency over the individual inverse models. Note that while the time window for this analysis was based on the GFP results (80–95 ms), the data for source reconstruction were based on the ERFs.

The peak coordinates of significant sources were localized with the Automated Anatomical Labeling (AAL) map and the Brodmann Areas image as provided by MRICron (<http://www.mccauslandcenter.sc.edu/mricron/>). The locations were additionally confirmed with SumsDB (<http://sumsdb.wustl.edu:8081/sums/searchload.do?dispatch=celldata>).

Results

Behavioral results

As shown in Fig. 1C, subjects’ conscious perception of the person in the images increased as a function of sensory evidence and the availability of prior knowledge. A two-way repeated measures ANOVA confirmed significant main effects of sensory evidence ($F(1,23) = 69.109$, $P = 2.206 \text{E-}08$) and prior knowledge ($F(1,23) = 70.167$, $P = 1.931 \text{E-}08$) on subjective visibility. The interaction between sensory evidence and prior knowledge was not significant ($F(1,23) < 1.0$). A comparable pattern of results was found for the objective discrimination task (male/female task): significant main effects of sensory evidence ($F(1,23) = 116.71$, $P < 1.746 \text{E-}10$) and prior knowledge ($F(1,23) = 77.812$, $P < 7.71 \text{E-}09$), with no interaction between the two factors ($F(1,23) < 1.0$). These results are in line with previous studies showing that prior knowledge can boost conscious perception (e.g. Aru et al., 2012a; Melloni et al., 2011; Mayer et al., 2015), and with theories such as predictive coding which postulate that conscious perception is determined both by sensory evidence and priors (e.g. Hohwy, 2013; Clark, 2013; Seth, 2015).

Correlating perceptual and neural gain

While prior information significantly enhances conscious perception, the extent of the benefit of prior knowledge on perception varies across subjects: some participants profit more than others. We capitalized on these interindividual differences to determine the earliest time at which the neural responses are affected by the availability of prior information. We computed a measure of *perceptual gain*, i.e., the difference in visibility between conditions with and without prior knowledge, and correlated that with a measure of *neural gain*, i.e., differences in the GFP of the MEG signal between the conditions with and without prior knowledge. If prior knowledge affects conscious perception, a significant correlation (across subjects) between the neural gain and the perceptual gain is expected. Importantly, the time point of this significant correlation provides a measure of how early prior knowledge directly affects conscious perception. The analysis idea can be seen as an extension of the interindividual differences approach exploited in fMRI (for review see Kanai and Rees, 2011) in which interindividual differences are correlated over space (i.e. different brain areas) to the time dimension. We observed a significant correlation between the neural gain in GFP and the perceptual gain in subjective perception due to prior knowledge (Fig. 2A) in an early time window, before 100 ms (uncorrected $P = 0.0007$, FDR corrected $P < 0.05$ over 80–95 ms; Fig. 2B).

As shown in Fig. 2A, the correlation between the perceptual and neural gain is negative, indicating that the greater the

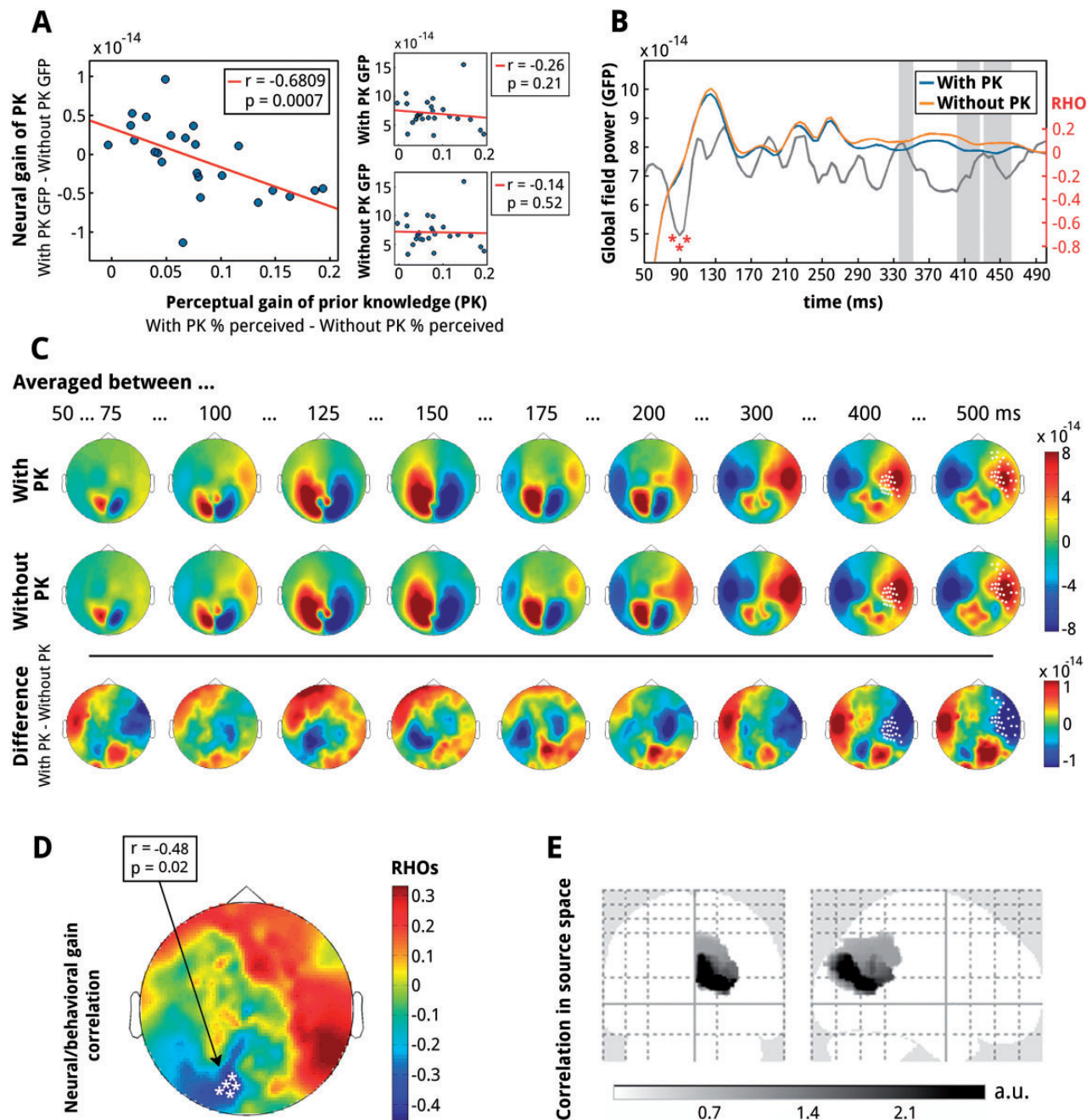


Figure 2. The correlation between neural gain and perceptual gain of prior knowledge (PK). (A) On the left: negative Spearman correlation between the neural gain (GFP with PK – GFP without PK) and perceptual gain (proportion of “seen” responses with PK – proportion of “seen” responses without PK), across the group of subjects in the time interval 80–95 ms after picture onset. Top panel on the right: correlation between the perceptual gain and the GFP from the condition with PK only; bottom panel on the right: correlation between the perceptual gain and the GFP for the condition without PK. (B) The GFP traces for trials with PK and without PK over the investigated time window (50–500 ms). Gray shaded areas depict significant main effect of PK from a respective ANOVA. The gray line depicts rho values of the Spearman correlation between the neural gain and perceptual gain of prior knowledge over the whole assessed time window (50–500 ms). The scale of rho values is on the right. The time points (80–95 ms) where there was a significant correlation between the neural gain and the perceptual gain of prior knowledge (panel A) are marked with a red asterisk. (C) Topographies over time of the MEG responses for conditions with PK and without PK and the topography of their differences (bottom row). Asterisks depict sensors that showed significant differences between the conditions with and without PK in the time windows where their GFPs were different (panel B). (D) Topography of the correlation between neural ERF gain and perceptual gain of PK. Sensors with a significant correlation ($P < 0.05$ uncorrected) are highlighted with white asterisks. (E) Neural sources underlying the temporal correlation observed in the GFP analysis. Plots show the localization of source activity differences between trials with and without prior knowledge in the occipital and parietal lobe which were negatively correlated with the perceptual gain of prior knowledge in the time interval 80–95 ms after picture onset (peak MNI coordinate at 10, –76, 18).

difference between the subject's GFP response to the pictures without and with prior knowledge, the bigger the perceptual effect of prior knowledge on conscious perception. This result fits well with the predictive coding framework, where top-down information suppresses expected sensory input (Friston, 2005; Rao and Ballard, 1999). The availability of precise top-down predictions should lead to weaker sensory responses to pictures with prior knowledge while also being accompanied by more efficient perceptual processing of these pictures, exactly as observed in our experiment.

Previous analysis investigated the correlation between the perceptual gain and the neural gain of prior knowledge, and thus capitalized on a difference. However, it could be argued that the observed effects are not related to the *difference* between the GFPs evoked by pictures with and without prior knowledge, but rather reflect the GFPs evoked by *either* the pictures with or without prior knowledge. Yet, when we correlated the perceptual effect of the subjects with the GFPs to either the pictures with or without prior knowledge, we did not observe a significant correlation in the respective time window (Fig. 2A on the right). Thus, the observed correlation between global brain responses and the effect of prior knowledge on conscious perception is specific to the *difference* between responses to pictures with and without prior knowledge.

The early neural effects are specific to prior knowledge, as a corresponding analysis based on the gain stemming from sensory evidence failed to reveal a significant effect. Also, the effect appears specific to subjective perception, as we observed no correlation between the gain in accuracy in the objective discrimination task and the GFP difference of pictures with and without prior knowledge (all $P > 0.1$).

Main effects of prior knowledge and degradation

The correlation approach adopted above links perception more directly to the neural measure than a direct contrast between trials with vs. without prior knowledge as a direct contrast does not only reflect perceptual effects but also additional cognitive process, e.g., memory or novelty (Aru et al., 2012b). Nevertheless, the correlation results neither exclude the possibility of an interaction between prior knowledge and sensory evidence nor do they provide information about the main effects alone.

Thus, in order to assess the GFP main effects and any possible interactions a two-way repeated measures ANOVA with factors sensory evidence and prior knowledge was conducted with a nonparametric cluster-based approach. The main effect of prior knowledge encompasses three short-lived clusters in close proximity (335–350 ms, 400–425 ms, and 430–460 ms, highlighted areas on Fig. 2B). Figure 2C shows the topographies across time for the conditions with and without prior knowledge, as well as the difference. In the time windows where we observed significant GFP differences for prior knowledge (Fig. 2B) we highlighted, for illustrative purposes, the sensors that showed significant differences between the conditions with and without prior knowledge (Fig. 2C). Similarly, the main effect of sensory evidence includes an early short-lived cluster between 180 ms and 195 ms and a late prominent cluster extending from 305 ms to 500 ms. Most importantly, however, there was no time epoch with an interaction between the two factors on GFP. Thus, prior knowledge and sensory evidence seem to not only influence behavior independently, but also their neural implementation appears to differ.

Positive values in the neural gain

An interesting aspect of the data presented thus far is that roughly half of the subjects show positive values in the neural gain (Fig. 2A), indicating higher responses for the condition with prior knowledge than for the condition without prior knowledge. If matching predictions should “explain away” sensory responses and thus weaken neural responses, then positive values in this subtraction seem odd. Why would trials with prior knowledge lead to stronger neural responses than trials without prior knowledge for some subjects? Importantly, according to the predictive coding theory, “explaining away” is only one side of the coin: the other, complementary process is sending top-down predictions which lead to “explaining away” sensory responses. In other words, prior knowledge should lead to suppression of predicted neural responses but at the same time to enhanced activity of the units that provide these predictions (Murray et al., 2002; Friston, 2005; Feldman and Friston, 2010; Egner, 2010; de Gardelle et al., 2012, 2013). It is then possible that the combined activity of these two types of processes leads to stronger activity in the condition where predictions are only partially formed. In other words, having imprecise predictions might lead to stronger neural responses than having no predictions at all (see Discussion for a detailed treatment of this issue and for another potential explanation).

It is notoriously difficult to separate these two processes—predictions and prediction errors (e.g. De Gardelle et al., 2012)—especially in a mixed signal like that of the MEG. However, the conjecture that subjects with positive values in the neural gain have imprecise predictions in the condition with prior knowledge can be tested on the basis of behavioral data: If some subjects formed imprecise priors during the familiarization phase, then these subjects should also perform worse on the recognition memory task (see Methods). In that task, subjects were asked to indicate whether the picture, now shown in the degraded fashion, had been presented during the familiarization phase. Indeed, when we correlated the neural GFP gain (the same as in Fig. 2A) with the accuracy in the recognition memory task, we observed a strong negative correlation (Fig. 3A). Subjects with more positive responses in the neural gain (Fig. 2A) performed worse in this recognition memory task, i.e. they remembered less well which images had been shown during the familiarization phase (correlation between recognition memory and neural gain, $r = -0.64$, $P = 0.001$). We also analyzed the relationship between perceptual gain and recognition memory. We predicted that subjects with smaller perceptual gain due to prior knowledge—thus more imprecise priors—should have worse recognition performance. When we correlated perceptual gain with recognition accuracy we observed a significant positive correlation ($r = 0.61$, $P = 0.002$) (Fig. 3B), indicating that subjects who were less able to recognize which degraded images had been presented during the familiarization phase benefited less from having prior knowledge. Those same subjects are the ones for whom we observed a more positive neural gain, as confirmed by a median split analysis over the perceptual gain of prior knowledge (Fig. 3C). When we further investigated those groups based on a median split over the perceptual gain, we observed that subjects with smaller perceptual gain were especially limited in recognizing images that were previously familiarized, as shown by an interaction between prior knowledge and recognition accuracy (interaction $F = 5.81$, $P < 0.05$) (Fig. 3D). While both groups have a similar recognition performance for images that were never familiarized (without PK condition), stronger differences in recognition performance

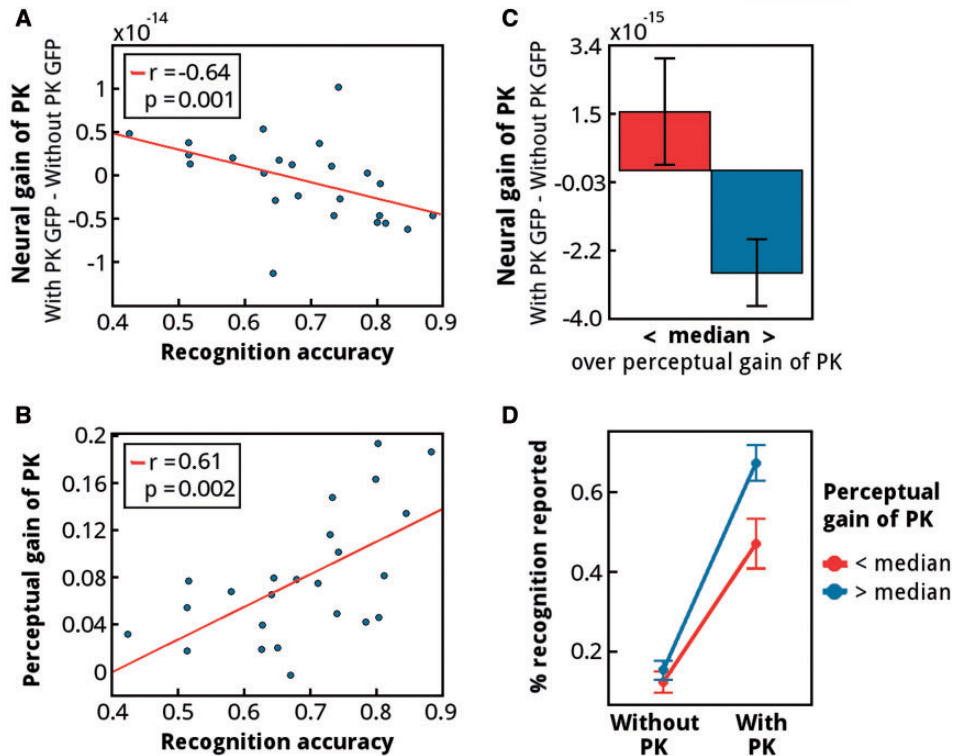


Figure 3. The relationship between perceptual gain and performance in the recognition task. (A) Negative correlation between the neural GFP gain (from Fig. 2A) and recognition accuracy in the time interval 80–95 ms after picture onset. Stronger neural suppression is observed for subjects who more accurately recognized the familiarized images. (B) Positive correlation between behavioral gain of prior knowledge and recognition accuracy. (C) Median split based on perceptual gain of prior knowledge: low perceptual gain of prior knowledge is associated with positive neural GFP gain values, while high perceptual gain is associated with negative neural GFP gain. (D) Median split based on perceptual gain of prior knowledge. Trials are further separated into those with and without prior knowledge. Subjects with low perceptual gain of prior knowledge were especially limited in recognizing familiarized images.

are observed for those images that were previously familiarized (with PK condition, post-hoc t -test: $P < 0.05$). Altogether, the analysis of the behavioral data supports the idea that the positive neural gain observed in a subset of subjects occurs in those subjects who performed worse on the recognition memory task, i.e., who had formed less precise priors during the familiarization phase (see Discussion).

Neural sources explaining the correlation between neural and perceptual gain

Finally, we investigated the neural sources underlying the observed correlation between the perceptual effect of prior knowledge and the measured differences in GFP responses. To this end, we focused on the sensor level data and aimed to reconstruct their underlying sources through a Multiple Sparse Priors (MSP) algorithm (Friston *et al.*, 2008) during the time window in which we observed a significant correlation, i.e. 80–95 ms post stimulus onset. Analogously to the previous analysis, we investigated neural sources where the activity difference between trials with and without prior knowledge (i.e. neural gain) was negatively correlated with the perceptual gain of prior knowledge. Figure 2D shows the topography of the sensor-level correlation, i.e. same data as on Fig. 2A only on the sensor level demonstrating that occipital sensors contribute to the correlation between the neural and perceptual gain of prior knowledge. On the level of neural sources (Fig. 2E), we found a significant cluster ($P = 0.006$, cluster level FDR corrected) spanning occipital

and posterior parietal regions (peak in the right cuneus/Brodmann area 17 at MNI coordinate 10, -76 , 18).

Discussion

The aim of the present study was to investigate *when* during the recognition process prior knowledge affects conscious perception. Answering this question is important for theories like predictive coding which postulate that prior knowledge affects conscious perception swiftly and directly. To investigate this issue, we made use of the interindividual variability of *perceptual effects* of prior knowledge to track the timing of the respective neural processes. We observed that prior knowledge boosts subjective visibility, and that this perceptual gain of prior knowledge is correlated with the neural gain early in time (80–95 ms post-stimulus). This effect was specific to subjective visibility, as there was no correlation with objective discrimination performance. Also, it was specific to prior knowledge as no comparable correlation was observed when investigating the effects of sensory evidence on conscious perception. Thus, prior knowledge exerts an early effect on conscious perception. The neural sources of this effect are localized in occipital and posterior parietal regions.

Predictive coding and the effects of prior knowledge

Our work concurs with and extends previous behavioral and fMRI studies showing that prior knowledge affects conscious

perception. For example, Hansen et al. (2006) demonstrated that participants perceived objects that were presented without any color in their usual color. Recent fMRI studies of this perceptual effect have indicated that previous knowledge about the color of the objects can be decoded from early visual areas (Bannert and Bartels, 2013; Vandenbroucke et al., 2014). Similarly, Schwiedrzik et al. (2007) have demonstrated that having a prediction about the motion path enhances detection performance of a moving stimulus. The neural correlates of this effect were also found to be in the early visual cortex (Alink et al., 2010). Finally, in an elegant study Van Loon et al. (2015) demonstrated that prior knowledge about Mooney images enhances their decodability from early visual areas (see also Hsieh et al., 2010). Interestingly, they observed that this beneficial effect of prior knowledge is disrupted when subjects receive Ketamine, implicating NMDA receptors in this form of visual plasticity, similar to what has been described for auditory predictions in the mismatch negativity literature (Javitt et al., 1996). However, none of these previous fMRI studies could unequivocally demonstrate that previous knowledge indeed affects conscious perception early after stimulus onset as for that time-resolved methods are necessary.

Electrophysiological studies have shown that prior knowledge affects neural processes already before 100 ms (e.g. Chaumon et al., 2008, 2009; Gamond et al., 2011; Todorovic and Lange, 2012). However, these studies provided no direct evidence that this early neural effect of prior knowledge has an early and enhancing effect on conscious perception as the perceptual effect of prior knowledge on conscious perception was not assessed in those studies. Hence, the effects reported in those studies may have reflected access to memory traces, novelty responses or other conceivable, general effects of prior knowledge.

Taken together, on the one hand some previous studies have clearly established that prior knowledge affects perception behaviorally and on the level of fMRI. On the other hand, other separate studies have documented early effects of prior knowledge with electrophysiological recordings. However, to our knowledge none of these prior studies assessed when exactly prior knowledge affects perception. Answering this question is relevant for the ongoing debate about whether prior knowledge influences perception directly (e.g. Firestone and Scholl, 2015; Lupyan, 2015). Our results now show that priors affect conscious perception early in time and at early processing stages, as suggested by the predictive coding theory. Note that the results do not imply that these early effects are part of the neural processes that underlie conscious experience of the target per se. Rather, it is likely that the present results reflect prerequisites of conscious perception (Aru et al., 2012b), i.e. neural processes that contribute to conscious perception but do not directly reflect it. Recent research by Pinto et al. (2015) also supports the present conclusion that predictions about stimuli affect their entry into conscious perception.

Predictive coding straightforwardly explains why the size of the perceptual gain of prior knowledge and the respective neural gain were negatively correlated across subjects in our study (Fig. 2A). By familiarizing half of the pictures, priors for these pictures and their contents are created. As the familiarization phase was rather short, there are interindividual differences in the quality (precision) of these priors. If the prior is adequate, it will help to resolve ambiguity and to perceive the contents of the otherwise degraded stimulus (e.g. Melloni et al., 2011). On the other hand, having a corresponding prior will also match the sensory input and help to “explain it away,” thus eliciting

smaller neural responses (Friston, 2005). In contrast, a mismatch between prior knowledge and sensory input will be signaled by higher prediction errors which in turn help to adjust the priors and learn about the world (e.g. Friston 2005; Hohwy, 2013; Kim et al., 2014). It has then been shown that predictable input leads to weaker neural responses already at early stages of visual processing (Alink et al., 2010; Kok et al., 2012). Thus, reduced responses to pictures with prior knowledge reflect the precision of the prediction and thus the quality of the prior that was established during familiarization. Hence, the weaker the response to pictures with prior knowledge, the better the prior that was formed during the familiarization phase. And the better the prior, the more it can support perception. Considering this, it is to be expected that those subjects who have weaker response to pictures with prior knowledge as compared to pictures without prior knowledge, benefit more perceptually. In contrast, subjects who benefited less perceptually might have less precise priors to support perception. In our data, those subjects also exhibited positive values in the neural gain (Fig. 2A). A finding that at first sight might seem at odd with the tenant of predictive coding that priors “explain away” sensory responses and therefore should weaken neural responses. However, according to the predictive coding theory, weakening of sensory responses is just one aspect of the inferential process, which is carried out by shutting down prediction error units (PE units). Another key feature is the establishment and maintenance of predictions, which in turn lead to suppression of PE units. Neurally this is implemented by so-called prediction units (P units, also called representational units). Thus, while the activity of the PE units is diminished as a function of the precision of the priors, the P units increase their responses (Friston, 2005; Feldman and Friston, 2010). Thus, theoretically two distinct classes of signals i.e., prediction error signals and prediction signals should contribute to the neural response, a tenant that has been established in fMRI (Egner, 2010; de Gardelle et al., 2012, 2013). The same mixture of responses is likely to be at place for magnetoencephalographic signals. This insight can then explain why some subjects show positive and some subjects show negative effects. We suggest that GFP captures the combined activity of P and PE units, resulting in stronger activity when subjects have imprecise predictions as compared to when they have no predictions at all. Specifically, establishing a precise and correct prior should lead to activity of P units and a reduction of activity in the PE units, amounting to smaller net responses after prior knowledge has become available (as compared to the condition without prior knowledge). In contrast, forming an imprecise prior could lead to activity in the P units but also to stronger activity in the PE units (as priors are not effectively explaining away sensory responses), resulting in a net increase in activity after prior knowledge has become available (as compared to the condition without prior knowledge). The analysis of the behavioral data supports the conjecture that subjects who formed less precise priors during the familiarization phase, as measured by the worse performance on the recognition memory task, were also the subset of subjects for whom we observed positive neural gain. An alternative possibility is that in the brains of those subjects several different priors are explored when explaining the sensory responses. Here, more priors are activated until the one that best explains the sensory input is found. The search for a matching prior is associated with the activity of more prediction units, leading to stronger neural responses in the condition with prior knowledge.

Given the challenges of separating the activity of predictions and prediction errors on mass neural signals (e.g. De Gardelle

et al., 2012, 2013), further experiments specifically investigating how priors are implemented while distinguishing the contribution of priors and prediction errors to the neural signals will be needed to provide a more definite test of those alternatives. Recent developments in the analysis of fMRI signals have made it possible to separate repetition suppression and repetition enhancement (De Gardelle et al., 2012, 2013), and may be applicable to MEG in future studies.

Differences between subjective and objective measures of behavior

Although it is tempting to think that subjective experience and objective performance should always go hand-in-hand, research over the last decades has shown that these two aspects of perception can indeed dissociate. It is well known that above chance performance can be achieved in tasks where subjects do not consciously perceive the stimuli, as observed e.g. in blind-sight and in subliminal priming (e.g. reviewed in Kouider and Dehaene, 2007). These findings indicate that objective performance can change without concomitant changes in subjective experience. The opposite pattern has also been observed, i.e. changes in subjective experience without concomitant changes in objective performance (e.g. Lau and Passingham, 2006; Schwiedrzik et al., 2011). Thus, objective performance and conscious experience can vary independently of each other and are most likely supported by different neural pathways (Lau and Passingham, 2006; Schwiedrzik et al., 2011). Our result, that an early negative correlation with the neural gain of prior knowledge is found only for subjective experience but not for objective performance, adds to this list of growing evidence for a dissociation of subjective experience from objective performance. In particular, our results suggest that prior knowledge can facilitate conscious experience by affecting neural processes early in time in the early visual cortex and that this facilitation is specific for conscious experience as objective performance does not benefit from it.

Effects of prestimulus activity and attention

Priors about the world could be activated already in the baseline time window during the anticipation of the upcoming stimuli (e.g. Hesselmann et al., 2010; Vetter et al., 2015; Mayer et al., 2015). Importantly, our present results cannot be explained by differences in baseline activity between trials with and without prior knowledge as the order of trials was randomized (i.e. in the baseline subjects could not know whether the upcoming image was in the familiarized image set or not). Nevertheless, our early effects of prior knowledge could indicate that these predictions are activated in the prestimulus time. In the case of images with prior knowledge, these predictions would be quickly matched with the incoming sensory evidence (Mayer et al., 2015; Myers et al., 2015). Importantly, the better the match, the smaller the evoked sensory responses, as postulated by predictive coding.

Studies on attention demonstrate enhanced responses to attended stimuli. Could the present results be explained by differential attentional processing? In our study, it is unlikely that stimuli with and without prior knowledge differed in their attentional loads. This is because there was no physical difference between images with prior knowledge vs. without prior knowledge, beyond the fact that half of those were previously familiarized. Hence, the images are unlikely to systematically differ in how they capture attention. Furthermore, trials with and

without prior knowledge were randomly intermixed; thus, subjects could not anticipate which trial type would be presented. Thus, any effect of attention should be similar across conditions. Furthermore, studies have systematically shown that predictions and attention can be dissociated and have differing effects: predictions suppress neural responses while attention enhances them (reviewed in Summerfield and Egner, 2009; Summerfield and De Lange, 2014). This effect has been captured by current computational models implementing predictive coding, e.g. by Feldman and Friston (2010) and by Spratling (2008). Those models assume that attention and predictions affect neural responses through different mechanisms. In particular, attention modulates the gain while predictions suppress prediction errors. The present results fit better with the role of predictions, as the subjects who showed the strongest benefit from prior knowledge also showed the weakest sensory responses (i.e. strongest suppression).

Conclusion

Our behavioral results demonstrate that prior knowledge has a beneficial effect on conscious perception. A time-resolved correlation analysis revealed that this effect of prior knowledge on conscious perception takes place early—between 80 ms and 95 ms after stimulus onset—thus unequivocally indicating that prior knowledge has a direct and rapid impact on conscious perception. These results are in line with the theory of predictive coding.

Acknowledgements

We are indebted to Caspar M. Schwiedrzik and two anonymous reviewers for insightful comments on the manuscript. This work was supported by the Max Planck Society, a Marie Curie International Outgoing Fellowship of the European Community's Seventh Framework Programme under project number [299372] (L.M.); and the LOEWE Neuronale Koordination Forschungsschwerpunkt Frankfurt, NeFF (L.M.). The authors also acknowledge the support of PUT438 (J.A.) and IUT20-40 (J.A. and R.R.) of the Estonian Research Council. The authors declare no competing financial interests. Data are available upon request.

References

- Alink, A, Schwiedrzik, CM, Kohler, A et al. Stimulus predictability reduces responses in primary visual cortex. *J Neurosci* 2010;30: 2960–66.
- Aru, J, Axmacher, N, Do Lam, AT et al. Local category-specific gamma band responses in the visual cortex do not reflect conscious perception. *J Neurosci* 2012a; 32: 14909–14.
- Aru, J, Bachmann, T et al. Distilling the neural correlates of consciousness. *Neurosci Biobehav Rev* 2012b; 36: 737–46.
- Bachmann, T. *Microgenetic Approach to the Conscious Mind. Advances in Consciousness Research*, Vol. 25.: Amsterdam: John Benjamins Publishing, 2000.
- Bannert, MM, Bartels, A. Decoding the yellow of a gray banana. *Curr Biol* 2013; 23: 2268–72.
- Benjamini, Y, Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B* 1995; 57:289–300.

- Chaumon, M, Drouet, V, Tallon-Baudry, C. Unconscious associative memory affects visual processing before 100 ms. *J Vis* 2008; 8: 10.
- Chaumon, M, Schwartz, D, Tallon-Baudry, C. Unconscious learning versus visual perception: dissociable roles for gamma oscillations revealed in MEG. *J Cogn Neurosci* 2009; 21: 2287–99.
- Clark, A. Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav Brain Sci* 2013; 36: 181–204.
- De Gardelle, V, Waszczuk, M, Egnér, T et al. Concurrent repetition enhancement and suppression responses in extrastriate visual cortex. *Cerebral Cortex* 2012; bhs211.
- De Gardelle, V, Stokes, M, Johnen, VM et al. Overlapping multi-voxel patterns for two levels of visual expectation. *Front Hum Neurosci* 2013; 7.
- Egnér, T, Monti, JM, Summerfield, C. Expectation and surprise determine neural population responses in the ventral visual stream. *The Journal of Neuroscience* 2010; 30(49), 16601–08.
- Feldman, H, Friston, KJ. Attention, uncertainty, and free-energy. *Front Hum Neurosci* 2010; 4.
- Firestone, C, Scholl, BJ. Cognition does not affect perception: evaluating the evidence for “top-down” effects. *Behav Brain Sci* Available on CJO2015 doi:10.1017/S0140525X15000965.
- Friston, K. A theory of cortical responses. *Philos Trans R Soc B Biol Sci* 2005; 360: 815–36.
- Friston, K, Harrison, L, Daunizeau, J et al. Multiple sparse priors for the M/EEG inverse problem. *NeuroImage* 2008; 39: 1104–20.
- Gamond, L, George, N, Lemarechal, JD et al. Early influence of prior experience on face perception. *Neuroimage* 2011; 54: 1415–26.
- Ghuman, AS, Bar, M, Dobbins, IG et al. The effects of priming on frontal-temporal communication. *Proc Natl Acad Sci* 2008; 105: 8405–9.
- Hansen, T, Olkkonen, M, Walter, S et al. Memory modulates color appearance. *Nature Neuroscience* 2006; 9: 1367–68.
- Henderson, JM, Hollingworth, A. High-level scene perception. *Ann Rev Psychol* 1999; 50: 243–71.
- Hohwy, J. *The Predictive Mind*. Oxford: Oxford University Press, 2013.
- Hsieh, PJ, Vul, E, Kanwisher, N. Recognition alters the spatial pattern of fMRI activation in early retinotopic cortex. *J Neurophysiol* 2010; 103: 1501–7.
- Javitt, DC, Steinschneider, M, Schroeder, CE et al. Role of cortical N-methyl-D-aspartate receptors in auditory sensory memory and mismatch negativity generation: implications for schizophrenia. *Proc Natl Acad Sci* 1996; 93: 11962–67.
- Kanai, R, Rees, G. The structural basis of inter-individual differences in human behaviour and cognition. *Nat Rev Neurosci* 2011; 12: 231–42.
- Kersten, D, Mamassian, P, Yuille, A. Object perception as Bayesian inference. *Ann Rev Psychol* 2004; 55: 271–304.
- Kim, G, Lewis-Peacock, JA, Norman, KA et al. Pruning of memories by context-based prediction error. *Proc Natl Acad Sci* 2014; 111: 8997–9002.
- Kok, P, Jehee, JF, de Lange, FP. Less is more: expectation sharpens representations in the primary visual cortex. *Neuron*, 2012; 75: 265–70.
- Kouider, S, Dehaene, S. Levels of processing during non-conscious perception: a critical review of visual masking. *Philos Trans R Soc Lond B Biol Sci* 2007; 362: 857–75.
- Lau, HC, Passingham, RE. Relative blindsight in normal observers and the neural correlate of visual consciousness. *Proc Natl Acad Sci USA* 2006; 103: 18763–68.
- Lupyan, G. Cognitive penetrability of perception in the age of prediction: predictive systems are penetrable systems. *Rev Philos Psychol* 2015; 6: 547–69.
- Maris, E, Oostenveld, R. Nonparametric statistical testing of EEG- and MEG-data. *J Neurosci Methods* 2007; 164: 177–90.
- Mayer, A, Schwiedrzik, CM, Wibral, M et al. Expecting to see a letter: alpha oscillations as carriers of top-down sensory predictions. *Cerebral Cortex* 2015; bhv146.
- Melloni, L, Schwiedrzik, CM, Müller, N et al. Expectations change the signatures and timing of electrophysiological correlates of perceptual awareness. *J Neurosci* 2011; 31: 1386–96.
- Murray, MM, Brunet, D, Michel, CM. Topographic ERP analyses: a step-by-step tutorial review. *Brain Topograp* 2008; 20: 249–64.
- Murray, SO, Kersten, D, Olshausen, BA et al. Shape perception reduces activity in human primary visual cortex. *Proc Natl Acad Sci* 2002; 99: 15164–69.
- Murray, SO, Schrater, P, Kersten, D. Perceptual grouping and the interactions between visual cortical areas. *Neural Networks* 2004; 17: 695–705.
- Myers, NE, Rohenkohl, G, Wyart, V et al. Testing sensory evidence against mnemonic templates. *eLife* 2015; e09000.
- Oostenveld, R, Fries, P, Maris, E et al. FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput Intell Neurosci* 2011; 2011: 156869.
- Pinto, Y, van Gaal, S, de Lange, FP et al. Expectations accelerate entry of visual stimuli into awareness. *J Vis* 2015; 15: 13.
- Rao, RP, Ballard, DH. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat Neurosci* 1999; 2: 79–87.
- Schwiedrzik, CM, Alink, A, Kohler, A et al. A spatio-temporal interaction on the apparent motion trace. *Vis Res* 2007; 47: 3424–33.
- Schwiedrzik, CM, Singer, W, Melloni, L. Subjective and objective learning effects dissociate in space and in time. *Proc Natl Acad Sci USA* 2011; 108: 4506–11.
- Seth, AK. Presence, objecthood, and the phenomenology of predictive perception. *Cogn Neurosci* 2015; 1–7.
- Sohoglu, E, Peelle, JE, Carlyon, RP et al. Predictive top-down integration of prior knowledge during speech perception. *J Neurosci* 2012; 32: 8443–53.
- Spratling, MW. Reconciling predictive coding and biased competition models of cortical function. *Front Comput Neurosci* 2008; 2.
- Summerfield, C, Egnér, T. Expectation (and attention) in visual cognition. *Trends Cogn Sci* 2009; 13: 403–9.
- Summerfield, C, de Lange, FP. Expectation in perceptual decision making: neural and computational mechanisms. *Nat Rev Neurosci* 2014; 15: 745–756.
- Todorovic, A, de Lange, FP. Repetition suppression and expectation suppression are dissociable in time in early auditory evoked fields. *J Neurosci* 2012; 32: 13389–95.
- Vandenbroucke, ARE, Fahrenfort, JJ, Meuwese, JDI et al. Prior knowledge about objects determines neural color representation in human visual cortex. *Cerebral Cortex* 2014, bhu224.
- van Loon, AM, Fahrenfort, JJ, van der Velde, B et al. NMDA Receptor Antagonist Ketamine Distorts Object Recognition by Reducing Feedback to Early Visual Cortex. *Cerebral Cortex* 2015; bhv018.
- Vetter, P, Grosbras, MH, Muckli, L. TMS over V5 disrupts motion prediction. *Cerebral Cortex* 2015; 25: 1052–59.