Neuronal correlates of continuous manual tracking under varying visual movement feedback in a virtual reality environment

Jakub Limanowski a,b,⁎, Evgeniya Kirilina a,b,c, Felix Blankenburg a,b

a Neurocomputation and Neuroimaging Unit, Department of Education and Psychology, Freie Universität Berlin, Berlin, Germany
b Center for Cognitive Neuroscience Berlin, Freie Universität Berlin, Berlin, Germany
c Department of Neurophysics, Max Planck Institute for Cognitive and Brain Sciences, Leipzig, Germany

ABSTRACT

To accurately guide one’s actions online, the brain predicts sensory action feedback ahead of time based on internal models, which can be updated by sensory prediction errors. The underlying operations can be experimentally investigated in sensorimotor adaptation tasks, in which moving under perturbed sensory action feedback requires internal model updates. Here we altered healthy participants’ visual hand movement feedback in a virtual reality setup, while assessing brain activity with functional magnetic resonance imaging (fMRI). Participants tracked a continually moving virtual target object with a photorealistic, three-dimensional (3D) virtual hand controlled online via a data glove. During the continuous tracking task, the virtual hand’s movements (i.e., visual movement feedback) were repeatedly periodically delayed, which participants had to compensate for to maintain accurate tracking. This realistic task design allowed us to simultaneously investigate processes likely operating at several levels of the brain’s motor control hierarchy. FMRI revealed that the length of visual feedback delay was parametrically reflected by activity in the inferior parietal cortex and posterior temporal cortex. Unpredicted changes in visuomotor mapping (at transitions from synchronous to delayed visual feedback periods or vice versa) activated biological motion-sensitive regions in the lateral occipitotemporal cortex (LOTC). Activity in the posterior parietal cortex (PPC), focused on the contralateral anterior intraparietal sulcus (aIPS), correlated with tracking error, whereby this correlation was stronger in participants with higher tracking performance. Our results are in line with recent proposals of a wide-spread cortical motor control hierarchy, where temporoparietal regions seem to evaluate visuomotor congruence and thus possibly ground a self-attribution of movements, the LOTC likely processes early visual prediction errors, and the aIPS computes action goal errors and possibly corresponding motor corrections.

1. Introduction

It has recently been established that the brain accomplishes accurate online action control by constructing internal models that continually predict the sensory consequences of motor commands, and which can be updated by sensory prediction errors, thus compensating for sensorimotor noise and conduction delays (Wolpert et al., 1998a, 1998b; Miall et al., 1993; Desmurget and Grafton, 2000; Todorov and Jordan, 2002; Shadmehr and Krakauer, 2008; Desmurget et al., 1999). Such predictive or “forward” models estimate the next state of the body given its current state and the predicted sensory consequences of the movement, and thereby need to appropriately weight and integrate sensory feedback from multiple modalities (Körding and Wolpert, 2004).

The underlying neuronal mechanisms have been investigated by experimentally controlled manipulations of sensory feedback during actions, introducing spatial perturbations or temporal delays to visual or proprioceptive movement feedback to investigate the sensorimotor adaptation of internal motor control models during goal-directed actions (Desmurget et al., 1999; Foulkes and Miall, 2000; Diedrichsen et al., 2005; Tseng et al., 2007; Grafton et al., 2008). A fruitful line of research has focused on the neuronal evaluation of visuomotor (in)congruence, i.e., on violations of predictions about the visual consequences of movements. Activity in the inferior parietal and temporal cortex has frequently been observed to increase during visuomotor discordance, which has been interpreted as indicating that these regions evaluate visuomotor congruence and by this evaluation could ground the “sense of agency” over one’s movements (Farrer and Frith, 2002; Leube et al., 2003; Ogawa et al., 2007; Farrer et al., 2008; Nahab et al., 2011). Recently, regions in the LOTC have also received...
much attention from action research, since they have been found activated by unseen hand movements (Astafiev et al., 2004) and are also modulated by visuomotor and visuopropraeptive congruence (David et al., 2007; Limanowski and Blankenburg, 2016) and by action planning (Zimmermann et al., 2012; Gallivan et al., 2016). However, the exact role of the LOTC in action is still subject to debate and could involve merely visual, multisensory, or even sensorimotor processing (see Lingnau and Downing, 2015, for a review).

Another line of research, based on neurophysiological work in monkeys (e.g. Taira et al., 1990; Sakata and Taira, 1994) and lately also neuroimaging in humans (e.g. Grekes et al., 2004; Diedrichsen et al., 2005; Tunik et al., 2007; Grafton et al., 2008) as well as investigations of related lesions (Wolpert et al., 1998b; Rushworth et al., 1997; Sirigu et al., 2004) has identified the posterior parietal cortex (PPC), specifically regions lining the intraparietal sulcus (IPS), and the cerebellum as potential sites of forward models for motor control that implement error-driven adaptation (for reviews see Desmurget and Grafton (2000), Grafton (2010), Shadmehr and Krakauer (2008), Scott (2012)). In sum, this recent empirical work suggests a wide-spread motor control hierarchy in the brain (Hamilton and Grafton, 2007; Grafton, 2010).

A drawback of most of the mentioned studies investigating action control is the use of abstract visual movement feedback in the form of e.g. a cursor (Desmurget et al., 1999; Diedrichsen et al., 2005; David et al., 2007; Ogawa et al., 2007; Grafton et al., 2008), which does not adequately reflect visual feedback of the body received during real life actions. By neglecting vision of the body, such studies may potentially miss important processes concerned with bodily action control (Makin et al., 2015; Salomon et al., 2016). Some experiments have used video recordings or virtual reality to provide realistic hand movement feedback (Leube et al., 2003; Nahal et al., 2011). However, as these studies focused on the explicit self-attribution of perturbed visual movement feedback (i.e., judgments of a "sense of agency"), they used isolated movements that were not goal-directed. These movements might have recruited only part of the network involved in visually guided action control—for example, there was no potential action goal- or performance-error.

In the present fMRI experiment, we therefore employed a continuous manual target-tracking task in a realistic virtual reality environment, which was designed to investigate dynamic manual action control by the brain’s motor hierarchy in an ecologically valid way. Participants wore an MR-compatible data glove controlling a photorealistic virtual hand model, presented in 3D via stereoscopic goggles, in real-time. Their task was to track a continually moving target: by closing and opening their hand, participants had to keep the fingertips of the virtual hand within the boundaries of a semi-transparent ellipse (the target) that moved up and down in a sine-like trajectory. During the task, we periodically introduced 20 s blocks of delayed visual hand movement feedback, i.e., the virtual hand’s movements lagged behind the actually performed movements. As the ellipse continued its regular (predictable) trajectory, participants had to compensate for the visuomotor delay in order to maintain accurate tracking. Notably, participants also had to adjust their movements when the visual feedback changed again from delayed to synchronous at the end of each delayed block. Thus while continually moving, participants repeatedly had to update their internal models estimating the relation between their motor output and the associated sensory (visual) consequences, while being presented with repeated but unpredictable changes in the visual feedback policy. In contrast to previous designs offering action feedback only on a trial-by-trial basis (see above), our task design enabled us to simultaneously investigate the neuronal correlates of various levels of the motor control hierarchy, involved in (i) the processing of delayed visual hand movement feedback per se (delayed versus synchronous feedback periods), which we assumed to activate inferior parietal and temporal regions implied by previous research on visuomotor discordance (ii) the processing of visual prediction error (unpredicted changes in visual feedback at transitions from synchronous to delayed feedback periods and vice versa), which should lead to activation of lower-level visual (self-) motion-selective areas, e.g. in the LOTS (we used separate functional localizers to specify the obtained activations e.g. as visual motion-selective), and (iii) the processing of tracking (i.e., action goal) error and resulting motor corrections, which we assumed would be reflected by activation of potential high-level sites of forward models for goal-directed hand actions such as the aIPS. Following the assumption that error-driven model update underlies motor learning and thus better performance, we also examined whether participants showing high tracking performance would also show increased activation of the same areas.

2. Methods

2.1. Participants

16 healthy, right-handed volunteers (11 male, mean age=27 years, range=21–37) participated in the experiment after signing informed consent. The experiment was approved by the ethics committee of the Freie Universität Berlin and conducted in accordance with the approved guidelines.

2.2. Experimental design and procedure

During the experiment, participants lay inside the scanner wearing an MR-compatible data glove (5DT Data Glove MRI) and digital stereoscopic goggles (VisuaSTIM, 800×600 pixels resolution, 30° eye field) for presentation of the virtual reality, which was instantiated in the open-source 3D computer graphics software Blender (http://www.blender.org) using its Python programming interface. The data glove was used to measure the participant’s finger flexions via sewn-in optical fiber cable sensors (1 sensor per finger, 8 bit flexure resolution per sensor). This information was fed to a photorealistic virtual 3D hand model, which was thus moveable by the participant in real-time (Fig. 1 a; the real hand was positioned across the participant’s chest in a corresponding posture). The sampling rate was 60 Hz, determined by the frame rate of the 3D graphics software. Before scanning, the glove was carefully calibrated to best fit each participant’s individual finger movement range (if necessary this was repeated between runs), and participants completed a practice session to get familiar with the task.

During the whole experiment, a semi-transparent ellipse (the target) was moving diagonally up and down in front of the virtual hand along a “biologically plausible”, approximately sinusoidal trajectory at 0.5 Hz (see Fig. 1 b and Supplementary material for details). Participants were instructed to close and open their real hand (i.e., to bend all fingers except the thumb) to continually keep the virtual hand’s index, middle, and ring finger tips (the little finger was also moved along) within the area of the moving semi-transparent ellipse. After an initial 60 s period of this task, the visual feedback of hand movements was periodically delayed, i.e., the movements of the virtual hand then lagged behind the movements actually performed by the participant. In each of 4 runs in total, 8 lengths of visual feedback delay (33, 67, 100, 133, 167, 200, 233, and 267 ms) were randomly presented in 20 s blocks, alternating with 20 s blocks of synchronous (non-delayed) visual feedback, resulting in approximately 7 min run length. Participants were instructed to fixate a white dot throughout the experiment, which was located in front of the ellipse, in the middle of its trajectory.

Throughout the experimental runs, we recorded the participant’s finger positions, the displayed finger positions of the 3D hand model, and the position of the target ellipse (Fig. 1 b). From these raw data we created six vectors, which were used as regressors in our fMRI analysis (see below and Fig. 1 c): vector 1 contained the parametric length of visual feedback delay (0–267 ms). Vectors 2 and 3 were calculated based on the temporal derivative of vector 1, thus representing visual
feedback transitions: vector 2 represented transitions from synchronous to delayed periods, vector 3 represented transitions from delayed to synchronous periods; these transitions were parametrically dependent on the length of the respective delay period (see Fig. 1c). Vector 4 contained the participant’s tracking error, calculated as the unsigned relative phase shift between the oscillatory phase of the target ellipse and the phase of the virtual hand’s fingertips within a 120 samples time window. Vector 5 was the temporal derivative of vector 4, which was used to examine the effects of the current rate of decrease in tracking error. Vector 6 was representing the motor output amplitude, calculated based on an envelope around the amplitude of current finger movements, and was included in the models as a regressor of no interest to account for motor output-related variance.

After the main experiment, participants completed three separate functional localizer runs, in which we identified areas responsive to the grasping movement per se, to a playback of one’s own movement, and to hand vision. Activations obtained from these localizer runs were used for functional specification of the activations obtained from the main analysis. In the first run, we identified brain areas responding to hand motion per se. Participants performed the same hand movements as in the main experiment, following the ellipse’s trajectory, but without visual hand feedback (in blocks of 20 s, separated by 20 s rest periods; total run duration approx. 4 min). As the movements were also cued with the moving ellipse (but without visual hand feedback), it also activated areas selective to visual motion. To account for this and to only obtain strictly motor-related activations, we masked the results exclusively with the visual motion localizer mask image. In the second run, we identified brain areas responding to visual (biological) motion.

To this end, an excerpt of one of the participant’s actual experimental runs was played back in blocks of 20 s, separated by 20 s blocks of a static image; total run duration approx. 4 min. In the third run we displayed static images of the virtual hand in different positions or virtual control objects, which were various cylindrical or squared shapes matching the virtual hand’s position and texture (in blocks of 20 s, separated by 20 s fixation blocks; total run duration approx. 8 min).

Participants also completed a separate behavioral experiment outside the scanner to identify the individual threshold at which each of the 8 delay lengths was consciously perceived as delayed visual feedback. This session consisted of two runs à 4.6 minutes of the same tracking task as the main experiment. However, now participants additionally had to indicate via left-hand keyboard button press whether they perceived the current visual feedback as delayed (‘D’ key) or synchronous (‘S’ key). Each delay length was presented twice per run in randomized order; to further render the onset of delays unpredictable, each delay presentation could last either 4 or 5 s, while the synchronous phases could last either 2, 3, 4, or 5 s. To account for reaction time (i.e., when participants had already detected the delay but not yet reported it via key press), we determined each participant’s average reaction time until button press for the two highest delay periods (233 and 267 ms, assuming that delays above 200 ms would be clearly perceived, following similar results from Leube et al., 2003), and we subtracted this value (1.08 s on average) individually from the recorded time points of key presses before evaluation. We chose an above-chance-level (>50%) correct detection rate as the threshold to classify a delay length as “consciously perceived” by that participant.

2.3. FMRI data preprocessing and analysis

The fMRI data were recorded using a 3 T scanner (Tim Trio, Siemens, Germany), equipped with a 12-channel head coil. T2*-weighted images were acquired using a gradient echo-planar imaging sequence (voxel size=3×3×3 mm³, 20% gap, matrix size=64×64, TR=2000 ms, TE=30 ms, flip angle=70°). For each participant, we recorded 1352 functional images in total (864 for the main experiment, 488 for the localizer runs), a field map (TE1=10.00 ms, TE2=12.46 ms), and a T1-weighted structural image (3D MPRAGE, voxel size=1×1×1 mm³, FOV=256×256 mm², 176 slices, TR=1900 ms, TE=2.52 ms, flip angle=9°). FMRI data were preprocessed and analyzed using SPM8 (www.fil.ion.ucl.ac.uk/spm/). Artifacts in individual slices were interpolated using the ArtRepair toolbox (Mazuka et al., 2009). Functional images were then realigned and unwarped, corrected for slice acquisition time differences, normalized to MNI space using DARTEL and resliced to 2 mm voxel size, spatially smoothed with an
8 mm full width at half maximum Gaussian kernel, detrended (Macey et al., 2004), and images featuring excessive scan-to-scan movement were interpolated (ArtRepair toolbox).

The six vectors extracted from the raw experimental data (see above) were convolved with SPM’s hemodynamic response function and entered as continuous mean-centered regressors for each run into the first-level general linear models (GLM, 300 s high-pass filter) fit to each participant. The regressors were not orthogonalized with respect to each other, in order to not assign shared variance to any single regressor and thus to obtain more clearly interpretable effects of the individual regressors (cf. Mumford et al., 2015). There were moderate correlations between the first-level regressors modeling the delay and the tracking error (mean Pearson’s $R=\.39$), the transition from synchronous to delayed visual feedback and the derivative of the error (mean $R=\.31$), and the delay and motor amplitude (mean $R=\.24$); all other correlations were minimal. The five principal components accounting for the most variance in the cerebrospinal fluid or white matter signal time course each (Behzadi et al., 2007) were added to these GLMs alongside the realignment parameters as regressors of no interest. Further, we tested for differences between consciously perceived versus not perceived delay periods and transitions. To this end, we calculated separate GLMs using the individual behavioral delay detection (see above) to model detected (above chance) delay periods and the corresponding transitions with +1, and the undetected ones with −1.

On the group-level, the resulting first-level contrast images were entered into one-way ANOVAs (main analysis and supra-vs-subthreshold analysis), one-sample $t$-tests (functional localizers), or two-sample $t$-tests (comparison of high versus low performers). We only report and display activations obtained from group-level contrasts that survived a statistical significance threshold of $p<\.05$, false discovery rate (FDR) corrected for multiple comparisons on the cluster level with an initial voxel-wise threshold of $p<\.001$, uncorrected. Where explicitly stated, we applied peak-level familywise error small volume correction within regions of interest defined by the significant activations obtained from the functional localizer runs. The resulting statistical parametric maps (SPMs) are projected onto the mean normalized structural image or rendered on SPM’s brain template, with cluster extent adjusted to display only clusters that survived correction for multiple comparisons. The SPMs can be inspected online at http://www.neurovault.org/ collections/1904/ (each SPM is also separately available at a threshold of $p<\.001$, uncorrected). The SPM Anatomy toolbox (Eickhoff et al., 2005) was used for anatomical reference.

3. Results

3.1. Behavioral results

All participants were able to perform the task, with an average tracking performance of $R=\.62$ (SD=\.21; quantified as participants’ tracking adaptation to the delay time course, i.e., as the correlation between green and grey curves in Fig. 1c). Half of the participants had a tracking performance above average (mean $R=\.79$, SD=\.12), and half below average (mean $R=\.44$, SD=\.10); these were thus classified as “high” and “low” performers, respectively (see Supplementary Fig. S1 for examples of high and low tracking performance). The results of the behavioral experiment showed that on average, the length of delay was proportional to its detection rate ($R=\.99$): participants detected longer delays more easily, with delays up to 133 ms below, and delays longer than 167 ms above chance level (see Fig. 2).

3.2. Functional localizer runs

For functional specification of the activations observed in the main experiment as e.g. grasp-related or biological motion-sensitive, we first analyzed the separate functional localizer runs. Visual playback (versus a static image) of each participant’s own movements produced significant ($p<\.05$, corrected for multiple comparisons, see Fig. 3a) activations in the bilateral LOTC, dorsal premotor cortex (PMd), superior parietal lobe (SPL), right supplementary motor area (SMA), right posterior superior temporal gyrus (pSTG), left primary somatosensory cortex, and left cerebellum. Grasping without visual hand feedback produced significant ($p<\.05$, corrected, see Fig. 3b) activations in the left primary motor and premotor cortex, the bilateral SMA and cerebellum, left putamen, left primary and secondary somatosensory cortex, and left cuneus. These areas were not activated by visual playback (they fell outside the mask image obtained from the visual playback contrast), but since movements were cued by the moving ellipse as in the main experiment, some further activations were located in visual motion-sensitive areas (see Fig. 3b). Finally, significantly ($p<\.05$, corrected) higher activation by hand versus object vision were observed in the bilateral lingual and fusiform gyri, and at a
more liberal threshold also in the left LOTC (x=-44, y=-86, z=0, t=3.09, p=.002, uncorrected).

3.3. Brain activity associated with visual feedback delay and transitions

In the main fMRI analysis, we first examined the parametric effect of the delay regressor (Fig. 1c) to identify brain activity modulations by delayed visual hand movement feedback per se. BOLD signal in the right temporal cortex, left inferior parietal cortex (IPC), right inferior frontal gyrus (IFG), superior frontal gyrus (SFG) and SMA, and in the right anterior lateral cerebellum correlated significantly (p < 0.05, corrected) with the length of visual feedback delay, i.e., activity in these regions increased parametrically with the length of delay. The right temporal cluster was focused on the posterior superior temporal sulcus (pSTS) and extended to the posterior middle and superior temporal gyrus (pMTG, pSTG); the left IPC cluster spanned the supramarginal and angular gyrus (SMG, AG). See Fig. 4a and Table 1 for details (Supplementary Fig. S2a shows further uncorrected activity in the left IFG and anterior insula, AI). Notably, regions in the right pSTG and pMTG were also activated by visual motion playback (contained within visual movement playback localizer mask image, see inlay in Fig. 4a).

The SMA activation was contained within areas activated by grasping per se but not by visual playback (contained within motor localizer but outside of visual playback localizer). Further, we contrasted supra- versus subthreshold delay periods (defined individually for each participant as delays detected above versus below chance level in the behavioral experiment), which revealed significantly stronger activation of the left IPC by delay periods that were reliably perceived (x=-60, y=-50, z=46, t=4.68, p < .05, corrected within activations obtained from the main GLM delay contrast).

We next tested for effects of an unpredicted change in the visual feedback policy. Note that in our design such changes occurred at transitions from synchronous to delayed visual movement feedback periods, and vice versa at transitions from delayed to synchronous feedback periods—either transition introduced a change in visuomotor mapping. To identify general effects independent of the direction of transition, we therefore calculated a “null” conjunction of both transition contrasts, which identified clusters of voxels that significantly responded to both transitions (see Supplementary Fig. S4 online for the individual transition contrasts). This analysis revealed significant activations in the bilateral LOTC that fell within areas activated by grasping (BA 44) and supplementary motor area (Supplementary Fig. S4a). Notably, regions in the right pSTS and pMTG were also activated by visual motion playback (contained within visual movement playback localizer mask image).

![Fig. 4. Brain activity associated with visual feedback delay and visual feedback transitions. (a) Render and overlays of brain areas in which activity was significantly (p < .05, corrected for multiple comparisons) correlated with parametrically delayed visual movement feedback: the right temporal cortex, focused on the pSTS, the left IPC, right IFG, bilateral SFG/SMA, and right lateral anterior cerebellum. The inlay shows overlays of areas activated by delayed visual feedback with activations by visual (biological) motion playback as obtained from the separate functional localizer run (p < .001, uncorrected). SPM available at http://www.neurovault.org/images/29225/. (b) Visual feedback transitions from synchronous to delayed visual feedback periods and vice versa significantly (p < .05, corrected; conjunction of both transition directions) activated the bilateral LOTC (identifed as strongly responsive to visual playback of one’s movements); these areas were also more strongly activated by transitions to and from supra- versus subthreshold delay periods. SPM available at http://www.neurovault.org/images/29226. See Table 1 for details.](85)

<table>
<thead>
<tr>
<th>Anatomical region</th>
<th>MNI x,y,z</th>
<th>Peak t</th>
<th>P (corrected)</th>
</tr>
</thead>
<tbody>
<tr>
<td>R. Temporal cortex (pSTS, pSTG, and pMTG)</td>
<td>56 -44 6</td>
<td>4.62</td>
<td>.001</td>
</tr>
<tr>
<td>R. Inferior frontal gyrus (BA 44)</td>
<td>58 16 2</td>
<td>4.86</td>
<td>.001</td>
</tr>
<tr>
<td>L. Inferior parietal cortex (SFG and pSTG)</td>
<td>-50 -56 54</td>
<td>4.08</td>
<td>.001</td>
</tr>
<tr>
<td>L. Superior frontal gyrus and supplementary motor area</td>
<td>-26 0 64</td>
<td>4.22</td>
<td>.002</td>
</tr>
<tr>
<td>R. Cerebellum (Lobule VI/VIIa)</td>
<td>38 -50 -36</td>
<td>4.31</td>
<td>.008</td>
</tr>
<tr>
<td>Correlations with visual feedback transitions (conjunction)</td>
<td>-50 -72 4</td>
<td>4.81</td>
<td>.005</td>
</tr>
<tr>
<td>L. Middle occipital gyrus (LOT/ V5)</td>
<td>-50 -72 4</td>
<td>4.81</td>
<td>.005</td>
</tr>
<tr>
<td>R. Middle occipital gyrus (LOT/ V5)</td>
<td>42 -72 6</td>
<td>4.33</td>
<td>.044*</td>
</tr>
</tbody>
</table>

*a Peak-level small volume correction within visual motion localizer mask image. BA=Brodman’s area.
3.4. Brain activity associated with tracking error

We next examined the effect of the unsigned tracking error regressor (cf. Fig. 1c) to identify brain areas potentially involved in error correction and thus in fine-tuned online motor control during goal-directed action. Tracking error significantly ($p < .05$, corrected) correlated with BOLD signal in the left PPC (the cluster was focused on the aIPS and spanning to SPL), right SPL, right SMA, and left LOTC (see Fig. 5a and Table 2 for details; at an uncorrected threshold, similar correlations were also observed in the bilateral cerebellum, the right IFG, right IPS, and in the bilateral PMd, see Supplementary Fig. S5). Of these activations, those in the left aIPS ($x=-48$, $y=-30$, $z=42$) fell within areas activated by grasping but not by visual playback (motor localizer activations outside of visual motion localizer mask), whereas those in the left SPL ($x=-38$, $y=-44$, $z=64$) and LOTC ($x=-54$, $y=-78$, $z=2$) fell within areas activated by visual playback; the SMA activations spanned across both (motor and visual playback) localizer mask images. See inlay of Fig. 5a for details. Notably, in high performers, activity in the left aIPS was more strongly correlated with tracking error than in low performers ($p < .001$, uncorrected, Fig. 5b). To validate these differences, we performed a separate analysis with individual tracking performance as a covariate of tracking error. Indeed, at a more liberal threshold, this revealed a positive correlation between individual tracking performance and tracking error-related activity at corresponding locations in the aIPS ($x=-50$, $y=-32$, $z=42$, $t=3.30$, $p=.003$, uncorrected). Both analyses revealed similar effects in the bilateral cerebellar regions activated by tracking error (Supplementary Fig. S5). Thus the left aIPS (and bilateral cerebellum) were more strongly activated by tracking error in participants who had a better tracking performance. No significant activation differences in these areas were found between tracking error during supra- versus subthreshold delay periods.

To identify brain areas responsive to the rate of error reduction, we moreover tested for BOLD signal correlations with the negative temporal derivative of the tracking error regressor. The rate of error reduction was significantly ($p < .05$, corrected) correlated with activity in the bilateral cerebellum, left putamen, right thalamus, and right IFG (Fig. 5c and Table 2); the activations in the putamen and cerebellum fell within areas activated by grasping (motor localizer). The left putamen moreover showed this effect more strongly in high versus low performers ($x=-18$, $y=-4$, $z=-8$, $t=4.66$, $p < .001$, uncorrected).

Finally, a control analysis of the effect of motor amplitude (included in the first-level GLMs as a regressor of no interest to account for motor output-related variance) revealed significant activations in the left putamen and bilateral cerebellum ($p < .05$, corrected, see Supplementary Fig. S6 and Table S1). Crucially, however, motor output amplitude was neither significantly correlated with tracking error (mean correlation between regressors across participants $R=.05$, $SD=.18$, n.s.) nor with the rate of tracking error reduction (mean $R=.05$, $SD=.08$, n.s.).

4. Discussion

Using a novel, continuous virtual reality-based hand-target tracking task with repeated alterations of delayed and synchronous visual movement feedback, we investigated the neuronal motor control hierarchy involved in online visually guided hand action control. Our
results suggest a distributed network of brain regions that process visuomotor conflicts, visual prediction errors, and action goal errors at different stages of such a motor hierarchy.

Parametrically varied visual feedback delay had the most prominent effect on activity in the right posterior temporal cortex and left inferior parietal cortex. Temporoparietal activation (particularly right-hemispheric) has been frequently observed under delayed or distorted visual movement feedback, and may indicate the detection of a mismatch between intended and actual movement outcome and a corresponding “loss” of a sense of agency, i.e., the self-attrbution of a movement (Farrer and Frith, 2002; Leube et al., 2003; Ogawa et al., 2007; Farrer et al., 2008; Nahab et al., 2011). It has been proposed that unpredicted visual movement consequences should especially be processed in biological motion-sensitive areas such as the pSTS (Grezes et al., 2001; Leube et al., 2003). Indeed, we observed that some of the delay-related right-hemispheric temporal activations fell within areas also activated by a playback of the participants’ own movements. Together with previous findings, our results imply the temporal and inferior parietal cortex in the processing of visuomotor mismatches and the prediction of sensory consequences of one’s actions.

An alternative but not exclusive interpretation of IPC activation by delay is that it reflected an updated set weights assigned to (conflicting) visual and proprioceptive information, as regions in the IPC seem to represent arm position based on integrated visual and proprioceptive information (Graziano et al., 2000; Limanowski and Blankenburg, 2016). Multisensory integration is an essential part of online motor control in the brain and relies on Bayes-optimal weighting of individual sensory information (van Beers et al., 1999; Körding and Wolpert, 2004). There is some evidence that the relative suppression of proprioceptive information, along with a corresponding visual dominance, may enable visuomotor adaptation (Bernier et al., 2009; Balslev et al., 2004). In our case this would likewise imply an up-weighting of visual information, which tentatively fits with the initially increased LOTC activations by transitions to delayed visual feedback. However, it is also possible that under delay, proprioception could be weighted more strongly. Rushworth et al. (2001) found that attention to upcoming movements during their preparation also produced left IPC, IFG (BA 44), and right cerebellar activity—these regions also responded to visual feedback delay in our study. Thus there may have been increased “motor attention” during delayed visual feedback, and this could involve increased proprioceptive weights. However, we did not observe any activity in primary proprioceptive areas (BA 3a or secondary somatosensory cortex). Interestingly, we found that a region in the left IPC was activated more strongly during delay periods in which the visual feedback was consciously perceived as delayed, which may reflect increased cognitive (intentional, cf. Sirigu et al., 2004) control under visuomotor discordance. These alternative or complementary interpretations are exciting questions for future research, however, all of them are compatible with a role of the IPC in self-attributing movements based on multimodal congruence.

Delay further significantly activated regions in the IFG, SMA, and cerebellum. The IFG has been implied in hand action control, specifically, in motor attention (Rushworth et al., 2001) and in conscious action selection based on learned associative rules (Toni et al., 2001). Both processes might have been engaged more strongly during delayed visual action feedback. The SMA has been associated with updating reaching plans to new task requirements or learning new visuomotor associations during action (Matsuzaka and Tanji, 1996; Sakai et al., 1999; Krakauer et al., 2004). In our case, a delay period likely introduced a novel visuomotor mapping that needed to be learned. The response to delay in the right lateral cerebellum matched the coordinates of a previously reported activation to delays between one’s movements and the corresponding tactile sensations in a modified self-touch paradigm (Blakemore et al., 2001), and thus is well in line with the proposed role of the cerebellum in predicting the consequences of one’s movements (Miall et al., 1993; Wolpert et al., 1998a).

Transitions from synchronous to delayed visual feedback periods and vice versa, i.e., unpredicted visual consequences of one’s movement, produced strong activity increases in the bilateral LOTC. These activations fell within areas identified as visual motion-sensitive in a separate localization run and corresponded to motion-sensitive area V5. However, the coordinates of these activations also match those of the visually body-selective “extrastrate body area” (EBA, Downing et al., 2001). LOTC activity has been associated with processing visual and proprioceptive information about arm position (Astafiev et al., 2004; cf. Limanowski et al., 2014; Gallivan et al., 2016; Limanowski and Blankenburg, 2016). Moreover, EBA activity has been observed during the preparation and execution of hand actions (Astafiev et al., 2004; Gallivan et al., 2016). However, there are several possible explanations for LOTC activations by action (cf. Lingnau and Downing, 2015), for instance dynamic visuopropiceptive comparison and integration, or visuomotor comparisons involving predictive motor signals (Astafiev et al., 2004; Gallivan et al., 2016). David et al. (2007) observed LOTC activation to incongruent mouse-cursor action feedback and speculate that LOTC might already process sensorimotor incongruence before parietal regions. Interestingly, we observed that the LOTC was also significantly activated by transitions from delayed to synchronous visual feedback periods, i.e., from conflicting to matching visual and proprioceptive information. Thus the LOTC did not merely respond to a conflict between visual and proprioceptive hand position information, but more generally to a mismatch between predicted and actual visual hand position. We therefore speculate that the LOTC activations observed during sudden visual feedback policy transitions reflect an early stage of motor control, specifically, the detection of visual (biological motion) prediction errors resulting from visual movement feedback that was not predicted by visuomotor mapping represented in higher-level motor areas. In line with previous speculations (cf. David et al., 2007; Makin et al., 2012), this would imply that prediction errors are passed from LOTC to PPC, and the PPC in turn may determine the weighting of visual self-motion information in LOTC depending on context. This could also explain the strong LOTC activation by tracking error as increased attention to visual motion information in response to action errors in a visually-guided tracking task. An alternative explanation could be that the LOTC provides a visual estimate of hand goal state for action planning, as proposed by Zimmermann et al. (2012), and that this estimate also needed updating in the light of performance error. In sum, our results align with those of others in indicating the LOTC at an early stage of action control.

Some activations by visual feedback transitions that did not survive correction for multiple comparisons included a region in the right pSTS. This activation matched the coordinates reported for activity during a subjective delay-detection task by Leube et al. (2003), and thus supports the proposed role of pSTS in predicting the visual consequences of one’s actions (see above). Further, the bilateral AI were responding to visual feedback transitions and also parametrically to visual feedback delay. The AI has been associated with change and error detection (Corbetta et al., 2008; Ullsperger et al., 2014), and could thereby contribute to a general awareness and self-attribute of the body (Craig, 2009; Limanowski et al., 2014; Allen et al., 2016). We hence encourage future research to follow up on our findings.

Participants’ tracking error was most prominently reflected by contralateral PPC activity, lining the anterior (aIPS/BA 2) to posterior IPS (SPL/BA 7). The PPC, specifically regions around the IPS, likely constitutes a high-level part of the brain’s motor control hierarchy (Cohen and Andersen, 2002; Tunik et al., 2007; Hamilton and Grafton, 2007). The aIPS is crucially involved in hand and grasp control and thus considered the putative human homologue of monkey area AIP (Rizzolatti et al., 1997; Frey et al., 2005; Culham et al., 2006; Grafton, 2010). Correspondingly, the aIPS activation we observed was contained within areas activated by grasping (motor localizer), and there was a
which helps correct the motor commands in the case of mismatching PPC uses a dynamic internal representation of the state of the body (the 2007). It has therefore been proposed that in visually guided action, the activated by errors during trial-by-trial visuomotor learning (Grafton 2005, cf. Della-Maggiore et al., 2004). Correspondingly, the PPC is similar results have been reported after TMS over SPL, Glover et al., 1998a; Miall et al., 1993 ). Thereby cerebellar models seem to be force scaling (Grafton, 2010) and gain learning (Krakauer et al., 2004). However, the putamen was not activated by performance error. In a

clear lateralization of the effects contralateral to the moving limb. Previous work has shown that disrupting left aIPS activity with transcranial magnetic stimulation (TMS) impairs right hand path corrections during visually guided reaching (Desmurget et al., 1999; similar results have been reported after TMS over SPL, Glover et al., 2005, cf. Della-Maggiore et al., 2004). Correspondingly, the PPC is activated by errors during trial-by-trial visuomotor learning (Grafton et al., 2008) and mouse-cursor curve tracking tasks (Ogawa et al., 2007). It has therefore been proposed that in visually guided action, the PPC uses a dynamic internal representation of the state of the body (the hand) and the world (the target) to compute a dynamic error signal, which helps correct the motor commands in the case of mismatching predicted final hand and target locations (Desmurget et al., 1999, 2000, 2001; cf. Rushworth et al., 1997; Wolpert et al., 1998b). We therefore propose that in our experiment, the aIPS was concerned with maintaining accurate tracking, i.e., with reaching the action goal, and that aIPS activation by error thus reflected the computation of a dynamic hand-target “action goal” error for a corrective motor response (cf. Diedrichsen et al., 1999; Desmurget and Grafton, 2000; Tunik et al., 2007). This interpretation is further supported by our finding that aIPS activation by errors was stronger in participants who maintained more accurate tracking. Conversely, the aIPS did not show differential error-related activity during supra- versus subthreshold delay periods, which may indicate that tracking error computation and correction was largely independent of conscious processing of delayed sensory feedback, as suggested by behavioral experiments (Fournier and Jeannerod, 1998; Weibel et al., 2015).

The more medial and superior parietal activation was also responsive to visual playback of hand action (visual motion localizer). The mIPS has previously been shown to be involved in visuomotor coordinate transformations during goal-directed movements with a mouse cursor (Eskandar and Assad, 2002; Grefkes et al., 2004; in these studies it also responded to visual playback of the movement). Our pattern of rather motor-related aIPS activation and visual motion-related SPL activation fits the speculation about somatosensory-to-visual gradients in the PPC (e.g., Grefkes and Fink, 2005). However, both state estimation and integration of appropriately weighted multisensory inputs are essential contributions to goal-directed action control (see above); correspondingly, recent work suggests that the aIPS and SPL are tightly coupled in their function (Grafton, 2010; Heed et al., 2011; Verhagen et al., 2013).

Tracking error further significantly activated the SMA. Previous human neuroimaging studies report similar SMA activation by performance error during visuo-motor learning (Grafton et al., 2008). Speculatively, the SMA activation by tracking error may indicate an update of movement plans. However, SMA activity was also correlated with delay length and could hence reflect learning novel visuomotor associations (see above), therefore its exact role remains to be specified by future research.

Further activations by tracking error were observed in the bilateral medial cerebellum. Although these activations did not survive corrected threshold, they are noteworthy given the often proposed importance of the cerebellum in online motor control (Wolpert et al., 1998a; Miall et al., 1993). Thereby cerebellar models seem to be informed by sensory prediction errors (Tseng et al., 2007; Diedrichsen et al., 2005) and possibly also by dynamical error signals computed by the PPC (Desmurget and Grafton, 2000). Given that, like in the aIPS, the cerebellar error-related activations were stronger in high versus low performers, our results tentatively support such proposals.

Finally, we observed significant correlations of cerebellar and putamen activity with the rate at which tracking error was reduced. The cerebellar activations are well in line with its role in fine-tuned motor control (see above). The putamen (the basal ganglia in general) have been associated with several functions in motor control, such as force scaling (Grafton, 2010) and gain learning (Krakauer et al., 2004). However, the putamen was not activated by performance error. In a visuomotor delay detection task (Leube et al., 2003), the putamen responded inversely to delay length; the authors speculate that the putamen might be involved in attenuating the predicted sensory consequences of actions by providing corresponding motor signals. This could also be a potential explanation of our result, which however should be addressed more specifically by future research.

Our experiment was designed as realistic as possible to target processes that work closely together as a motor control hierarchy in real life settings. This is at the same time a potential limitation of some of our results’ interpretability: participants made more errors during delayed visual feedback periods (the tracking task was harder under delayed visual feedback, cf. Foulkes and Miall, 2000), and consequently the regressors modeling delay and tracking error were somewhat correlated (mean R²=.39), which means that our results might have missed some of their shared variance. Therefore these results will need to be validated by specific, more restricted experiments.

Finally, while our interpretation largely draws on previous proposals of a forward-inverse model architecture of the motor system, it is also compatible with recent alternative accounts of motor control based on active inference that appeal to a general free-energy principle of brain function (Friston, 2010). These accounts propose a common mechanism governing perceptual and motor systems, namely, prediction error minimization across hierarchical generative models (Killner et al., 2007; Adams et al., 2013). Active inference thus offers compelling explanations of recurrent exchange of predictions and errors in the motor control hierarchy, which could readily accommodate our results, e.g. by the interpretation of LOTC processing visual prediction errors and parietal regions processing higher-level visuomotor mappings. A challenging task for future research is to design experiments that may explicitly test the assumptions of active inference in the domain of motor control.

5. Conclusion

We investigated the neuronal correlates of adaptive manual action control using a novel virtual reality-based hand-target tracking task with manipulated photorealistic visual hand movement feedback. We observed that inferior parietal and temporoparietal regions responded to visuomotor mismatches, thereby possibly grounding the self-attribute of a movement as previously suggested. Unpredicted visual movement consequences were processed in biological motion-sensitive occipitotemporal regions, which may imply the encoding of visual prediction errors in these regions. Hand-target tracking errors engaged anterior intraparietal areas; these areas may have computed corresponding motor corrections in response to errors in order to still reach the desired action goal. Our findings support previous electrophysiological work in monkeys and neuroimaging work in humans. Moreover, by the simultaneous investigation of processes that likely operate at different levels of the motor control hierarchy, our results improve the understanding of the wide-spread brain network involved in the online control of goal-directed manual actions. Such an understanding is crucial for developing (ultimately directly brain-controlled) avatars for use in virtual reality and telepresence, where sensory action feedback delays are often unavoidable.

Acknowledgments

The authors thank Richard Limanowski, Jan Herding, and Dennis Haupt for technical assistance with the virtual reality environment.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.neuroimage.2016.11.009.