Priming of visual cortex by temporal attention? The effects of temporal predictability on stimulus(-specific) processing in early visual cortical areas

Rico Fischer*, Franziska Plessow, Hannes Ruge

Technische Universität Dresden, Germany

A R T I C L E   I N F O

Article history:
Accepted 30 October 2012
Available online 8 November 2012

Keywords:
Temporal predictability
Temporal attention
Alerting signal
Accessory

A B S T R A C T

In recent studies it has been shown that temporal predictability of expected events alters processing in perception and action. Yet, the neural mechanism(s) by which temporal predictability biases this processing is as of yet little understood. Therefore, in the present fMRI study we investigated how temporal predictability affects neural processing in visual cortical areas. For this, thirty-four participants either categorized the gender or the movement direction of vertically or horizontally moving faces in different blocks of trials. Temporal predictability of stimulus onset was manipulated by the presence or absence of an auditory alerting signal validly predicting stimulus onset. The behavioral data revealed a clear performance benefit for the presence of an alerting signal. Neuroimaging results showed that irrespective of the currently performed task temporal predictability significantly reduced activation in the primary visual cortex. This activation reduction correlated with the alerting signal-related performance benefit. Furthermore, we did not find a selective influence of increased temporal predictability on target-specific visual processing (faces or movement) in the respective material-specific visual brain areas. Together, these findings suggest an increased task-unspecific readiness by the alerting signal that might result in more efficient transmission of stimulus codes into response codes.

Introduction

The successful deployment of temporal attention to the onset of an expected event is a major aspect in action control and the prerequisite for efficient interaction with and in the environment. Predicting and therefore preparing for the onset of a future event reduces temporal uncertainty and allows for an optimal engagement in information processing (Niem and Näätänen, 1981). Thus, it is not surprising that manipulations of temporal predictability of expected events have been shown to bias perception and action (for an overview see Nobre et al., 2007). Much research has been conducted over the last decade to determine neural mechanisms responsible for allocation of attention in time (Coull, 2004; Coull et al., 2000, 2001; Fan et al., 2005; Hackley et al., 2009) and to detect similarities, for example, to allocation of attention in space (Coull and Nobre, 1998). Yet, given the importance of the ability to anticipate the onset of future events, it appears surprising that to date only little is known about the neural mechanisms by which temporal predictability might bias perception and action (cf. Nobre, 2010).

The deployment of attention (in time or space) is generally thought to support the selection of task-relevant features and thus enables selective attention. Previous research in this field has predominantly been dedicated to studying the effects of spatial attention on stimulus processing. The conclusion that spatial attention improves perceptual analysis can be derived from the finding that contrast perception (i.e., responding to the orientation of contrast gratings) is enhanced at the attended spatial location (Carrasco et al., 2004). Early influential work demonstrated modulating effects of spatial attention on neural activity in extrastriate neurons (area V4) in non-human primates (Moran and Desimone, 1985).

More recent evidence of selective spatial attention gating visual processing in human participants was provided by an fMRI study in which target-specific enhancement of neural activity in visual cortical areas was related to the deployment of visual spatial attention (Hopfinger et al., 2000). The selective allocation of visual attention toward a location in space includes top-down attentional control in terms of attentional disengagement and voluntary orienting. In a spatial cuing paradigm, the cue triggered the allocation of attention to the indicated location. Isolating cue-related neural activity, the authors were able to identify a network for voluntary attentional control. In addition, they also showed the consequences of allocating spatial attention. In particular, cue-related attentional control biased activity in those visual cortical areas that were later dedicated to process the visual target. In other words, the deployment of spatial attention resulted in selective sensory processing of relevant visual targets (Hopfinger et al., 2000).

To date, it is completely unclear to which extend such findings related to spatial attention can be generalized to the field of temporal attention, e.g., whether temporal predictability can also trigger selective...
sensory processing. It is agreed upon that temporal predictability facilitates information processing and thus optimizes behavior by speeding up performance. While traditional views propose a late locus of effects of temporal predictability at motor stages (Coull and Nobre, 1998; Kiesel and Miller, 2007; Mattes et al., 1997; Minussi et al., 1999; Tandonnet et al., 2006), recent research has been accumulated demonstrating effects of temporal predictability on early perceptual and pre-motor stages of information processing (Correa et al., 2005, 2006; Hackley et al., 2007; Lange et al., 2003; Martens and Johnson, 2005; Rolke and Hofmann, 2007).

On the assumption that temporal predictability modulates perceptual processing, few studies have conducted research investigating the neural mechanisms of perceptual biases by temporal predictability. Anderson and Sheinberg (2008), for example, could show that spiking responses of neurons in the inferior temporal (IT) cortex of monkeys were modulated by manipulations of temporal predictability. In a temporal orienting task, each picture of a number of objects served as cue (indicating an early or late target onset with 80% validity) or as target (indicating the required response). For early target onset trials, a valid cue was associated with more spiking than an invalid cue. That is, the early spiking response decreased when the stimulus onset was not expected (invalid cue).

In human participants, changes in the temporal predictability of stimulus onset were found to result in altered neural processing in the primary visual cortex (V1, Alink et al., 2010; see also Bueti et al., 2010). Alink and colleagues, for example, argued that stimuli with a predictable onset require less neural activation to transmit its information from lower to higher cortices (Rao and Ballard, 1999). Accordingly, in an fMRI study Alink et al. (2010) showed that neural responses in V1 were smaller when the stimulus onset was temporally predictable than when it was unpredictable by participants. Although this demonstration of predictable stimuli being processed with less neural activation in visual cortical areas is an intriguing result, the predictability of stimulus onset was provided by means of the trajectory of surrounding illusory motion. It therefore remains an open question as to whether these findings can be generalized to and replicated with more conventional methods of reducing temporal uncertainty in the environment, such as explicit temporal cues predicting the stimulus onset. This is especially important since other fMRI studies demonstrated increased activations of the visual cortex areas such as the right occipital pole (BA 18) that were associated with exogenous shifts of temporal attention (Coull et al., 2000).

Although the research outlined above has provided important opening work on understanding the neural underpinnings of the allocation of attention to predict upcoming stimulus events, a number of important questions remain. For example, as aforementioned, can the effects of temporal predictability on early visual cortical areas as reported by Alink et al. (2010) be generalized to different kinds of temporal predictability? How selective is the sensory priming of neural activity in visual areas? Especially the latter question appears captivating. It has been argued, for example, that the selection of task-relevant information is guided by the expected temporal onset of events (Correa, 2010). Does the allocation of temporal attention to the onset of the stimulus event result in selective sensory processing of relevant visual targets analogous to the allocation of spatial attention (Hopfinger et al., 2000)? Indeed, spatial and temporal attention share many (neural) commonalities (Coull and Nobre, 1998) and it has been argued that orienting attention in time results in virtually the same changes in neural activity as observed for orienting attention in space suggesting that both types of allocating attention are neurophysiologically similar (cf. Anderson and Sheinberg, 2010). Therefore, while previous studies suggest a “ubiquitous system for allocating attentional resources in general” (Coull and Nobre, 1998, p.7431), it is yet to be tested whether the deployment of temporal attention to stimulus onset (temporal predictability) leads to the same consequences as the deployment of spatial attention — namely, the selective (biased) sensory processing of relevant visual targets.

The present study

The present study aims at providing further information whether and how temporal predictability of the upcoming stimulus onset affects neural processing in visual cortical areas. First, we aim at extending the findings by Alink et al. (2010) by investigating the generality of the observed reduction in neural activity of V1 neurons in response to temporal predictability. Alink and colleagues explained their finding by means of feedback signals from higher motion-related areas as temporal onset predictability in their paradigm was closely linked with motion perception. Previous studies mostly applied temporal visual cues to manipulate temporal predictability (e.g., Coull, 2004; Coull and Nobre, 1998; Fan et al., 2005), which however, may confound neural activity of target stimulus processing in V1. Therefore, we will test whether patterns of reduced neuronal activity in V1 can similarly be found when task-irrelevant acoustic alerting signals contain temporal information about stimulus onset.

Our second aim relates to assumptions in the domain of spatial attention proposed by Hopfinger et al. (2000). Here we follow the idea that, if not only the allocation of spatial attention but also the deployment of temporal attention results in selective sensory processing of relevant visual targets, it should be possible to obtain modulations of neural activity in stimulus-material specific neural areas in accordance with manipulations of temporal attention.

In the present study, participants were presented with pictures of male and female faces that could move either in a left-right dimension (horizontally) or in an up-down dimension (vertically). In one block of trials participants had to respond to movement direction (horizontal versus vertical) whereas in the other block of trials they had to discriminate gender (female versus male). We particularly implemented a choice–reaction task instead of often used stimulus-detection tasks, because it was suggested that effects of temporal predictability on sensory processing are more likely to be detected when a perceptual analysis of the visual stimulus features is required (cf. Correa, 2010). Temporal uncertainty was realized by large variations in the response–stimulus interval (RSI) between consecutive trials. In this context of unpredictable stimulus onset, an acoustic signal validly predicted the temporal stimulus onset in half of the trials by appearing always at a constant 250 ms prior to stimulus onset — a forperiod interval known for optimum preparation for stimulus onset (Fischer et al., 2007; Gottsdanker, 1980) and for revealing performance benefits in terms of speeded responses in a variety of choice-reaction tasks (Fischer et al., 2010, 2012). We particularly chose this manipulation of temporal predictability because 1) the presence versus absence of an alerting signal is known to determine temporal predictability (Bernstein et al., 1973) and 2) this approach allowed differentiating between neural activity of the auditory temporal cue (alerting signal) and the neural activity of visual target stimulus processing in V1. This distinguishes the current approach from more classical manipulations of temporal predictability (e.g., Coull, 2004; Coull and Nobre, 1998; Fan et al., 2005), in which the neural activity of visual temporal cue may be difficult to dissociate from neural activity of the target stimulus.

Therefore, within the present paradigm we expect alerting signal-based response speeding in the behavioral data for both, face- and motion-categorization task, alike. We also expect the alerting signal to activate the temporal orienting or alerting network that comprises a number of specified cortical areas associated with the allocation and directing of temporal attention to an expected point in time (Coull et al., 2000, 2001; Fan et al., 2005; Hackley et al., 2009; Thiel et al., 2004). Typical areas include the left intra-parietal cortex and pre-supplementary motor area (pre-SMA) possibly supporting attentional allocation towards an expected event and superior temporal
As a first hypothesis, we expect a stimulus-materia specific effect of alerting signal presentation in early visual areas, such as reduced neural activity in V1 neurons as reflected in reduced BOLD activation V1 or other early downstream visual areas (Alink et al., 2010). Second, if temporal predictability results in selective material-specific sensory processing in cortical areas, manipulations of temporal predictability by alerting signal presence/absence should lead to task specific neural modulations. Naturally, irrespective of the currently implemented task, we expect moving face stimuli to elicit activations within visual brain areas in both face-processing (e.g., fusiform face area, FFA) and movement-processing (e.g., medio-temporal cortex, MT). Crucially, however, in the face-categorization task we assume focused temporal attention to specifically alter activations in areas associated with task-relevant stimulus features (i.e., face-processing areas). We expect the opposite in the motion-categorization task. Here, temporal predictability is thought to modulate activations in movement-sensitive areas.

Method

Participants

34 students of the Technische Universität Dresden (11 male, 18–30 years; mean age 23.0 years) participated in the study. They gave their written informed consent prior to their inclusion in the study in accordance with the Declaration of Helsinki and the guidelines of the local ethics committee. All participants had normal or corrected-to-normal vision.

Stimuli and apparatus

Stimuli consisted of 14 male and 14 female familiar faces, that were taken from the stimulus set previously used by Engst, Martín-Loeches, and Sommer (2006).1 Face pictures were presented in black and white in a dark red frame (176 pixels × 213 pixels) to improve discrimination from the black background. Visual stimuli were projected to participants via Visuastim digital goggles (Resonance Technology, Inc., Northridge, BC, USA) simulating a viewing distance of 100 cm. The alerting signal was presented binaurally at 700 Hz (approximately 50 dB) via headphones (NordicNeuroLab). A fiber-optic, light-sensitive key press was used to record participants’ behavioral responses, which were carried out with the index and middle finger of the right hand, respectively. Stimulus presentation and data recording were controlled by Presentation 12.1 software (Neurobehavorial Systems, Inc., Albany, CA, USA) running on a Windows XP personal computer.

Procedure

Experimental procedure

In separate blocks of trials, participants responded to one of two stimulus features, that is, faces (i.e., male versus female) or direction of movement (vertical versus horizontal). Half of the participants started with the face-categorization task, in which they were required to categorize moving faces as either male or female. Subsequently, the same participants performed the movement direction task, judging the vertical or horizontal movement of the face stimuli. The other half of the participants received the reversed task order. Task order was counterbalanced across participants. Response mapping was kept constant between participants. That is, participants responded with the left key to male faces and to horizontal movements and with the right key to female faces and to vertical movements.

Each trial began with the presentation of a central fixation sign (plus sign) for a variable interval between 1 and 5 s (i.e., 1, 1.5, 2, 2.5, 3, 3.5, 4, and 5 s, respectively) after which the target stimulus was presented centrally for 80 ms. Following target offset, the same target was presented for 80 ms repeatedly either slightly above or below the central position (x=0; y=−60/60 pixels), creating an up- and downward movement. For a horizontal movement, the target was presented for 80 ms repeatedly slightly to the left or right of fixation (x=−60/60; y=0 pixels). Following the response or a maximum of 1840 ms after target onset, the fixation sign was presented. In case of a missing or erroneous response, the word “Fehler” (wrong) was presented for 300 ms as error feedback. In half of the trials, an alerting signal preceded the central target onset by 250 ms for 80 ms. Trials with an alerting signal were presented randomly throughout an experimental block. Participants were told that this acoustic signal is completely task-irrelevant and could be ignored.

The experimental session consisted of three experimental blocks. The scanner paused between blocks. In the first two blocks, the main experiment was implemented, either starting with the face-categorization task or with the movement-categorization task. Both, face- and motion-categorization task, started with eight trials of practice to familiarize with the response mapping. Each task consisted of 224 trials, in which the complete combination of 28 (faces) × 2 (movement) × 2 (alerting signal) was presented twice. The third block was always the functional localizer, comprising 17 stimulation blocks of 20 s duration, separated by 16 fixation blocks of 15 s duration. There were four types of stimulation blocks including (1) moving faces, (2) stationary faces, (3) moving scrambled faces, and (4) stationary scrambled faces.

Imaging procedure

Whole-brain images were acquired on a Siemens 3 T whole-body Trio System (Erlangen, Germany) with a 16 channel circularly polarized head coil. Headphones dampened scanner noise and enabled the presentation of the auditory alerting signal. Both, structural and functional images, were acquired for each participant. High-resolution structural images (1.0 mm × 1.0 mm × 1.0 mm) were acquired using an MP-RAGE T1-weighted sequence (TR=1900 ms, TE=2.26 ms, TI=900 ms, flip=9°). Functional images were acquired using a gradient echo planar sequence (TR=2000 ms, TE=30 ms, flip=80°, interleaved slice acquisition, slice gap=0). Each volume contained 26, 5.0 mm thick slices (in-plane resolution 4.0 mm × 4.0 mm).

Data analysis

Behavioral data

Erroneous responses2 (5.0 %) and RTs below 150 ms and above 1000 ms (2.8 %) were excluded prior to RT analysis. Repeated measures ANOVAs with the factors alerting signal (AS on, AS off) and task (face, motion) were conducted on RTs and error rates.

Imaging data

Pre-processing. The empirical data set was analyzed with SPM5 running within MATLAB 7.8. Pre-processing included slice-time correction, rigid body movement correction (3 translation, 3 rotation parameters), normalization of the functional images by registering the mean functional image to the standard MNI EPI template image provided by SPM5 (the resulting interpolated spatial resolution was

1 Face stimuli consisted of a sample of national and international actors, celebrities, and politicians, generally known to German participants. However, since familiarity was not an issue in the present study, it was not individually assessed for each participant prior to participation.

2 Due to a technical problem, erroneous trials of thirteen participants could not be identified. Yet, the data of these participants were included in the analyses to increase statistical power. This procedure seems justified as the RT data pattern remains the same even when excluding the complete data set of those participants.
GLM analyses. The pre-processed imaging data were analyzed using the General Linear Model (GLM) approach as implemented in the SPM5 software package. Model regressors were created by convolving neural input functions for the different block types (localizer scans) or event types (main experimental scans) with the assumed canonical hemodynamic response function used by SPM5. A 1/128 Hz high-pass filter was used. Three separate GLMs were estimated, one for the functional localizer and two different GLMs for the main experimental task.

A first GLM was implemented for the functional localizer to determine the material-specific sensory/perceptual brain areas related to visual face processing and visual motor processing. Block-related regressors were created for (1) faces/motion blocks, (2) faces/no-motion blocks, (3) scrambled faces/motion blocks, (4) scrambled faces/no-motion blocks, and (5) fixation blocks. Two contrasts were computed, one comparing faces versus scrambled faces and one comparing motion versus no-motion. Brain areas specifically involved in face processing were determined via the faces versus scrambled faces contrast at p<.05 (cluster-size-corrected for whole-brain volume at thresholded at p<.001), exclusively masked by the motion versus no-motion contrast at a very lenient threshold of p<.05 (uncorrected). Thereby, only those voxels were included which were sensitive to faces and at the same time not significantly sensitive to motion. Similarly, brain areas specifically involved in motion processing were determined via the motion versus no-motion contrast at p<.05 (cluster-size-corrected for the whole-brain volume at an initial height threshold of p<.001), exclusively masked by the faces versus scrambled faces contrast at a very lenient threshold of p<.05 (uncorrected).

A second GLM was implemented for the analysis of the main experiment. Four event-related regressors were included for (1) motion-task trials with an alerting signal (AS on), (2) motion-task trials without an alerting signal (AS off), (3) face-task trials with an alerting signal (AS on), and (4) face-task trials without an alerting signal (AS off). Two contrasts were computed for each subject, including the main-effect contrast alerting signal, and the interaction contrast alerting signal X task. Additionally, to reveal a possible correlational relationship between BOLD activation and behavioral, the subject-specific mean RT differences of AS off–AS on, indicating the reduction of RT with AS on, were entered as a covariate for these contrasts at the group-analysis level. The AS on–AS off contrast was computed to identify the general “alerting network”, mainly for comparison with results from related previous studies. The reverse AS off – AS on contrast was computed to reveal the predicted reduction of activation within early visual brain areas. Activation maps were thresholded at p<.05 (cluster-size-corrected for the whole-brain volume at an initial height threshold of p<.001).

Furthermore, the more specific alerting signal X task interaction contrast was computed to test the second central hypothesis of the present experiment that the alerting signal would alter neural processing in material-specific brain areas depending on the task currently performed. Thus, when participants implemented the face-categorization task, we expected an effect of the alerting signal in face-sensitive areas as identified via the functional localizer. Vice versa, when participants implemented the motion-categorization task, we expected an effect of the alerting signal in motion-sensitive areas as identified via the functional localizer. The hypothesized task-specific modulation of BOLD activation within material-specific ROIs was examined in two different ways to make sure that true effects would not be missed due to overly conservative correction procedures. In a first analysis, we used group-level ROIs determined via the functional localizer. Within these group-level ROIs, we searched for voxels exhibiting a significant alerting signal X task interaction. Significant voxels were identified by using small volume correction and assessing significance according to (1) a corrected cluster-level threshold (p<.05) at different initial height thresholds of p<.001; p<.01; p<.05 and (2) voxel-level family-wise-error (FWE) threshold (p<.05). In a second analysis, ROIs were determined for each subject separately within the boundaries of the group-level ROIs (thresholded at p<.01, uncorrected) and defined by the respective localizer contrasts. Within these single-subject ROIs, we extracted two values for each cell of the alerting signal X task interaction, including (1) the mean beta values averaged across voxels within an 8 mm sphere centred on the peak-activation voxel of the respective localizer contrast within each subject-specific ROI, and (1) the mean beta values averaged across all voxels within each subject-specific ROI. These mean beta values were then assessed on the group level at p<.05; uncorrected for number of ROIs assessed. Finally, the alerting signal X task interaction was tested on the whole-brain level (corrected cluster-level threshold p<.05 at different initial height thresholds of p<.001; p<.01) to be able to identify significant voxels outside the pre-determined ROIs.

Results

Behavioral data

RTs

Results are presented in Fig. 1. Temporal predictability by means of an alerting signal revealed a strong influence on RTs, F(1, 33) = 247.76, p<.001, η² = .882. Responses were 83 ms faster when stimulus onset was predictable (AS on) than when it was non-predictable (AS off). RTs were also shorter responding to motion (563 ms) compared to responding to faces (583 ms), F(1, 33) = 5.53, p = .025, η² = .143. Both factors interacted, F(1, 33) = 12.51, p = .001, η² = .275. Alerting-signal effects on RTs were more pronounced in the motion (–94 ms) than in the face-categorization task (71 ms).

Error rates

An effect of temporal predictability was also observed in the error data with more accurate responses when the alerting signal was present than when it was absent, F(1, 20) = 31.29, p<.001, η² = .610. No other effects were significant, all p’s >.31.

Imaging data

Localizer

The result of the two localizer conditions are summarized in Tables 1 and 2 and visualized in Fig. 2. Specifically face-sensitive
activations included, as expected, the fusiform gyri (Kanwisher et al., 1997). Specifically motion-sensitive activations included, as expected, visual area MT within the posterior middle temporal gyrus (Born and Bradley, 2005).

**Main experiment**

Results are presented in Tables 3 and 4 and in Fig. 3. The AS on > AS off contrast was computed to identify the general “alerting network”. This contrast revealed strong activations in cortical areas including the middle and superior temporal lobes (MTL and STL) and the posterior intraparietal sulcus (pIPS), the precentral gyrus, and in the left middle lateral PFC (mid LPFC). Additionally, there were non-cortical regions activated including the Putamen and the Thalamus, as well as the Cerebellum. The reverse AS on < AS off contrast was computed to reveal the predicted reduction of activation within early visual brain areas. Indeed, as predicted this contrast revealed circumscribed activations within bilateral visual cortex. The only other significant activation was found in the anterior cingulate cortex.

Furthermore, activation within these visual areas correlated with RT in the way visualized in Fig. 4, showing that a stronger acceleration of RT with AS on relative to AS off is associated with a stronger suppression of BOLD activation in the occipital pole for AS on relative to AS off trials (see also Table 5).

The more specific alerting signal X task contrast was computed to test the second central hypothesis of the present experiment that the alerting signal would alter neural processing in material-specific brain areas depending on the task currently performed. As shown graphically in Fig. 2 (lower panel) for exemplary coordinates, there was no such effect irrespective of the analysis procedure used. The absence of such an effect was even confirmed when using an extremely lenient threshold at $p < 0.05$, uncorrected.

**Discussion**

In the present study, we aimed to demonstrate that focused temporal attention results in biased sensory processing in visual cortical areas. For this, we presented alerting signals to reduce temporal uncertainty by predicting the temporal onset of stimulus presentation. We hypothesized that temporal predictability by means of the alerting signal results in reduced neural processing in early visual areas (i.e., V1). Furthermore, we tested whether effects of temporal predictability are task-specific and thus associated with increased selectivity in stimulus-material specific sensory processing.

First of all, the present results revealed a clear performance benefit in terms of faster and more accurate responses when stimulus onset was predictable (AS on) than when it as unpredictable (AS off), representing the typical effect of alerting. This pattern was more pronounced in the motion-categorization task compared to the face-categorization task. The fact that both dependent measures (i.e., RT and error rates) benefited from the presence of an alerting signal excludes interpretations in terms of processing-criterion shifts that result in premature responding to the alerting signal at the cost of increased error rates (i.e., speed-accuracy trade off). This two-fold performance benefit on the behavioral level was accompanied by alerting signal-related effects on the level of brain activation. Not surprisingly, the alerting signal activated typical brain areas that have previously been described to be involved in similar experimental situations (Coull, 2004; Fan et al., 2005) including the middle temporal lobe (MTL)/superior temporal lobe (STL) and Thalamus linked to “alerting” processes, the posterior intraparietal sulcus (pIPS) linked to attentional “orienting” processes, and frontal areas related to executive control processes in alerting situations (Fan et al., 2005; Hackley et al., 2009; see Raz and Buhle, 2006 for an overview). Additionally, we also observed activations in the Basal Ganglia (e.g., Putamen) and in the Cerebellum, both of which have been suggested to be related to timing processes in action control (Dreher and Grafman, 2002; Ivry and Spencer, 2004).

In addition to alerting-signal related activations and in line with our first hypothesis, temporal predictability of stimulus onset significantly reduced neural activity in the primary visual cortex irrespective of the task-relevant stimulus material (face versus motion). We could show that the behavioral performance benefit induced by the alerting signal negatively correlated with the strength of the BOLD response in V1. Put differently, the larger the alerting-signal impact on behavioral performance the stronger the reduction of BOLD activation in the occipital pole. This pattern extends results recently reported in a study by Alink et al. (2010), in which participants predicted the onset of visual stimuli on the basis of the trajectory of surrounding illusory motion. These

---

**Table 1**

<table>
<thead>
<tr>
<th>Brain region</th>
<th>Number of voxels</th>
<th>Local minima</th>
<th>z-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>L fusiform face area</td>
<td>148</td>
<td>40 - 56 - 24</td>
<td>7.0</td>
</tr>
<tr>
<td>R fusiform face area</td>
<td>153</td>
<td>40 - 28 - 16</td>
<td>7.8</td>
</tr>
<tr>
<td>L (para-)hippocampus</td>
<td>38</td>
<td>28 - 12 - 24</td>
<td>4.7</td>
</tr>
<tr>
<td>R/L hippocampus</td>
<td>184</td>
<td>32 - 10 - 12</td>
<td>5.4</td>
</tr>
<tr>
<td>R posterior MTL/STL</td>
<td>38</td>
<td>52 - 12 - 12</td>
<td>4.8</td>
</tr>
<tr>
<td>R posterior IPS</td>
<td>23</td>
<td>28 - 56 - 48</td>
<td>4.2</td>
</tr>
<tr>
<td>L precentral</td>
<td>41</td>
<td>40 - 20 - 24</td>
<td>4.4</td>
</tr>
<tr>
<td>R precentral</td>
<td>131</td>
<td>44 - 64 - 48</td>
<td>4.4</td>
</tr>
<tr>
<td>L mid LPFC</td>
<td>34</td>
<td>48 - 20 - 24</td>
<td>4.5</td>
</tr>
<tr>
<td>R mid LPFC</td>
<td>44</td>
<td>44 - 24 - 20</td>
<td>4.2</td>
</tr>
<tr>
<td>L mid VLPCF/anterior insula</td>
<td>30</td>
<td>36 - 24 - 8</td>
<td>4.0</td>
</tr>
</tbody>
</table>

**Table 2**

<table>
<thead>
<tr>
<th>Brain region</th>
<th>Number of voxels</th>
<th>Local minima</th>
<th>z-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>L OCC</td>
<td>1892</td>
<td>-8 - 100 - 20</td>
<td>8.2</td>
</tr>
<tr>
<td>R OCC</td>
<td>16</td>
<td>-96 - 16 - 8</td>
<td>7.4</td>
</tr>
<tr>
<td>L cuneus</td>
<td>4</td>
<td>-84 - 16 - 8</td>
<td>7.4</td>
</tr>
<tr>
<td>R cuneus</td>
<td>8</td>
<td>-96 - 16 - 8</td>
<td>7.4</td>
</tr>
<tr>
<td>L mid STL</td>
<td>43</td>
<td>28 - 12 - 12</td>
<td>4.1</td>
</tr>
<tr>
<td>R mid STL</td>
<td>65</td>
<td>32 - 12 - 12</td>
<td>3.9</td>
</tr>
<tr>
<td>L posterior MTL/MT</td>
<td>33</td>
<td>60 - 12 - 12</td>
<td>4.8</td>
</tr>
<tr>
<td>L mid IPS</td>
<td>38</td>
<td>44 - 12 - 12</td>
<td>4.2</td>
</tr>
</tbody>
</table>

Note: IPS = intraparietal sulcus; L = left; MTL = middle temporal lobe; OCC = occipital cortex; R = right; STL = superior temporal lobe.
authors argued that temporal predictability of stimuli reduces responses in primary visual cortex, because less neural activation is required to transmit information from lower to higher visual cortical areas when stimuli are predictable than when they are unpredictable. Therefore, by anticipating forthcoming sensory stimuli, the brain is able to process expected stimuli with less neural “effort” (cf. Alink et al., 2010). Importantly, this “benefit” in visual processing is not bound to temporal predictions derived from processing of visual information but can be triggered also by auditory (alerting) signals inducing temporal predictability.

At present, however, it is not clear how temporal predictability results in reduced neural processing in early sensory areas. Striking similarities of reduced neural responses in early sensory areas, however, have been found in response to expected visual input. Here it is argued that knowing what visual input is likely to occur, helps reducing the computational demand of visual processing. The neural signature of visual expectation has been captured in so-called predictive coding models (e.g., Friston, 2005; Rao and Ballard, 1999), that postulate an interaction between two classes of neurons, representational

Table 3

<table>
<thead>
<tr>
<th>Brain region</th>
<th>Number of voxels</th>
<th>MNI coordinate</th>
<th>z-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>L MTL/STL</td>
<td>1031</td>
<td>-64 -32 12</td>
<td>6.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-52 -24 4</td>
<td>6.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-44 -16 -12</td>
<td>6.3</td>
</tr>
<tr>
<td>L posterior IPS</td>
<td></td>
<td>-32 -64 48</td>
<td>4.3</td>
</tr>
<tr>
<td>L Thalamus</td>
<td></td>
<td>-20 16 0</td>
<td>4.8</td>
</tr>
<tr>
<td>R MTL/STL</td>
<td>815</td>
<td>64 -36 8</td>
<td>6.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>60 -24 4</td>
<td>5.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>52 -12 -4</td>
<td>5.4</td>
</tr>
<tr>
<td>R Putamen</td>
<td>24</td>
<td>8 0</td>
<td>5.0</td>
</tr>
<tr>
<td>L precentral</td>
<td>243</td>
<td>-40 -4 60</td>
<td>4.7</td>
</tr>
<tr>
<td>L mid LPFC</td>
<td></td>
<td>-56 20 24</td>
<td>4.7</td>
</tr>
<tr>
<td>R precentral</td>
<td>26</td>
<td>40 4 52</td>
<td>3.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>48 4 56</td>
<td>3.4</td>
</tr>
<tr>
<td>L Cerebellum</td>
<td>31</td>
<td>-12 -72 -28</td>
<td>4.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-32 -60 -32</td>
<td>3.6</td>
</tr>
</tbody>
</table>

Note. IPS = intraparietal sulcus; L = left; LPFC = lateral prefrontal cortex; MTL = middle temporal lobe; R = right; STL = superior temporal lobe.

Therefore, by anticipating forthcoming sensory stimuli, the brain is able to process expected stimuli with less neural “effort” (cf. Alink et al., 2010). Importantly, this “benefit” in visual processing is not bound to temporal predictions derived from processing of visual information but can be triggered also by auditory (alerting) signals inducing temporal predictability.

At present, however, it is not clear how temporal predictability results in reduced neural processing in early sensory areas. Striking similarities of reduced neural responses in early sensory areas, however, have been found in response to expected visual input. Here it is argued that knowing what visual input is likely to occur, helps reducing the computational demand of visual processing. The neural signature of visual expectation has been captured in so-called predictive coding models (e.g., Friston, 2005; Rao and Ballard, 1999), that postulate an interaction between two classes of neurons, representational
neurons (encoding what is possible or likely as upcoming sensory input) and error neurons (representing the prediction error that results from the discrepancy between expected and actual sensory input). These two classes of interacting neurons iteratively reconcile expected and observed information within several hierarchical levels of visual processing by continuously reducing the prediction error (for a detailed review see Summerfield and Egner, 2009). In this framework, cortical responses are directly related to (visual) predictability. The activity of error neurons will be minimized when the stimulation is in accordance to expectation, but will be increased when sensory input conflicts with expectation. Thus, the more higher sensory regions are able to facilitate the interpretation of the naturally ambiguous perceptual input, the smaller the activity elicited by this perceptual input in lower cortical areas (cf. Summerfield and Egner, 2009). Although, predictive coding models are powerful to account for empirical findings of reduced neural activity to expected identity of visual input, currently it seems unclear in which extent they may also serve for explaining similar findings of the expected onset of the visual input (i.e., temporal predictability). It is conceivable that representational units not only code the expected identity but also the expected onset of an upcoming visual event, as different types of expectation often occur in combination (cf. Doherty et al., 2005). Knowing when something is likely to occur may help reducing perceptual noise providing a less ambiguous and thus, clearer perceptual signal in the first place. Of course future research will be needed to show in which respect predictive coding models may also account for findings of temporal predictability.

The present observation that early visual areas showed reduced activation after an acoustic alerting signal contrasts with Bueti et al. (2010) who report a positive relationship between temporal expectation and preparatory activation in early visual areas. The combination of two major procedural differences can explain this discrepancy. First, while Bueti et al. measured ongoing preparatory processes prior to the onset of the prepared perceptual operation (color change detection), our study measured the compound of ongoing preparation (alerting stimulus onset until target onset = 250 ms) and visual target processing including the potential impact of preparation on
target processing. Hence, the activation increase in Bueti et al. likely reflects the pure impact of ongoing preparation in early visual areas whereas the activation decrease in the present study might primarily reflect the facilitating impact of preparation on subsequent visual target processes (i.e., the accelerated visuo-motor transmission). Second, Bueti et al. correlated fluctuations in endogenously controlled temporal expectation with preparation-related BOLD activation across comparably long preparation periods in the range of seconds. In comparison, a potential similar BOLD activation increase associated with ongoing preparation in the present study was exogenously triggered and sustained over a much shorter preparation interval in the range of hundreds of milliseconds. This might explain why this potential preparation-related BOLD activation component was outweighed by the BOLD activation component associated with subsequent facilitation of target processing.

With respect to our second aim of our study, we did not find evidence that temporal predictability of stimulus onset resulted in selective stimulus-specific sensory processing. Evidence for selective sensory processing due to the deployment of attention was reported in the domain of spatial attention (Hopfinger et al., 2000). Therefore, we hypothesized that the temporal predictability might bias neural activity in those visual areas that are associated with task-relevant visual processing (i.e., faces versus motion). However, neural activity in face-processing and motion-processing areas was not at all affected by the presence or absence of the alerting signal. Of course, interpretations on the basis of null findings have to be handled with care. Nevertheless, we took a number of measures that render an explanation in terms of insufficient statistical power rather unlikely. First, we measured a comparably large number of 34 participants. Second, we ran several ROI-based analyses with increasingly lenient thresholding criteria to be sure that we did not miss even the weakest indication of possible task-related alerting effects. Even for participants-specific ROIs at voxelwise p < .05 (uncorrected), we were not able to demonstrate any effect.

Taken together, the present study revealed that task irrelevant acoustic signals expectedly showed neural activation pattern associated with temporal predictability and alerting (Coull, 2004; Fan et al., 2005). Furthermore, we could also demonstrate that the predictable onset of the imperative stimulus by the presence of an auditory alerting signal results in reduced neural processing in early visual areas (e.g., V1), suggesting visual processing with less neural “effort” (Alink et al., 2010) potentially resulting in a more efficient transmission of perceptual stimulus information into corresponding motor commands (e.g., facilitated stimulus–response translation processes, Fischer et al., 2010, 2012).

In the paradigm of the present study we incorporated large response–stimulus intervals to maximize temporal uncertainty about the expected stimulus onset. Despite the irrelevance of the alerting signals, their presence reflects hereby a valid predictor of the exact temporal stimulus onset and thus, resulted in temporal preparation to reduce temporal uncertainty (see also Müller-Gethmann et al., 2003; Niemi and Näätänen, 1981). Even though we highlighted the advantages of manipulating temporal predictability by means of the presence versus absence of an alerting signal in the introduction, it comes also at a cost that should be mentioned. For example, alerting signals have not only been associated with increased stimulus onset predictability, but also elicit brief increases in arousal, both of which are difficult to dissociate. Even though, in a recent study we found evidence for stimulus-specific effects of alerting signals that are unlikely the result of unspecific increases in arousal (Fischer et al., 2012), data of the current study cannot dissociate between these possibilities.

On the notion that effects of arousal and stimulus intensity are closely linked (Pribram and McGuinness, 1975), one could argue, for example, that phasic increases in arousal might modulate and affect the “energy” of the visual stimulus. In fact, according to the so-called “energy integration hypothesis”, the presence of an acoustic alerting signal would be comparable to increasing the intensity of the visual target stimulus (Bernstein et al., 1970). The seminal work by Stein et al. (1996) directly showed that auditory stimuli can increase the perceived intensity of additionally presented visual stimuli. At the same time, however, we suggest that an alerting-signal-based increase in the intensity of the visual input would rather predict an increased neural response of neurons in the primary visual cortex and not a decrease of neural activity in conditions of auditory and visual stimulation. Such reasoning, however, is

---

**Table 5**

<table>
<thead>
<tr>
<th>Brain region</th>
<th>Number of voxels</th>
<th>MNI coordinate (x, y, z)</th>
<th>z-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>L OCC</td>
<td>11 (p &lt; .034; corrected)</td>
<td>−16 −100 8</td>
<td>4</td>
</tr>
<tr>
<td>R OCC</td>
<td>12 (p &lt; .043; corrected)</td>
<td>20 −100 12</td>
<td>2.9</td>
</tr>
</tbody>
</table>

Note: L = left; OCC = occipital cortex; R = right.

---

**Fig. 4.** Correlations between AS-related BOLD suppression in the occipital poles and the AS-related benefit in response times (RT).
rather speculative and it remains to be shown to which extent effects of phasic arousal may be involved in the present findings. For this, subsequent studies are needed that can disentangle potential effects of temporal predictability and arousal (see e.g., Weinbach and Henik, 2012). Because of the elusive role of brief phasic arousal in the present design, at present we cannot exclude that, for the specific demonstration of selective (biased) sensory processing of relevant visual targets, other manipulations of temporal predictability (e.g., cued temporal orienting including non-informative neural cues, Coull and Nobre, 1998) might be more promising.

References


