Perceiving your hand moving: BOLD suppression in sensory cortices and the role of the cerebellum in the detection of feedback delays

B. Ezgi Arikan

Department of Psychology,
Justus-Liebig University Giessen, Giessen, Germany ✉

Bianca M. van Kemenade

Department of Psychiatry and Psychotherapy,
Philipps University Marburg, Marburg, Germany

Kornelius Podranski

Department of Psychiatry and Psychotherapy,
Philipps University Marburg, Marburg, Germany
Core Facility Brain Imaging, Faculty of Medicine,
Philipps University Marburg, Marburg, Germany
Max Planck Institute for Human Cognitive and Brain
Sciences, Leipzig, Germany

Olaf Steinsträter

Department of Psychiatry and Psychotherapy,
Philipps University Marburg, Marburg, Germany
Core Facility Brain Imaging, Faculty of Medicine,
Philipps University Marburg, Marburg, Germany

Benjamin Straube

Department of Psychiatry and Psychotherapy,
Philipps University Marburg, Marburg, Germany

Tilo Kircher

Department of Psychiatry and Psychotherapy,
Philipps University Marburg, Marburg, Germany

Sensory consequences of self-generated as opposed to externally generated movements are perceived as less intense and lead to less neural activity in corresponding sensory cortices, presumably due to predictive mechanisms. Self-generated sensory inputs have been mostly studied in a single modality, using abstract feedback, with control conditions not differentiating efferent from reafferent feedback. Here we investigated the neural processing of (a) naturalistic action–feedback associations of (b) self-generated versus externally generated movements, and (c) how an additional (auditory) modality influences neural processing and detection of delays. Participants executed wrist movements using a passive movement device (PMD) as they watched their movements in real time or with variable delays (0–417 ms). The task was to judge whether there was a delay between the movement and its visual feedback. In the externally generated condition, movements were induced by the PMD to disentangle efferent from reafferent feedback. Half of the trials involved auditory beeps coupled to the onset of the visual feedback. We found reduced BOLD activity in visual, auditory, and somatosensory areas during self-generated compared with externally generated movements in unimodal and bimodal conditions. Anterior and posterior cerebellar areas were engaged for trials in which action–feedback delays were detected for self-generated movements. Specifically, the left cerebellar lobule IX was functionally connected with the right superior occipital gyrus. The results indicate efference copy-based predictive mechanisms specific to self-generated movements, leading to BOLD suppression in sensory areas. In addition, our results support the cerebellum's role in the detection of temporal prediction errors during our actions and their consequences.

Introduction

Our actions shape the way we perceive the world around us (Blakemore, Wolpert, & Frith, 1998; Desantis & Haggard, 2016b; Haggard, Clark, & Kalogeras, 2002; Wolpert, Miall, & Kawato, 1998). For example, it has been well established that sensory stimuli associated with self-generated movements are experienced as less intense and produce less neural activity in corresponding sensory cortices than if they were externally delivered (Blakemore, Rees, & Frith, 1998; Blakemore, Wolpert, & Frith, 1999; Roussel, Hughes, & Waszak, 2014; Shergill et al., 2013). This is known as sensory suppression. Recent theories of motor control explain suppression within the forward model framework, in which an internal model anticipates the sensory consequences of self-generated movements via efference copies and monitors whether predicted consequences match actual sensory input (Blakemore, Frith, et al., 1999; Flanagan, Vetter, Johansson, & Wolpert, 2003; Wolpert, Ghahramani, & Flanagan, 2001). Efference copies then modulate neural processing in the corresponding sensory cortex, resulting in reduced sensation for these consequences (Pynn & DeSouza, 2013). Such predictive processing is crucial to discriminate events of external origin from those initiated by our own movement, allowing us to attribute agency over sensory inputs associated with our own actions (Blakemore, Wolpert, & Frith, 2002; Brown, Adams, Parees, Edwards, & Friston, 2013; Hoover & Harris, 2012; Shergill, Samson, Bays, Frith, & Wolpert, 2005).

These studies, along with others investigating sensorimotor prediction of voluntary actions and their consequences, have mostly focused on one sensory modality at a time. In general, sensory suppression has been observed for somatosensory, auditory, and visual domains, both on a behavioral (Bays, Flanagan, & Wolpert, 2006; Cardoso-Leite, Mamassian, Schutz-Bosbach, & Waszak, 2010; Roussel et al., 2014; Sato, 2008) as well as at a neural level (Blakemore, Frith, & Wolpert, 2001; Blakemore, Wolpert, et al., 1998; Desantis, Roussel, & Waszak, 2014; Lange, 2011; Leube et al., 2003; Shergill et al., 2013; Waszak, Cardoso-Leite, & Hughes, 2012). However, our actions in the real world often lead to multiple sensory consequences; i.e., when we knock on a door, we would get visual, auditory, and somatosensory information simultaneously. Moreover, each of these modalities might be affected by the presence of the other. Research on cross-modal processing points to the facilitatory effect of a stimulus modality on another modality when they are spatially and/or temporally in synchrony (Arabzadeh, Clifford, & Harris, 2008; Bresciani, Dammeier, & Ernst, 2008; Diederich & Colonius, 2004; McDonald, Teder-Sälejärvi, & Hillyard, 2000; Teder-Sälejärvi, Di Russo,

McDonald, & Hillyard, 2005). On the other hand, there is evidence that a specific modality can have an inhibitory effect on another modality (Colavita, 1974; Kawashima, O'Sullivan, & Roland, 1995; Meredith, 2002; Morein-Zamir, Soto-Faraco, & Kingstone, 2003; Sinnett, Soto-Faraco, & Spence, 2008; Wang et al., 2012). These context-dependent cross-modal interactions might begin at relatively early stages of information processing in regions known as modality-specific (Calvert, Campbell, & Brammer, 2000; Laurienti et al., 2002; Macaluso, 2006; Macaluso, Frith, & Driver, 2000; Shimojo & Shams, 2001). In the case of action-related effects, studies focusing on the role of cross-modal processing of self-generated movement consequences suggest facilitation of behavioral performance with the inclusion of an additional modality (Desantis & Haggard, 2016a; Desantis, Mamassian, Lisi, & Waszak, 2014; Farrer, Valentin, & Hup, 2013; Kawabe, Roseboom, & Nishida, 2013; van Kemenade, Arikan, Kircher, & Straube, 2016). With respect to neural correlates, our group has previously shown that perceiving single auditory or visual (unimodal) and audiovisual (bimodal) consequences of voluntary button presses lead to reduced BOLD activity in somatosensory, visual, and auditory cortices for self-generated stimuli compared with passive viewing of these stimuli. In addition, participants were more sensitive to delays between their button press and the sensory consequence of the button press in the bimodal condition, suggesting that the additional modality enhanced delay detection for self-generated movements and their consequences (Straube et al., 2017).

Despite these findings, a number of issues remain to be explored that are crucial in advancing our understanding of why the processing of voluntary actions and their sensory consequences are distinct. First, most studies investigating cross-modal perception and action so far used passive viewing of sensory stimuli as a control condition (Desantis & Haggard, 2016a; Desantis, Waszak, Moutsopoulou, & Haggard, 2016; Kawabe et al., 2013; Straube et al., 2017; van Kemenade, Arikan, Kircher, & Straube, 2017). Although this condition might address the influence of self-generated movements on the perception of sensory inputs by factoring out mere sensory processing, the role of efferent feedback cannot be addressed. Externally generated movements, on the other hand, might prove more efficient in examining the role of efference copies specific to self-generated movements. Self-generated sensory stimuli produce both efferent and reafferent information, whereas sensory consequences of externally generated movements provide only reafferent information (Weiskrantz, Elliott, & Darlington, 1971). By contrasting both movements, the influence of efferent information can be identified. Second, actions and their consequences in the real world are more complex than button presses and abstract dots

as action outcomes. Although button presses as well as computer-generated sensory stimuli are easier to control and manipulate, more studies are needed to demonstrate that similar mechanisms apply to real life (Ingram & Wolpert, 2011). Importantly, by examining naturalistic action–feedback associations, one can also take into account ownership of the body, which contributes to action control (Gentile et al., 2011; Hoover & Harris, 2012; Kilteni & Ehrsson, 2017; Leube et al., 2003; Limanowski, Kirilina, & Blankenburg, 2017).

Our main aim in the present study was twofold: to investigate (a) whether naturalistic action–feedback associations produce suppression in corresponding sensory cortices as opposed to movements without efferent feedback and (b) whether cross-modal processing influences detection of temporal prediction errors for and concurrent neural processing of sensory stimuli associated with self-generated movements. More specifically, we aimed to examine BOLD suppression regarding naturalistic visual feedback of self-generated wrist movements and the impact of an additional (auditory) modality on the detection of temporal discrepancies between self-generated movements and their visual feedback. To this end, participants were asked to perform wrist movements using the handle of a custom-made passive movement device (PMD) (see van Kemenade et al., 2019). Half of the trials involved an externally generated movement condition in which the PMD moved the hand automatically. This condition allowed us to address the role of efferent feedback specific to self-generated movements as opposed to reafferent feedback arising as a result of both movements. The movements were recorded with a camera and displayed to the participants in real time, which allowed us to better represent a naturalistic feedback of the action. Temporal discrepancies were induced by presenting systematic delays between the movement and the visual display of the movement. The task of the participants was to report whether they detected a delay between the actual movement and the visual feedback of the movement (unimodal condition). This task would address detection of temporal prediction errors associated with one's own action. Half of the trials involved an auditory beep coupled to the onset of the visual feedback (bimodal condition). The task remained the same in these trials. Based on our previous finding of BOLD suppression (Straube et al., 2017), we expected to observe reduced activity in sensory areas for self-generated compared with externally generated movements. Because unimodal and bimodal conditions would be equally likely, we expected to find reduced BOLD activity in auditory regions in both unimodal and bimodal conditions when the movement was self-generated due to sensorimotor prediction. We also hypothesized that the auditory stimulus triggered by self-generated movements would provide additional information about temporal discrepancies between the movement and its visual feedback, aiding in subjective detection of delays. Finally, we explored possible differences in BOLD activity associated with detecting versus not detecting action–feedback delays across movements and modalities in order to assess whether detection of a temporal discrepancy between the movement and its feedback would lead to distinct brain activation when the movement was self-generated.

Methods

Data from all participants was taken from a recent study by our group in which we examined contributions of different regions in action–feedback monitoring of naturalistic self- versus externally generated movements (van Kemenade et al., 2019). The experimental design and data acquisition procedures have been reported previously in van Kemenade et al. (2019) and restated below for the reader's convenience.

Participants

Twenty-three right-handed students (11 females, mean age = 25.83 ± *SD* 3.09) from Philipps University Marburg participated in the behavioral training and fMRI scanning. One participant who took part in the behavioral training was not included in the fMRI study due to an increased bias to report delays even for the 0-ms delay (see Experimental design and procedure for details). Five participants were excluded from the final analyses due to excessive head movement ($N = 1$) or technical issues ($N = 2$), and insufficient number of subjective delay detection trials ($N = 2$), resulting in a final sample of 18 participants (eight females, age = 26.22 ± *SD* 3.34). Written informed consent was obtained from all participants. Participants reported normal or corrected-to-normal vision and normal hearing as well as no history of psychiatric or neurological disorders and no current use of psychoactive medications. Right-handedness was confirmed by the Edinburgh Handedness Inventory (Oldfield, 1971). The experiment was approved by the local ethics committee and performed in accordance with the Declaration of Helsinki. The participants were paid for their participation.

Apparatus

A custom-made magnetic resonance (MR)-compatible PMD was used for the execution of both self- and
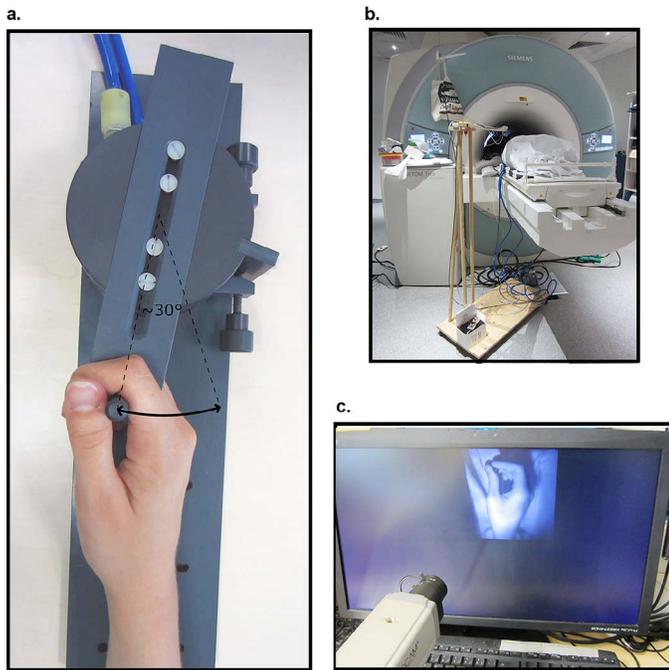
Figure 1. (a) Custom-made PMD. Movement trajectory was horizontal, and the range of movement from one end to another was ~30. (b) Outlook of the experimental setup. MRI-compatible camera recorded and displayed hand movements in real time during self- and externally generated movements. (c) Visual display shown to the participant, comprising the visual feedback of the movement.

externally generated movements. The device was built using nonferromagnetic materials (PVC, titan). A handle was designed for the execution of the movement with a horizontal trajectory. Importantly, the movement range of the handle and, therefore, the to-be-executed movement was restricted so that movement range would remain constant throughout trials and for different movements. Approximate movement range was ~30° and ~5.5 cm (see Figure 1a). Motion to the device could be induced automatically with compressed air (6 bar). Approximate force used when the device was working automatically was 20 N.

An MR-compatible camera (MRC High Speed, MRC Systems GmbH, Heidelberg, Germany) with a 2-ms refresh rate was used to record movements of the hand. The camera was mounted to an adjustable custom-made support. Auditory beeps were presented via MR-compatible headphones (MR-Confon, Optimel, Magdeburg, Germany). Responses were given using MR-compatible button pads (Cedrus, Lumina, San Pedro, CA). Visual display of the movement recorded by the camera in real time was presented on a mirror-projected computer screen (60 Hz refresh rate). Five LEDs were attached to the PMD, not visible to the participant, for the purpose of detecting movement onsets. A motion-detection algorithm written specifi-

cally for this study was applied to the camera images, detecting the spatial position of one moving LED with regard to the four fixed LEDs in each camera frame. For each trial, movement onset was defined as a minimum of 1° change in the relative position of the moving LED within one frame (requiring an increase of at least 0.5° per frame). Delays were presented from movement onset onward. All equipment was controlled by custom written software running on a personal computer (Dell Optiplex 9020, Intel Core i5-4570, four cores, 3.2 GHz, 4 GB RAM).

Experimental design and procedure

An event-related design was used with a delay-detection paradigm as shown in Figure 2. On each trial, the participant was asked to perform a wrist movement (extension and flexion) with the right hand using the handle of the PMD. The movements were recorded by the camera and presented to the participant in real time via mirror-projection, constituting the visual feedback of the movement (see Figure 1b & c). Six delays (0, 83, 167, 250, 330, and 417 ms, determined based on the screen refresh rate) were introduced between the movement and the real-time display of the movement. The participants were asked to judge whether there was a delay between actual movement of their hand and visual feedback of the movement. Half of the trials contained 440-Hz sine wave pure tones (500 ms), which were presented with the same delay used to delay the visual feedback. The task remained the same in these trials.

One week prior to scanning, participants completed a behavioral training in order to familiarize themselves with the task. First, they were trained to perform hand movements both in the self- and externally generated movement conditions. Accordingly, they were instructed to grab the handle of the PMD and move the handle from left to right and then back in approximately 1,500 ms while trying to maintain a constant pace with the help of a metronome. They then practiced externally generated movements by grabbing the handle of the PMD while trying to relax the hand and wrist as much as possible and letting the device move their wrist. In order to prevent participants from using visual cues from the hand, a curtain was placed to hide the right hand from participants' sight. Participants executed both movements first without and then with the visual display recorded by the camera and presented in real time via a computer screen. The participants were instructed to observe the displayed movement on the screen and to judge whether there was a delay between their actual movement and the visual feedback of the movement or not. They were also told that, in some trials, they would additionally receive
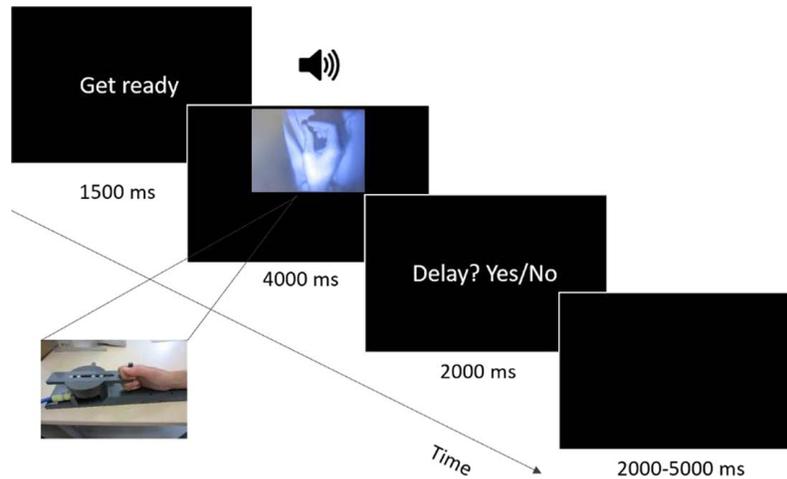
Figure 2. Timeline of an experimental trial. After the "Get ready" cue, the camera was turned on for 4,000 ms, during which the participant had to perform either a self- or an externally generated movement. They were asked to watch a visual display of their movements on a mirror screen and judge whether there was a delay between the actual movement and the visual feedback of the movement. Auditory beeps coupled to the onset of the visual feedback were presented during bimodal trials.

auditory beeps, but this was not related to the task. Participants first completed trials with no delay and a 417-ms delay and received feedback indicating whether they answered correctly. They then completed three runs that were similar in length to the main experimental runs in the scanning sessions. Each run contained a self and an external movement block (miniblocks) in alternating order (144 trials in total). The order of movement was counterbalanced across participants. Because being able to detect the delays was crucial for our paradigm, only participants who detected at least 50% of delays or more at the most extreme delay and who reported detecting a delay less than 50% of the 0-ms delay trials were invited to the fMRI experiment. One participant did not fulfill these criteria and was, therefore, not invited to the fMRI experiment. The training procedure took approximately 1 hr.

In the fMRI experiment, participants were positioned inside the bore in supine position with their right hand placed on the PMD handle, which was located beside them. Each participant received five experimental blocks, each of which contained two miniblocks (self- and externally generated movement conditions, each having 24 trials) in counterbalanced order. At the beginning of each miniblock, the participant was informed about the specific movement they had to perform (self- or externally generated). Each trial started with a cue, "Ready," that lasted for 1,500 ms. This cue indicated that participants had to prepare themselves for the upcoming movement. Immediately after the disappearance of the cue, the camera was turned on for 4,000 ms. This timing frame was chosen to give participants *some freedom* in performing a self-generated, intentional movement. The participant was

instructed to perform the movement (in the self-generated miniblock) or let the device move their hand (in the externally generated miniblock). The onset of the externally generated trial was jittered (500–1,000–1,500 ms). The participants saw their hand movement on the screen in real time or with one of five delays. After a 500-ms interval, the question "Delay?" appeared on the screen. Participants used their left index and middle fingers to respond "Yes" or "No," respectively. Button assignment was counterbalanced across participants. The maximum time allowed for the response was 2,000 ms. Following the response, an intertrial interval of 2,000–5,000 ms was presented. Unimodal visual and bimodal audiovisual trials were randomized within blocks. Total number of trials for each participant was 240. Participants were allowed to have short breaks between the blocks. Each movement was monitored online and recorded for the purpose of ensuring compliance with instructions and for post hoc screening of movement parameters. The scanning procedure took approximately 1 hr.

Analysis

Behavioral data analysis

Statistical analyses (repeated-measures ANOVA and paired samples *t* tests) were performed using SPSS21 (IBM Corp., released 2012). Behavioral data were grouped based on participant's responses regarding delay-detection judgments, referred to as the detection of delays. The proportion of delay responses was calculated for each condition by dividing the number of

trials in which the participant detected a delay between the movement and its visual feedback by the total number of trials in that condition.

fMRI Data Acquisition and Analysis

Functional imaging was carried out in a 3 T TIM Trio scanner (Siemens, Erlangen, Germany), using a 12-channel head coil. A gradient echo EPI sequence was used (TR: 1,650 ms, TE: 25 ms, flip angle: 70°). The whole brain was covered in 34 axial slices (gap: 15%, slice thickness: 4 mm), each of which was acquired as a 64×64 matrix (FoV: 192×192 mm^2, voxel size: $3 \times 3 \times 4.6$ mm) in descending order. Three hundred thirty volumes were obtained during each experimental block. Anatomical images were obtained using a T1-weighted MPRAGE sequence (TR: 1,900 ms, TE: 2.26 ms, flip angle: 9°, FoV: 256×256, slice thickness: 1 mm, matrix: 256×256, voxel size: $1 \times 1 \times 1$ mm).

In order to account for possible stimulus-correlated motion artifacts in the self-generated movement condition, we performed independent components analysis (ICA) on the raw data using the MELODIC tool (Beckmann & Smith, 2004) in FSL (Jenkinson, Beckmann, Behrens, Woolrich, & Smith, 2012; Smith et al., 2004). The data were prepared for the ICA with the implementation of slice-time correction, realignment using the MCFLIRT procedure, and spatial smoothing with an 8-mm full-width-at-half-maximum isotropic Gaussian kernel. After ICA, components that seem to indicate motion artifacts were removed. We also checked whether head motion differences were similar across self- and externally generated movements by conducting statistical analyses on head motion parameters using a freely available R script to estimate scan-to-scan motion of fMRI data (https://github.com/kesslerr/motionEstimator). The script estimates changes in absolute head motion over time by calculating differences on the x-, y-, and z-axes between successive scans and transforming these values into a single scan-to-scan difference (see Mazaika, Hoeft, Glover, & Reiss, 2009; Wilke, 2012). For each run, we initially divided the head motion parameters obtained after realignment based on movement type (head motion parameters during self- vs. externally generated movements). We then calculated maximum scan-to-scan values for each run and movement type. These values were then averaged over all runs for each participant. Finally, we conducted a paired samples t test on the maximum scan-to-scan values averaged across runs for self- and externally generated movements. Our results showed no significant differences in scan-to-scan head motion values between the self-generated and externally-generated movements, $t(17) = 1.76$, $p = 0.97$.

The remaining preprocessing steps and analyses were performed using statistical parametric mapping (SPM12, Wellcome Department of Imaging Neuroscience, University College London, UK) implemented in MATLAB 2012b (MathWorks, Natick, MA). Accordingly, each participant's anatomical image was coregistered to their first functional images, segmented and normalized to the standard Montreal Neurological Institute (MNI) standard space, and resampled to $2 \times 2 \times 2$ mm^3 voxels.

Prior to statistical analyses, we screened recorded video images from all participants to examine whether the movements were properly and/or actually performed. We excluded those trials in which the participant did not execute the instructed movement and bimodal trials in which the movement detection algorithm did not work (1.4% of all trials).

Statistical analyses were performed based on the following purposes: First, we aimed to identify possible BOLD-suppression effects during self-generated movements and the influence of auditory modality on BOLD activity when the movement was self-generated. As a proof of principle, we initially examined BOLD activity during the execution of both movements as these movements have been shown to engage overlapping as well as different brain regions (Ciccarelli et al., 2005; Miall, Christensen, Cain, & Stanley, 2007; Szameitat, Shen, Conforto, & Sterr, 2012; Weiller et al., 1996). We then examined BOLD-suppression effects in the self-generated condition by contrasting it with the externally generated condition. We also explored whether BOLD-suppression effects were similar across different modalities, i.e., when an auditory modality was presented. Second, we assessed possible differences in BOLD activity across movements and modalities for trials in which a delay was detected versus undetected. Here, we were interested in examining neural responses for detected versus undetected delays for self-generated movements and their sensory consequences and whether such processing would be modulated by the presence of an additional modality. Importantly, delay detection could be confounded by the amount of physical delay; i.e., participants are more likely to detect action–feedback discrepancies when these delays are large. In our statistical analyses, we also attempted to account for this confound by adopting a parametric approach in which both delay detection performance (detected vs. undetected) and the effect of physical delay were accounted for (see below for details).

For the statistical analyses, we implemented a standard summary statistic approach (Friston et al., 1995; Holmes & Friston, 1998; Worsley & Friston, 1995). This method rests on a mass univariate general linear model (GLM) in which time series from each voxel are modeled as a linear combination of experimental variables (effects of interest) plus an error term

(Friston et al., 1995; Pernet, 2014). For both analyses, time series data from each participant and session were initially modeled using a design matrix that consisted of six regressors: four experimental conditions time-locked to the onset of camera and two additional regressors modeling the cue and response periods, respectively. In addition, we included two parametric modulators for each of the four experimental conditions in the following order: physical delay (delay in milliseconds for that trial) and delay detection (defined as a binary variable: 1 = detected, 0 = not detected). As our focus was on BOLD activity linked with subjective delay detection, we put the regressor for physical delay first, to which SPM attributes higher variance to physical delay by orthogonalizing subjective delay detection with regard to the physical delay. This way, we aimed to capture variance associated specifically with the subjective detection of delay (Mumford, Poline, & Poldrack, 2015). The regressors were convolved with the canonical hemodynamic response function employed by SPM. In addition, six realignment parameters were included as nuisance regressors to account for residual motion artifacts. Time series data were high-pass filtered with a cutoff frequency of 128 s. Temporal autocorrelation was modeled by an autoregressive model of order 1 (AR(1)). After this model was estimated, parameter estimates from regressors of interest were used to compute contrast images corresponding to eight conditions of interest: self-generated unimodal ($Self_{Uni}$), self-generated bimodal ($Self_{Bi}$), externally generated unimodal (Ext_{Uni}), externally generated bimodal (Ext_{Bi}), delay detection (detected vs. undetected) during self-generated unimodal ($DetSelf_{Uni}$), delay detection during self-generated bimodal ($DetSelf_{Bi}$), delay detection during externally generated unimodal ($DetExt_{Uni}$), delay detection during externally generated bimodal ($DetExt_{Bi}$). These contrast images were passed to a second-level random effects flexible-factorial analysis. As a proof of principle, a conjunction of self- and externally generated movements was calculated to identify areas commonly activated during both movements (($Self_{Uni} + Self_{Bi}$) ∩ ($Ext_{Uni} + Ext_{Bi}$)). We also examined activity differences in the self- compared with externally generated movements using the contrast [($Self_{Uni} + Self_{Bi}$) − ($Ext_{Uni} + Ext_{Bi}$)]. BOLD suppression effects were assessed by conducting a *t* test for externally versus self-generated conditions using the contrast [($Ext_{Uni} + Ext_{Bi}$) − ($Self_{Uni} + Self_{Bi}$)]. Moreover, we examined whether the same regions showed reduced BOLD activity during self-generated movements in both unimodal and bimodal conditions with a conjunction analysis [($Ext_{Uni} − Self_{Uni}$) ∩ ($Ext_{Bi} − Self_{Bi}$)]. Differences in BOLD activity linked with the presentation of additional modality were tested by contrasting bimodal with the unimodal condition [($Self_{Bi} + Ext_{Bi}$) −

($Self_{Uni} + Ext_{Uni}$)]. Because we did not have a specific hypothesis regarding the direction of the interaction between movement and modality, an *F* test was initially conducted to see whether there was any interaction between movement and modality, which would presumably be followed by a *t* test when necessary. Two *t* tests were conducted to assess whether BOLD activity linked with delay detection (detected vs. undetected) differed when the movement was self- as opposed to externally generated. These were [($DetSelf_{Uni} + DetSelf_{Bi}$) − ($DetExt_{Uni} + DetExt_{Bi}$)] for self-generated movements and [($DetExt_{Uni} + DetExt_{Bi}$) − ($DetSelf_{Uni} + DetSelf_{Bi}$)] for externally generated movements. The influence of auditory modality on BOLD activity related to the detection of delays were examined using the contrasts [($DetSelf_{Uni} + DetExt_{Uni}$) − ($DetSelf_{Bi} + DetExt_{Bi}$)] for unimodal versus bimodal and [($DetSelf_{Bi} + DetExt_{Bi}$) − ($DetSelf_{Uni} + DetExt_{Uni}$)] for bimodal versus unimodal conditions. We also examined whether activity associated with detection of delays was different when the movement was self-generated and in the presence of the auditory modality. The respective contrasts were [($DetSelf_{Uni} − DetSelf_{Bi}$) − ($DetExt_{Uni} − DetExt_{Bi}$)] and [($DetSelf_{Bi} − DetSelf_{Uni}$) − ($DetExt_{Bi} − DetExt_{Uni}$)].

We additionally assessed psychophysiological interactions (PPIs) between posterior cerebellar regions, namely bilateral cerebellar lobules IX and visual areas in the brain as a function of delay detection and movement type (see Results section for details). We restricted our region of interest to the posterior cerebellum, namely bilateral cerebellar lobules IX, as this region has been found to be highly involved in the visual guidance of movement (Stoodley & Schmahmann, 2010). We initially concatenated each participant's session data using our original GLM on subjective delay detection, in which the number of sessions was modeled as additional block regressors. Accordingly, for each participant, we used two mask images of the left and right cerebellar lobules IX using the automated anatomical labeling (AAL) atlas (Tzourio-Mazoyer et al., 2002) in order to look at the time course of each voxel. We then extracted the first eigenvariate from all voxels within a 4-mm radius of the seed region based on the local maxima of the participant within the mask and adjusted by the effects of interest. Extracted eigenvariates were passed to a GLM at the subject level, at which signal time course from left and right cerebellar lobules IX, our effects of interest (movement, modality) along with parametric modulators defining physical delay and delay detection, their psychological factors and PPI regressors were modeled. We also included cue and responses as well as movement parameters as regressors of no interest. "Detected > undetected" parametric modulators (i.e., the parameter estimates of the interaction term) of
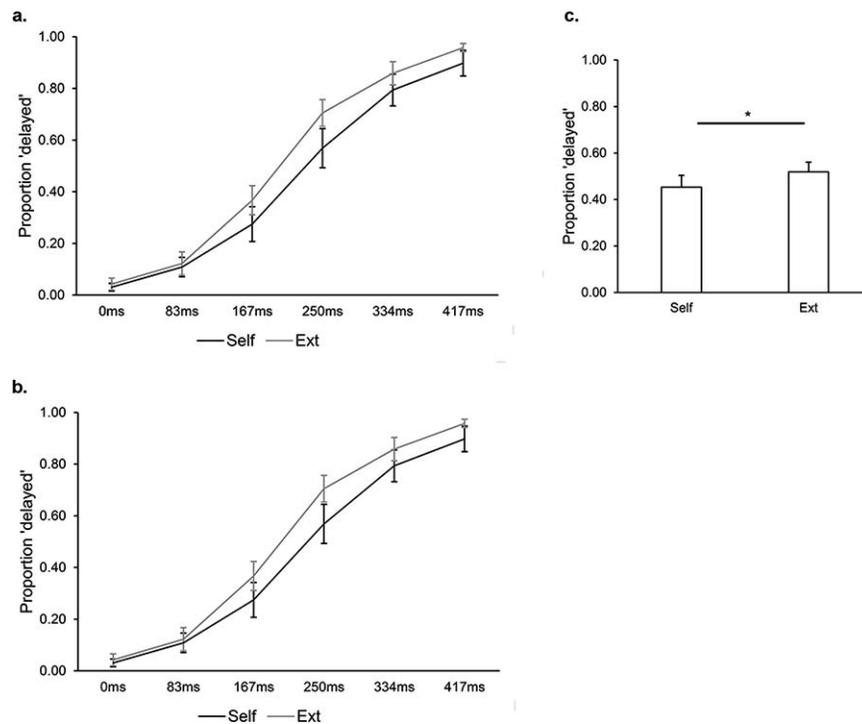
Figure 3. Plots illustrating (a) proportion of delay responses for each delay as a function of movement for unimodal and (b) bimodal trials and (c) significant effect of externally generated movement with regard to delay detection compared with self-generated movements. Error bars denote standard error of the mean ($\pm SEM$). See also Supplementary Figure S1 for individual data from a representative participant.

respective PPI regressors were then passed to a random effects flexible-factorial GLM at the group level.

To correct for multiple comparisons at the whole-brain level, we used family-wise error (FWE) correction at $p < 0.05$ based on Gaussian random field theory (Worsley, Evans, Marrett, & Neelin, 1992; Worsley et al., 1996) implemented in SPM12. In addition, to avoid type II errors, Monte Carlo simulations were run using the estimated smoothness of our functional data (14.6 mm). This resulted in a minimum cluster size of 88 resampled voxels to achieve correction for multiple comparisons at $p < 0.05$, assuming an individual voxel type I error of $p = 0.001$ (Slotnick, Moo, Segal, & Hart, 2003). This method also allows one to detect moderate effects (Forman et al., 1995; Slotnick et al., 2003). The AAL atlas (Tzourio-Mazoyer et al., 2002) and SPM Anatomy Toolbox (Eickhoff et al., 2005) were used to label significant activations based on peak activation voxels. Visualization of significant activations was performed by SPM12 and MRICron (Rorden & Brett, 2000), and GraphPad Prism version 7.00 for Windows (GraphPad Software, La Jolla, CA). Effect sizes were reported as mean beta estimates using the MarsBar toolbox (Brett, Anton, Valabregue, & Poline, 2006). We calculated the beta values separately for each participant and then obtained an average value for each effect of interest and participant.

Results

Behavioral data

Figure 3a and b displays proportion of delay responses as a function of movement for unimodal and bimodal conditions, respectively (see also Supplementary Figure S1). A 2 (Movement: Self- vs. Externally generated) × 2 (Modality: Unimodal vs. Bimodal) repeated-measures ANOVA was conducted on the proportion of delay responses. There was a main effect of movement, $F(1, 19) = 6.83$, $p < 0.05$, $\eta^2 = 0.21$. Accordingly, the proportion of delay responses was significantly higher in the externally generated ($M = 0.52$, $SEM = 0.03$) than in the self-generated ($M = 0.45$, $SEM = 0.04$) condition (Figure 3c). There was no main effect of modality, $F(1, 19) = 2.24$, $p = 0.15$, nor was there an interaction, $F(1, 19) = 0.08$, $p = 0.77$.

In order to test whether movement duration was similar across self- and externally generated movements, we used recorded trials to determine the onset and offset of each movement (18 participants). A post hoc paired-samples t test showed that average movement durations across self- and externally generated trials were not significantly different from each other, $t(17) = 1.23$, $p = 0.23$; average movement duration for
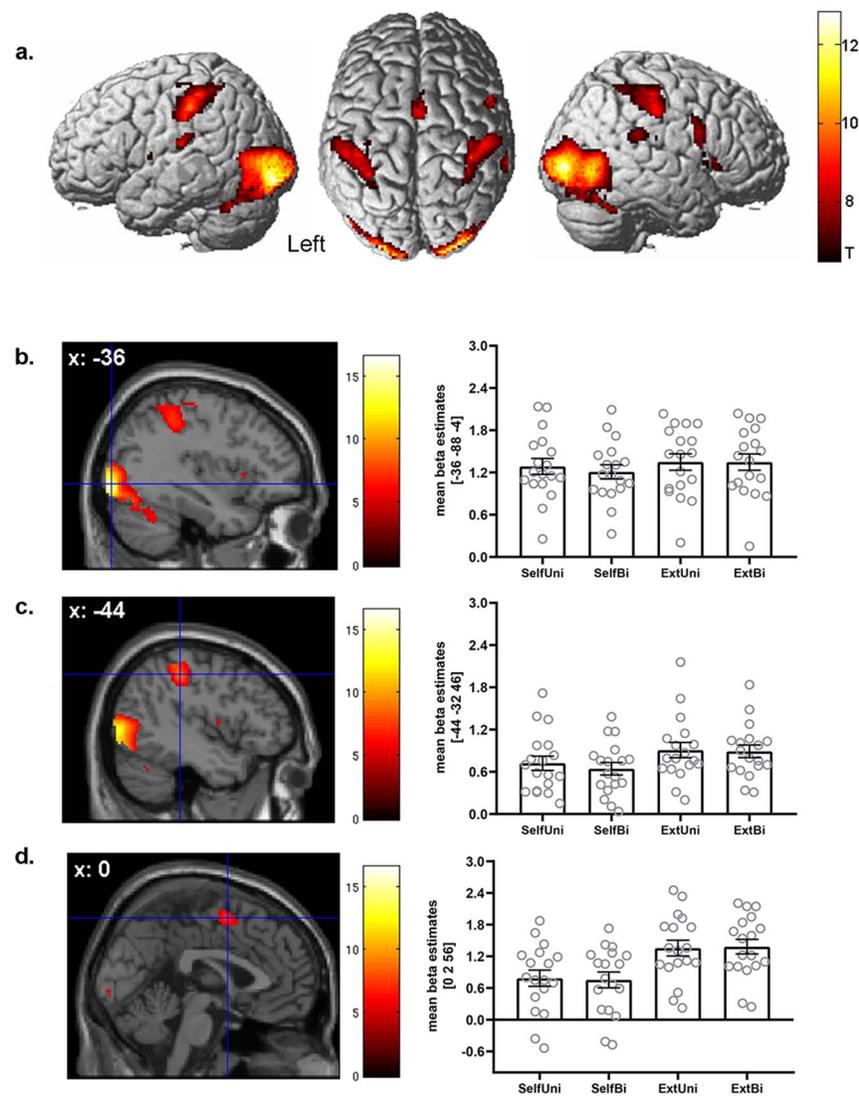
Figure 4. fMRI results showing (a) commonly activated regions during self- and externally generated movements and corresponding plots of (b) left inferior occipital area (MNI coordinates: *x, y, z* = −36, −88, −4), (b) left postcentral gyrus (MNI coordinates: *x, y, z* = −44, −32, 46), and (c) left supplementary motor area (MNI coordinates: *x, y, z* = 0, 2, 56) overlaid on a standardized T1-weighted image. Dots and bar graphs show mean beta estimates from individual participants and across conditions (±*SEM*) for the corresponding peak voxel, respectively. All maps were thresholded at $p_{FWE} < 0.05$. See also Supplementary Figure S2 for individual data from a representative participant.

self-generated movements (*M* = 1,373.82, *SEM* = 81.84), average movement duration for externally generated movements (*M* = 1,282.78, *SEM* = 44.22).

fMRI data

We first explored common activations during self- and externally generated conditions by using a conjunction of self- and externally generated movements that were contrasted against an implicit baseline. We found that both movements led to activity in visual and somatosensory areas, including bilateral somatosensory areas along with bilateral inferior and middle occipital gyri. In addition, both movements led to activations in premotor and motor areas, inferior frontal and parietal areas, and bilateral thalamus and insula (see Figure 4, and Supplementary Figure S2). The anatomical locations of these clusters are shown in Table 1. Examination of common areas across different movements and modalities revealed similar results.

BOLD-suppression effects were assessed by contrasting externally with self-generated conditions. The contrast revealed bilateral activations in somatosensory cortices; middle and superior occipital gyri, including the V5; and middle and superior temporal gyri (MTG and STG, respectively). Other areas involved premotor and motor regions, inferior parietal lobules (IPL),

Anatomical label	Hemisphere	Coordinates			t value	No. voxels
		x	y	z		
Occipital mid R	Right	36	−90	2	16.53	2,835
Temporal mid R	Right	48	−72	−2	13.46	
Calcarine R	Right	20	−100	2	13.18	
Occipital mid L	Left	−36	−88	−4	14.99	2,518
Calcarine L	Left	−14	−94	−6	13.65	
Occipital mid L	Left	−30	−96	2	13.63	
Postcentral R	Right	50	−24	48	8.85	692
Postcentral R	Right	44	−30	54	7.8	
Postcentral R	Right	32	−36	46	6	
Postcentral L	Left	−44	−32	46	8.49	841
Parietal inf L	Left	−50	−28	50	8.09	
Postcentral L	Left	−44	−36	54	7.94	
Supp motor area L	Left	0	2	56	7.11	183
Supp motor area R	Right	4	14	50	4.8	
Temporal sup R	Right	64	−36	22	6.64	79
Frontal inf oper R	Right	50	8	22	5.77	136
Precentral R	Right	52	6	34	5.54	
Supramarginal L	Left	−50	−26	20	5.72	110
Supramarginal L	Left	−60	−22	16	5.46	
Insula R	Right	36	20	6	5.48	93
Thalamus L	Left	−12	−18	6	5.39	26
Rolandic oper L	Left	−46	−2	8	5.15	9
Cerebellum VIII L	Left	−12	−70	−44	5.06	41
Cerebellum VIII L	Left	−20	−66	−48	4.94	
Insula L	Left	−34	18	4	4.92	8
Frontal inf oper L	Right	54	14	4	4.84	6
Rolandic oper R	Right	56	−22	20	4.71	3
Frontal inf oper R	Right	56	12	10	4.71	2

Table 1. Anatomical locations of common peak activations during self- and externally generated movements ($p_{FWE} < 0.05$, height threshold = 4.67).

precuneus, thalamus, and insula (see Figure 5a and Supplementary Figure S3a). Anatomical locations of these clusters are shown in Table 2.

We also identified common BOLD-suppression effects in sensory areas across unimodal and bimodal conditions. Significant BOLD-suppression effects were found in bilateral primary and secondary somatosensory cortices, middle and inferior occipital gyri, MTG, and STG. Additional activations were observed in supplementary motor area (SMA), bilateral IPL, putamen, angular gyri, supramarginal gyri, and thalamus (see Figure 5b and Supplementary Figure S3b). Contrasting self- with externally generated movements did not lead to significant activations (both at $p_{FWE} < 0.05$ and $p < 0.001$ uncorrected level).

As expected, bimodal compared with unimodal trials led to significant activations in bilateral temporal cortices with two large clusters in bilateral STG (MNI coordinates: $x, y, z = 64, −24, 10, T = 13.76, k_E = 2,429, p_{FWE} < 0.05$; MNI coordinates: $x, y, z = −50, −26, 8, T$ = 11.53, $k_E = 1,588, p_{FWE} < 0.05$). However, when relative strength of these activations was examined separately for each movement, activations in the self-generated bimodal condition were weaker than in the externally generated bimodal condition. Figure 6 illustrates relative BOLD amplitudes in these areas across different movements and modalities (see also Supplementary Figure S4). Using a mask created from the ($[\text{Self}_{Bi} + \text{Ext}_{Bi}] − [\text{Self}_{Uni} + \text{Ext}_{Uni}]$) contrast, we further conducted a t test ($[\text{Ext}_{Bi} − \text{Self}_{Bi}]$) to confirm whether activation in auditory areas were significantly higher in the externally than in the self-generated condition. We found significant activations in bilateral STG (MNI coordinates: $x, y, z = 46, −32, 22, T = 8.95, k_E = 1,747, p_{FWE} < 0.05$; MNI coordinates: $x, y, z = −44, −32, 18, T = 7.65, k_E = 986, p_{FWE} < 0.05$). Finally, there were no significant interactions between movement and modality.

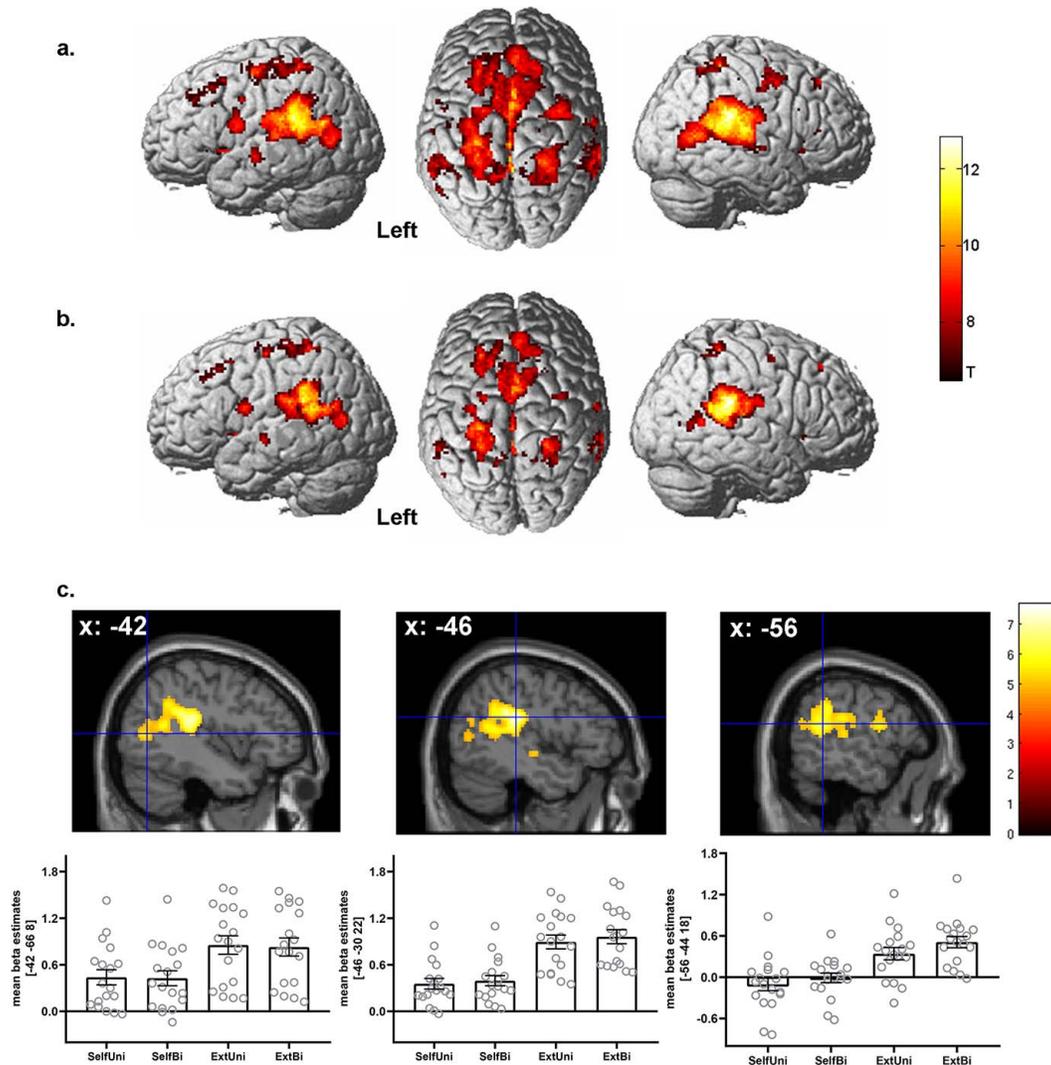Our second analysis regarding delay detection resulted in a main effect. First, delay detection

Figure 5. fMRI results showing (a) BOLD suppression during self-generated movements compared with externally generated movements, (b) commonly suppressed regions across unimodal and bimodal trials during self- versus externally generated movements. Plots of (c) left rolandic operculum (MNI coordinates: x, y, $z = -46$, -30, 22), (b) left middle occipital gyrus (MNI coordinates: x, y, $z = -42$, -66, 8), and (c) left superior temporal gyrus (MNI coordinates: x, y, $z = -56$, -44, 18) showing commonly suppressed regions across unimodal and bimodal trials during self- versus externally generated movements in overlaid on a standardized T1-weighted image. Dots and bar graphs show mean beta estimates from individual participants and across conditions ($\pm SEM$) for the corresponding peak voxel, respectively. All maps were thresholded at $p_{FWE} < 0.05$. See also Supplementary Figure S3 for individual data from a representative participant.

(detected vs. undetected trials) during self- as opposed to externally generated movements revealed bilateral regions in the anterior and posterior cerebellar areas including bilateral cerebellar lobules IV–VI (MNI coordinates: x, y, $z = 24$, -44, -36, $T = 4.31$; MNI coordinates: x, y, $z = 12$, -46, -20, $T = 4.32$) as well as lobules IX and the cerebellar vermis (MNI coordinates: x, y, $z = -2$, -36, -38, $T = 4.51$, $k_E = 1,592$, $p < 0.05$ cluster corrected; see Figure 7 and Supplementary Figure S5). We found no significant activations when we contrasted externally with self-generated movements. Finally, there were no interaction effects.

In order to examine whether activity linked with detecting versus not detecting a delay in the cerebellum was correlated with activity in visual areas during self-generated movements, we employed a PPI analysis. PPI employs a simple regression method in which task-related (psychological factor) modulations in the time series of a seed region (physiological factor) can be investigated in terms of connectivity changes in other voxels of the brain (Friston et al., 1997). Our hypothesis was that, if the cerebellum were specifically involved in detecting discrepancies between self-generated movements and their visual feedback, increased

Anatomical label	Hemisphere	Coordinates			t value	No. voxels
		x	y	z		
Supramarginal R	Right	44	−32	24	8.57	3,408
Temporal sup R	Right	58	−36	20	7.83	
Temporal sup R	Right	48	−38	12	6.74	
Rolandic oper L	Left	−46	−30	22	8.25	3,317
Supramarginal L	Left	−44	−38	26	7.78	
Temporal sup L	Left	−56	−44	18	6.84	
Precentral L	Left	−28	−28	60	7.26	7,567
Supp motor area L	Left	−6	4	58	7.08	
Supp motor area R	Right	8	6	52	6.78	
Insula R	Right	30	22	8	6.78	230
Precentral L	Left	−54	2	16	6.09	160
Insula L	Left	−30	20	2	5.98	584
Putamen L	Left	−24	6	6	5.67	
Putamen L	Left	−22	18	0	5.62	
Frontal mid R	Right	36	−4	52	5.65	258
Precentral R	Right	42	−8	46	5.11	
Temporal sup L	Left	−48	−14	−6	5.51	94
Cuneus R	Right	20	−68	36	5.2	57
Insula L	Left	−30	12	−12	5.13	6
Lingual R	Right	26	−68	−2	5.04	25
Thalamus L	Left	−6	−16	0	4.93	18
Precentral L	Left	−46	−2	36	4.88	18
Frontal sup R	Right	20	−8	64	4.87	4
Postcentral L	Left	−48	−16	50	4.86	5
Precentral R	Right	54	−12	42	4.84	5
Frontal mid L	Left	−36	20	32	4.84	9
Cerebellum crus II R	Right	14	−82	−36	4.83	8
Putamen R	Right	26	−2	14	4.82	4
Frontal inf tri R	Right	36	16	28	4.8	18
Postcentral L	Left	−44	−10	50	4.79	1
Frontal mid R	Right	30	44	14	4.77	2
Precentral R	Right	20	−18	66	4.74	1
Occipital mid L	Left	−32	−72	28	4.73	1
Putamen R	Right	24	12	4	4.72	4
Thalamus R	Right	10	−14	0	4.67	1

Table 2. Anatomical locations of peak activations during externally versus self-generated movements (Ext − Self contrast) ($p_{FWE} < 0.05$, height threshold = 4.67).

connectivity between this area and the visual cortex would be evident only for self-generated movements. Due to our a priori hypothesis, the connectivity analysis was restricted to visual areas previously found to be involved in perceiving visual feedback of one's own movement (Leube et al., 2003). Accordingly, after an initial cluster defining threshold of $p < 0.001$, we applied peak-level family-wise error small volume correction within a 10-mm radius region of interest (ROI) centered on the coordinates corresponding to the right and left occipital gyri reported in Leube et al. (2003). As a statistical threshold, we chose $p_{FWE} < 0.025$ because we additionally applied a Bonferroni correction for two tests (a PPI for the left and right

cerebellar lobules IX as seed, respectively). The analysis revealed a significant cluster in the right superior occipital gyrus (SOG; MNI coordinates: $x, y, z = 18, -94, 14$, $T = 3.65$, $k_E = 15$, small volume correction $p_{FWE} = 0.014$; see Figure 8). This suggests that activity in the right superior occipital gyrus correlated significantly with activity in the left cerebellar lobule IX when a delay was detected during self- as opposed to externally generated movements. For the right cerebellar lobule IX, the analysis revealed a region in the left SOG; however, this did not reach significance after correcting for the number of tests performed ($p_{FWE} = 0.038$).
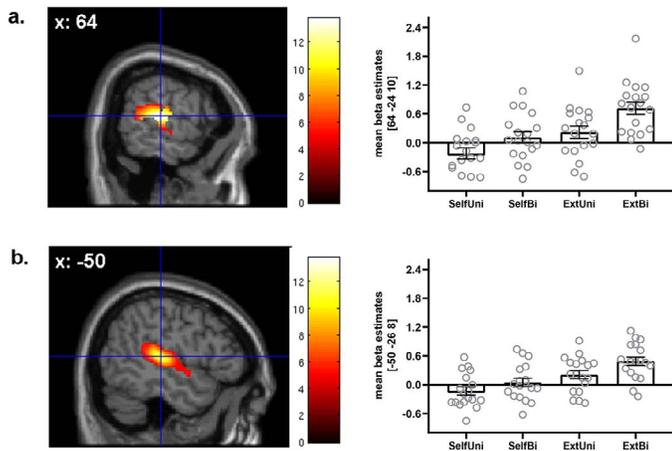
Figure 6. Main effect of auditory modality in the (a) right STG (MNI coordinates: *x, y, z* = 64, −24, 10; $p_{FWE} < 0.05$) and the (b) left STG (MNI coordinates: *x, y, z* = −50, −26, 8; $p_{FWE} < 0.05$) overlaid on a standardized T1-weighted image. Dots and bar graphs show mean beta estimates from individual participants and across conditions (±*SEM*) for the corresponding peak voxel, respectively. See also Supplementary Figure S4 for data from a single representative participant.

Discussion

In the present study, we assessed neural correlates of perceiving self-generated movements and resulting visual feedback and whether and how an additional (auditory) modality would alter corresponding processing in the brain associated with these movements. By contrasting self- with externally generated movements of the wrist, we were able to address efference copy-based predictive processes. Moreover, our experimental setup allowed us to examine these processes with naturalistic action–feedback associations, which are usually more complex than simple button presses. Behaviorally, we found significantly higher proportion of delay responses for externally compared with self-generated movements. fMRI analysis revealed activations in several motor-related and sensory areas for both movements. As hypothesized, we found reduced BOLD activity in somatosensory, visual, and auditory regions of the brain during self- compared with externally generated movements, demonstrating BOLD suppression in sensory cortices. This effect was similar across unimodal and bimodal conditions, implying predictive processing and subsequent suppression for self-generated sensory inputs. Detection of delays was found to be linked with increased recruitment of bilateral anterior and posterior regions of the cerebellum including lobules IV–VI as well as lobules IX when the movement was self-generated as opposed to when it was externally generated. PPI analysis showed enhanced functional connectivity in the left cerebellar lobule IX with right SOG, further supporting cerebellum's involvement in detection of discrepancies between self-generated movements and accompanying visual feedback. Together, these new findings extend previous findings regarding BOLD suppression in sensory cortices associated with voluntary movements
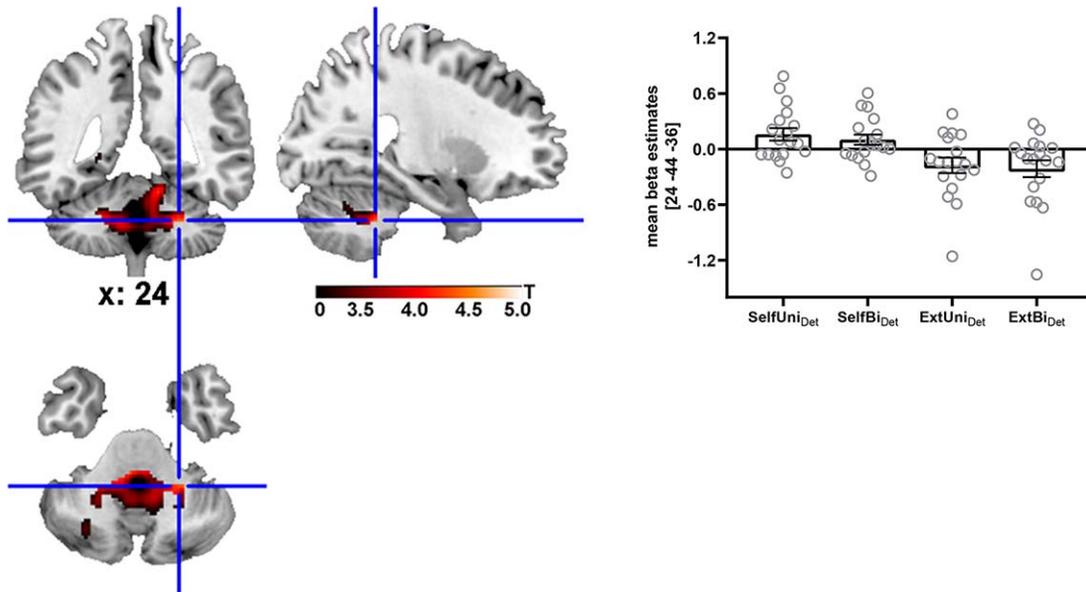

Figure 7. Significant activation cluster in the cerebellum including bilateral cerebellum IV–VI and bilateral cerebellum IX (peak MNI coordinates: *x, y, z* = 24, −44, −36) associated with detecting a delay during self- compared with externally generated movements (*p* < 0.05 cluster corrected) overlaid on a standardized T1-weighted image. Dots and bar graphs show mean beta estimates from individual participants and across conditions (±*SEM*) for the corresponding peak voxel. See also Supplementary Figure S5 for data from a single representative participant.
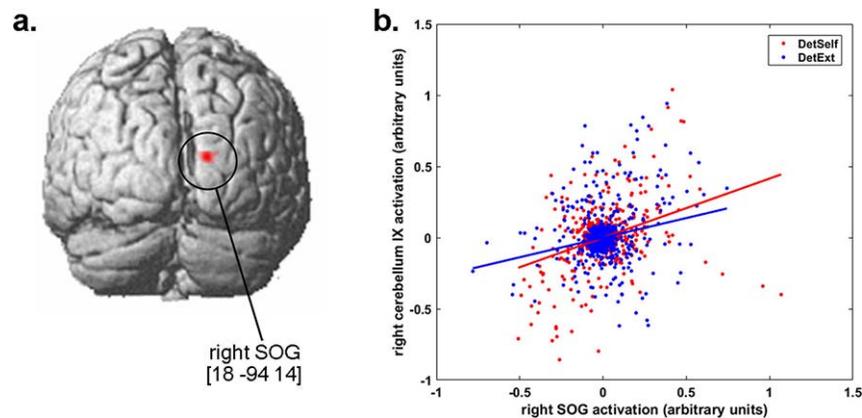
Figure 8. Psychophysiological interactions. (a) SOG (MNI coordinates: *x*, *y*, *z* = 18, −94, 14) showing functional connectivity with the left cerebellum IX (MNI coordinates: *x*, *y*, *z* = −6, −52, −32) in detected versus undetected trials during self- movements compared with externally generated movements (cluster defining threshold $p < 0.001$, small volume correction at the peak level $p_{FWE} < 0.05$ within a priori specified ROI). (b) Scatterplot from a single participant's data illustrating correlation between the left cerebellum IX and the right SOG for self- and externally generated movements.

to the processing of naturalistic action–feedback associations and specifically as a result of efference copy-based processes. Moreover, our results support the relevance of the cerebellum in the detection of delays between self-generated movements and accompanying visual feedback. We now discuss our findings in detail.

BOLD suppression in sensory cortices during self-generated movements

Both self- and externally-generated conditions led to activations in areas involved in the planning and control of movements, such as the SMA, M1, thalamus, and insula. These results are in line with existing evidence showing similar activations in premotor and motor structures during self- and externally generated movements (Balslev, Nielsen, Lund, Law, & Paulson, 2006; Ciccarelli et al., 2005; Miall et al., 2007; Szameitat et al., 2012; Weiller et al., 1996). As hypothesized, self-generated movements, in contrast with externally generated movements led to reduced BOLD activity in somatosensory and visual regions. The BOLD-suppression effect was specific to higher order visual areas, including V5, which has been implicated in the processing of visual motion (Zeki et al., 1991). This result points to a highly specific downregulation of visual feedback associated with self-generated movements. Together, our results suggest increased predictive processing of sensory signals associated with self-generated movements compared to when these signals were externally generated.

Strikingly, BOLD-suppression effects were not only evident in somatosensory and visual areas, but also in the auditory cortex in both unimodal and bimodal

conditions. In other words, activity in auditory cortices was suppressed during self-generated movements independent of whether an auditory stimulus occurred or not. This is similar to our previous finding, demonstrating suppressed BOLD signal in auditory cortices during the presentation of audiovisual and visual-only stimuli (Straube et al., 2017). These findings can be explained in terms of expectation-based processing. Expected sensory inputs have been shown to result in reduced neural processing, implying a predictive processing for such inputs (Alink, Schwiedrzik, Kohler, Singer, & Muckli, 2010; den Ouden, Daunizeau, Roiser, Friston, & Stephan, 2010; Kok & de Lange, 2015; Kok, Jehee, & de Lange, 2012; Murray, Kersten, Olshausen, Schrater, & Woods, 2002). In our study, participants were likely to anticipate the auditory stimulus, which led to reduced BOLD activity in auditory cortices because unimodal and bimodal trials were intermixed. This effect should be specific for self-generated trials as opposed to externally generated trials as we propose that sensory inputs associated with our own actions are more likely to be expected due to increased control we can exert on these inputs, which led to BOLD suppression in corresponding sensory areas. In other words, sensory consequences of self-generated movements might have led to heightened expectation and subsequent down-regulation for these consequences (Blakemore, Frith, et al., 1999; Blakemore, Wolpert, et al., 1998; Blakemore, Wolpert, & Frith, 2000; Wolpert, 2007; Wolpert & Flanagan, 2001). Our results are in line with these findings and suggest that people might be more ready to form expectations about upcoming sensory inputs when they are in control of the likely consequences.

Suppression effects observed in sensory regions are in line with our previous study (Straube et al., 2017)

along with several other studies demonstrating reduced activity in sensory areas linked with one's own action (Benazet, Thénault, Whittingstall, & Bernier, 2016; Blakemore, Frith, & Wolpert, 1999; Shergill et al., 2013). Together, these findings support the notion of efference copy-based predictions in processing sensory consequences of voluntary actions (Blakemore et al., 2000; Cui et al., 2014; Wolpert et al., 2001; Wolpert et al., 1998). Externally generated movements also yielded increased activity in the cingulate cortex, MTG, precuneus, and putamen compared with self-generated movements. The MTG, precuneus, and putamen have been implicated in time perception (Matell & Meck, 2004; Stevens, Kiehl, Pearlson, & Calhoun, 2007) and processing of visuomotor incongruencies (Balslev et al., 2006; David et al., 2007; Dreher & Grafman, 2002; Leube et al., 2003; MacDonald & Paus, 2003). Moreover, externally generated movements led to increased BOLD activity in the IPL, an area found to be relevant in integrating visuomotor signals (Balslev et al., 2006; Farrer et al., 2008; Limanowski et al., 2017; Schnell et al., 2007), distinction between self- and other-generated actions (Blakemore & Frith, 2003; Jackson, 2004; Leube et al., 2003; Shimada, Hiraki, & Oda, 2005; Weiller et al., 1996), and mental imagery and complex actions (Blakemore & Sirigu, 2003; Haggard, 2009; Shergill et al., 2013; Sirigu, Daprati, Pradat-Diehl, Franck, & Jeannerod, 1999). Our results, therefore, provide support for the involvement of these areas in processing visuomotor incongruencies, possibly contributing to self–other distinction.

Cerebellum's role in the detection of delays associated with self-generated movements and accompanying feedback

fMRI analysis on contrasting trials in which a delay was detected versus not detected revealed two important findings. First, regions of the bilateral anterior and posterior cerebellum, including lobules IV–V and IX, were found to be significantly engaged for trials in which a delay was detected versus not detected during self-generated movements. This is in line with the notion of the cerebellum as a state estimator, modulating predicted sensory consequences of motor commands and providing error signal for the unpredicted consequences of these commands (Blakemore & Frith, 2003; Blakemore et al., 2001; Blakemore, Wolpert, et al., 1999; Leube et al., 2003; Miall, Weir, Wolpert, & Stein, 1993; Miall et al., 2007; van Kemenade et al., 2019; Wolpert et al., 1998). Our finding of higher involvement in the cerebellar areas for subjectively delayed trials during self- compared with externally-generated movements not only supports the cerebellum's role as generating predictions and detecting

errors during sensorimotor processing, but it also provides a subjective component to the involvement of the cerebellum in processing these discrepancies. It has been claimed that awareness of action–feedback discrepancies might not be available to the cerebellum (Blakemore & Sirigu, 2003; Farrer et al., 2008). In this sense, our finding suggests the opposite, that the cerebellum might be involved in the detection of discrepancies between predicted and actual feedback. This was further confirmed by our PPI analysis demonstrating increased functional connectivity between the left cerebellum IX and the right SOG for self-compared with externally generated movements when a delay was detected.

Limitations and future directions

In the current study, we restricted the movement range by the PMD in order to have self- and externally generated movements as similar as possible regarding movement kinematics. We also trained participants prior to the fMRI experiment and monitored their movements throughout the fMRI sessions. Moreover, we tracked movement recordings of 18 participants to check compliance with the instructions and found no significant differences in the average movement duration across movements. Nevertheless, it could be that self-generated movements might be more demanding to execute than externally generated movements, which might, in turn, influence attentional orienting to accompanying sensory stimuli (Hughes, Desantis, & Waszak, 2012).

Second, our main hypothesis concerned BOLD suppression in sensory cortices during self-generated movements as opposed to externally generated movements. We, therefore, focused on differences in sensory regions. However, our main effect also revealed increased BOLD activity in premotor and motor structures. Together with the possibility that self-generated movements might influence processes other than motor control, such as attention, future experiments might benefit from including these movements alone as baseline conditions or scrutinizing different components that define volition, such as what or when to perform an action in terms of their neural correlates (Brass & Haggard, 2008; Hoffstaedter, Grefkes, Zilles, & Eickhoff, 2013).

Another potential limitation of the current study concerns BOLD activity associated with delay detection judgments. We aimed to overcome the covariance of physical delay and delay detection by accounting for the effect of physical delay first and attributing only the remaining variance to subjective delay detection (Mumford et al., 2015). Despite this, it is not entirely possible to disentangle the effect of physical delays

from that of delay detection judgments (Wearden & Jones, 2007). Therefore, our results with regard to detection of delays must be interpreted with caution.

We did not find a behavioral advantage in bimodal conditions, which goes against previous findings on voluntary actions and multisensory processing (Arabzadeh et al., 2008; Desantis & Haggard, 2016b; Desantis, Mamassian, et al., 2014; van Kemenade et al., 2016). This could be explained by the fact that the auditory stimulus in the current study was not perceived as part of the movement because the visual feedback was naturalistic, whereas the auditory stimulus consisted of a beep, which was much more abstract. We chose a simple sine wave instead of a naturalistic sound recording because a sound recording would have introduced unwanted additional delays between action and its feedback due to additional computational processing. However, this means, at the same time, that we cannot rule out the possibility that this might have reduced integration of the sound to the movement and its visual feedback. Another possible explanation lies in the different movement and feedback types between the different studies. In previous studies in which multisensory processing enhanced behavioral performance, the feedback occurred at the end of the movement, i.e., a button press leading to a visual dot or a beep. Here, the visual feedback occurred during movement. In their study, David, Skoruppa, Gulberti, Schultz, and Engel (2016) found that participants were more sensitive to outcome- as opposed to movement-related feedback. It could, therefore, be the case that multisensory processing aids particularly behavioral performance when the feedback is outcome-related and less when it is movement-related to which participants are generally less sensitive. We suggest future studies thoroughly investigate the interplay between different types of movements and multiple sensory modalities associated with these movements as well as the role of the naturalness of the sensory feedback.

Although perception of the sensory consequences triggered by voluntary movements has been extensively investigated, studies using naturalistic action–feedback associations that would better mimic real-world situations are limited. A study by Benazet et al. (2016) focused on this aspect by using visual feedback of a moving limb. In line with our results, they found that cortical processing of visual reafferent signals are suppressed, but only when the signals are correctly predicted. We believe that integrating real-world situations to experimental paradigms is of crucial importance, and more studies are needed to investigate natural movements executed in everyday life. In addition, recent investigations have found contradictory evidence for the reduced sensation and neural processing of self-generated movement consequences (Mifsud et al., 2016; Reznik, Henkin, Levy, & Mukamel, 2015; Reznik, Henkin, Schadel, & Mukamel, 2014; Roussel, Hughes, & Waszak, 2013; Yon & Press, 2017). Therefore, more studies are needed to clarify mechanisms involved in perceiving sensory inputs associated with our own actions, whether and how predictions are influenced by these actions, and the impact of cross-modal processing on subsequent perceptual experience.

Conclusions

In conclusion, we investigated for the first time action-related BOLD suppression during self-generated movements leading to unimodal visual versus bimodal audiovisual feedback using naturalistic action–feedback associations. Externally generated movements allowed us to effectively distinguish the influence of efference copy-based predictive processes from reafferent feedback. We found reduced BOLD activity in somatosensory, visual, and auditory regions of the brain during self- compared with externally generated movements and across different modalities, indicating heightened suppression of upcoming multisensory stimuli associated with self-generated movements. Detection of delays was found to be linked with increased cerebellar processing specifically for self-generated movements. Together, these findings extend previous accounts regarding BOLD suppression to realistic action–outcome associations and for multiple sensory inputs associated with these actions. Furthermore, our results highlight the cerebellum's role in the detection of temporal discrepancies between voluntary actions and accompanying feedback, supporting its relevance in action–perception coupling.

Keywords: self-generated movements, suppression, cross-modal, cerebellum

Acknowledgments

Corresponding author: B. Ezgi Arikan.
Address: Department of Psychology, Justus-Liebig University Giessen, Giessen, Germany.
Email: arikan.ezgi@gmail.com.

References

Alink, A., Schwiedrzik, C. M., Kohler, A., Singer, W., & Muckli, L. (2010). Stimulus predictability reduces responses in primary visual cortex. *Journal of Neuroscience*, *30*(8), 2960–2966, https://doi.org/10.1523/JNEUROSCI.3730-10.2010.

Arabzadeh, E., Clifford, C. W. G., & Harris, J. A. (2008). Vision merges with touch in a purely tactile discrimination. *Psychological Science*, *19*(7), 635–641, https://doi.org/10.1111/j.1467-9280.2008.02134.x.

Balslev, D., Nielsen, F. Å., Lund, T. E., Law, I., & Paulson, O. B. (2006). Similar brain networks for detecting visuo-motor and visuo-proprioceptive synchrony. *NeuroImage*, *31*(1), 308–312, https://doi.org/10.1016/j.neuroimage.2005.11.037.

Bays, P. M., Flanagan, J. R., & Wolpert, D. M. (2006). Attenuation of self-generated tactile sensations is predictive, not postdictive. *PLoS Biology*, *4*(2), 281–284, https://doi.org/10.1371/journal.pbio.0040028.

Beckmann, C. F., & Smith, S. M. (2004). Probabilistic independent component analysis for functional magnetic resonance imaging. *IEEE Transactions on Medical Imaging*, *23*(2), 137–152, https://doi.org/10.1109/TMI.2003.822821.

Benazet, M., Thénault, F., Whittingstall, K., & Bernier, P.-M. (2016). Attenuation of visual reafferent signals in the parietal cortex during voluntary movement. *Journal of Neurophysiology*, *116*(4), 1831–1839, https://doi.org/10.1152/jn.00231.2016.

Blakemore, S.-J., & Frith, C. (2003). Self-awareness and action. *Current Opinion in Neurobiology*, *13*(2), 219–224, https://doi.org/10.1016/S0959-4388(03)00043-6.

Blakemore, S.-J., Frith, C. D., & Wolpert, D. M. (1999). Spatio-temporal prediction modulates the perception of self-produced stimuli. *Journal of Cognitive Neuroscience*, *11*(5), 551–559, https://doi.org/10.1162/089892999563607.

Blakemore, S.-J., Frith, C. D., & Wolpert, D. M. (2001). The cerebellum is involved in predicting the sensory consequences of action. *Neuroreport*, *12*(9), 1879–1884, https://doi.org/10.1097/00001756-200107030-00023.

Blakemore, S.-J., Rees, G., & Frith, C. D. (1998). How do we predict the consequences of our actions? A functional imaging study. *Neuropsychologia*, *36*(6), 521–529, https://doi.org/10.1016/S0028-3932(97)00145-0.

Blakemore, S.-J., & Sirigu, A. (2003). Action prediction in the cerebellum and in the parietal lobe. *Experimental Brain Research*, *153*(2), 239–245, https://doi.org/10.1007/s00221-003-1597-z.

Blakemore, S.-J., Wolpert, D. M., & Frith, C. D. (1998). Central cancellation of self-produced tickle sensation. *Nature Neuroscience*, *1*(7), 635–640, https://doi.org/10.1038/2870.

Blakemore, S.-J., Wolpert, D. M., & Frith, C. D. (1999). The cerebellum contributes to somatosensory cortical activity during self-produced tactile stimulation. *NeuroImage*, *10*(4), 448–459, https://doi.org/10.1006/nimg.1999.0478.

Blakemore, S. J., Wolpert, D., & Frith, C. (2000). Why can't you tickle yourself? *Neuroreport*, *11*(11), R11–R16.

Blakemore, S. J., Wolpert, D. M., & Frith, C. D. (2002). Abnormalities in the awareness of action. *Trends in Cognitive Sciences*, *6*(6), 237–242.

Brass, M., & Haggard, P. (2008). The what, when, whether model of intentional action. *The Neuroscientist*, *14*(4), 319–325, https://doi.org/10.1177/1073858408317417.

Bresciani, J.-P., Dammeier, F., & Ernst, M. O. (2008). Tri-modal integration of visual, tactile and auditory signals for the perception of sequences of events. *Brain Research Bulletin*, *75*(6), 753–760, https://doi.org/10.1016/j.brainresbull.2008.01.009.

Brett, M., Anton, J.-L., Valabregue, R., & Poline, J.-B. (2006, January 1). Region of interest analysis using an SPM toolbox. [Abstract]. Presented at the 8th International Conference on Functional Mapping of the Human Brain, June 2–6, 2002, Sendai, Japan.

Brown, H., Adams, R. A., Parees, I., Edwards, M., & Friston, K. (2013). Active inference, sensory attenuation and illusions. *Cognitive Processing*, *14*(4), 411–427, https://doi.org/10.1007/s10339-013-0571-3.

Calvert, G. A., Campbell, R., & Brammer, M. J.

(2000). Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. *Current Biology: CB*, *10*(11), 649–657.

Cardoso-Leite, P., Mamassian, P., Schutz-Bosbach, S., & Waszak, F. (2010). A new look at sensory attenuation: Action-effect anticipation affects sensitivity, not response bias. *Psychological Science*, *21*(12), 1740–1745, https://doi.org/10.1177/0956797610389187.

Ciccarelli, O., Toosy, A. T., Marsden, J. F., Wheeler-Kingshott, C. M., Sahyoun, C., Matthews, P. M., . . . Thompson, A. J. (2005). Identifying brain regions for integrative sensorimotor processing with ankle movements. *Experimental Brain Research*, *166*(1), 31–42, https://doi.org/10.1007/s00221-005-2335-5.

Colavita, F. B. (1974). Human sensory dominance. *Perception & Psychophysics*, *16*(2), 409–412, https://doi.org/10.3758/BF03203962.

Cui, F., Arnstein, D., Thomas, R. M., Maurits, N. M., Keysers, C., & Gazzola, V. (2014). Functional magnetic resonance imaging connectivity analyses reveal efference-copy to primary somatosensory area, BA2. *PLoS One*, *9*(1): e84367, https://doi.org/10.1371/journal.pone.0084367.

David, N., Cohen, M. X., Newen, A., Bewernick, B. H., Shah, N. J., Fink, G. R., & Vogeley, K. (2007). The extrastriate cortex distinguishes between the consequences of one's own and others' behavior. *NeuroImage*, *36*(3), 1004–1014, https://doi.org/10.1016/j.neuroimage.2007.03.030.

David, N., Skoruppa, S., Gulberti, A., Schultz, J., & Engel, A. K. (2016). The sense of agency is more sensitive to manipulations of outcome than movement-related feedback irrespective of sensory modality. *PLoS One*, *11*(8): e0161156, https://doi.org/10.1371/journal.pone.0161156.

den Ouden, H. E. M., Daunizeau, J., Roiser, J., Friston, K. J., & Stephan, K. E. (2010). Striatal prediction error modulates cortical coupling. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *30*(9), 3210–3219, https://doi.org/10.1523/JNEUROSCI.4458-09.2010.

Desantis, A., & Haggard, P. (2016a). Action-outcome learning and prediction shape the window of simultaneity of audiovisual outcomes. *Cognition*, *153*, 33–42, https://doi.org/10.1016/j.cognition.2016.03.009.

Desantis, A., & Haggard, P. (2016b). How actions shape perception: Learning action-outcome relations and predicting sensory outcomes promote

audio-visual temporal binding. *Scientific Reports*, *6*(1): 39086, https://doi.org/10.1038/srep39086.

Desantis, A., Mamassian, P., Lisi, M., & Waszak, F. (2014). The prediction of visual stimuli influences auditory loudness discrimination. *Experimental Brain Research*, *232*(10), 3317–3324, https://doi.org/10.1007/s00221-014-4001-2.

Desantis, A., Roussel, C., & Waszak, F. (2014). The temporal dynamics of the perceptual consequences of action-effect prediction. *Cognition*, *132*(3), 243–250, https://doi.org/10.1016/j.cognition.2014.04.010.

Desantis, A., Waszak, F., Moutsopoulou, K., & Haggard, P. (2016). How action structures time: About the perceived temporal order of action and predicted outcomes. *Cognition*, *146*, 100–109, https://doi.org/10.1016/j.cognition.2015.08.011.

Diederich, A., & Colonius, H. (2004). Bimodal and trimodal multisensory enhancement: Effects of stimulus onset and intensity on reaction time. *Perception & Psychophysics*, *66*(8), 1388–1404, https://doi.org/10.3758/BF03195006.

Dreher, J.-C., & Grafman, J. (2002). The roles of the cerebellum and basal ganglia in timing and error prediction. *The European Journal of Neuroscience*, *16*(8), 1609–1619.

Eickhoff, S. B., Stephan, K. E., Mohlberg, H., Grefkes, C., Fink, G. R., Amunts, K., & Zilles, K. (2005). A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *NeuroImage*, *25*(4), 1325–1335, https://doi.org/10.1016/j.neuroimage.2004.12.034.

Farrer, C., Frey, S. H., Van Horn, J. D., Tunik, E., Turk, D., Inati, S., & Grafton, S. T. (2008). The angular gyrus computes action awareness representations. *Cerebral Cortex*, *18*(2), 254–261, https://doi.org/10.1093/cercor/bhm050.

Farrer, C., Valentin, G., & Hup, J. M. (2013). The time windows of the sense of agency. *Consciousness and Cognition*, *22*(4), 1431–1441, https://doi.org/10.1016/j.concog.2013.09.010.

Flanagan, R. R., Vetter, P., Johansson, R. S., & Wolpert, D. M. (2003). Prediction precedes control in motor learning. *Current Biology*, *13*(2), 146–150, https://doi.org/10.1016/S0960-9822(03)00007-1.

Forman, S. D., Cohen, J. D., Fitzgerald, M., Eddy, W. F., Mintun, M. A., & Noll, D. C. (1995). Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. *Magnetic Resonance in Medicine*, *33*(5), 636–647.

Friston, K. J., Buechel, C., Fink, G., Morris, J., Rolls, E., & Dolan, R. (1997). Psychophysiological and

Arikan et al.

modulatory interactions in neuroimaging. *Neuro-Image*, *6*(3), 218–229, https://doi.org/10.1006/nimg.1997.0291.

Friston, K. J., Holmes, A. P., Poline, J.-B., Grasby, P. J., Williams, S. C. R., Frackowiak, R. S. J., & Turner, R. (1995). Analysis of fMRI time-series revisited. *NeuroImage*, *2*(1), 45–53, https://doi.org/10.1006/nimg.1995.1007.

Gentile, G., Petkova, V. I., Ehrsson, H. H., Gentile, G., Petkova, V. I., & Ehrsson, H. H. (2011). Integration of visual and tactile signals from the hand in the human brain: An fMRI study. *Journal of Neurophysiology*, *105*(2), 910–922, https://doi.org/10.1152/jn.00840.2010.

Haggard, P. (2009, May 8). The sources of human volition. *Science*, *324*(5928), 731–733, https://doi.org/10.1126/science.1173827.

Haggard, P., Clark, S., & Kalogeras, J. (2002). Voluntary action and conscious awareness. *Nature Neuroscience*, *5*(4), 382–385, https://doi.org/10.1038/nn827.

Hoffstaedter, F., Grefkes, C., Zilles, K., & Eickhoff, S. B. (2013). The "what" and "when" of self-initiated movements. *Cerebral Cortex*, *23*(3), 520–530, https://doi.org/10.1093/cercor/bhr391.

Holmes, A., & Friston, K. (1998, January 1). Generalisability, random effects and population inference. *Neuroimage*, *7*(4 (2/3)), S754. *Proceedings of Fourth International Conference on Functional Mapping of the Human Brain*, June 7-12 1998, Montreal, Canada.

Hoover, A. E. N., & Harris, L. R. (2012). Detecting delay in visual feedback of an action as a monitor of self recognition. *Experimental Brain Research*, *222*(4), 389–397, https://doi.org/10.1007/s00221-012-3224-3.

Hughes, G., Desantis, A., & Waszak, F. (2012). Mechanisms of intentional binding and sensory attenuation: The role of temporal prediction, temporal control, identity prediction, and motor prediction. *Psychological Bulletin*, *139*(1), 133–151, https://doi.org/10.1037/a0028566.

IBM Corp. (2012). IBM SPSS Statistics for Windows, Version 21.0. Armonk, NY: IBM Corp.

Ingram, J. N., & Wolpert, D. M. (2011). Naturalistic approaches to sensorimotor control. *Progress in Brain Research*, *191*, 3–29, https://doi.org/10.1016/B978-0-444-53752-2.00016-3.

Jackson, P. (2004). Motor cognition: A new paradigm to study self–other interactions. *Current Opinion in Neurobiology*, *14*(2), 259–263, https://doi.org/10.1016/j.conb.2004.01.020.

Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W., & Smith, S. M. (2012). FSL. *NeuroImage*, *62*(2), 782–790, https://doi.org/10.1016/j.neuroimage.2011.09.015.

Kawabe, T., Roseboom, W., & Nishida, S. (2013). The sense of agency is action-effect causality perception based on cross-modal grouping. *Proceedings of the Royal Society B: Biological Sciences*, *280*(1763): 20130991, https://doi.org/10.1098/rspb.2013.0991.

Kawashima, R., O'Sullivan, B. T., & Roland, P. E. (1995). Positron-emission tomography studies of cross-modality inhibition in selective attentional tasks: Closing the "mind's eye." *Proceedings of the National Academy of Sciences, USA*, *92*(13), 5969–5972.

Kilteni, K., & Ehrsson, H. H. (2017). Sensorimotor predictions and tool use: Hand-held tools attenuate self-touch. *Cognition*, *165*, 1–9, https://doi.org/10.1016/j.cognition.2017.04.005.

Kok, P., & de Lange, F. P. (2015). Predictive coding in sensory cortex. In B. U. Forstmann & E.-J. Wagenmakers (Eds.), *An Introduction to Model-Based Cognitive Neuroscience* (pp. 221–244). New York, NY: Springer, https://doi.org/10.1007/978-1-4939-2236-9_11.

Kok, P., Jehee, J. F. M., & de Lange, F. P. (2012). Less is more: Expectation sharpens representations in the primary visual cortex. *Neuron*, *75*(2), 265–270, https://doi.org/10.1016/j.neuron.2012.04.034.

Lange, K. (2011). The reduced N1 to self-generated tones: An effect of temporal predictability? *Psychophysiology*, *48*(8), 1088–1095, https://doi.org/10.1111/j.1469-8986.2010.01174.x.

Laurienti, P. J., Burdette, J. H., Wallace, M. T., Yen, Y.-F., Field, A. S., & Stein, B. E. (2002). Deactivation of sensory-specific cortex by cross-modal stimuli. *Journal of Cognitive Neuroscience*, *14*(3), 420–429, https://doi.org/10.1162/089892902317361930.

Leube, D. T., Knoblich, G., Erb, M., Grodd, W., Bartels, M., & Kircher, T. T. J. (2003). The neural correlates of perceiving one's own movements. *NeuroImage*, *20*(4), 2084–2090.

Limanowski, J., Kirilina, E., & Blankenburg, F. (2017). Neuronal correlates of continuous manual tracking under varying visual movement feedback in a virtual reality environment. *NeuroImage*, *146*, 81–89, https://doi.org/https://doi.org/10.1016/j.neuroimage.2016.11.009.

Macaluso, E. (2006). Multisensory processing in sensory-specific cortical areas. *The Neuroscientist*, *12*(4), 327–338, https://doi.org/10.1177/1073858406287908.

Macaluso, E., Frith, C. D., & Driver, J. (2000, August 18). Modulation of human visual cortex by cross-modal spatial attention. *Science, 289*(5482), 1206–1208.

MacDonald, P. A., & Paus, T. (2003). The role of parietal cortex in awareness of self-generated movements: A transcranial magnetic stimulation study. *Cerebral Cortex,13*(9), 962–967.

Matell, M. S., & Meck, W. H. (2004). Cortico-striatal circuits and interval timing: Coincidence detection of oscillatory processes. *Cognitive Brain Research, 21*(2), 139–170, https://doi.org/10.1016/j.cogbrainres.2004.06.012.

Mazaika, P. K., Hoeft, F., Glover, G. H., & Reiss, A. L. (2009). Methods and software for fMRI analysis of clinical subjects. *NeuroImage, 47*, S58, https://doi.org/10.1016/S1053-8119(09)70238-1.

McDonald, J. J., Teder-Sälejärvi, W. A., & Hillyard, S. A. (2000, October 19). Involuntary orienting to sound improves visual perception. *Nature, 407*(6806), 906–908, https://doi.org/10.1038/35038085.

Meredith, M. A. (2002). On the neuronal basis for multisensory convergence: A brief overview. *Brain Research. Cognitive Brain Research, 14*(1), 31–40.

Miall, R. C., Weir, D. J., Wolpert, D. M., & Stein, J. F. (1993). Is the cerebellum a Smith Predictor? *Journal of Motor Behavior, 25*(3), 203–216, https://doi.org/10.1080/00222895.1993.9942050.

Miall, R. C., Christensen, L. O. D., Cain, O., & Stanley, J. (2007). Disruption of state estimation in the human lateral cerebellum. *PLoS Biology, 5*(11), 2733–2744, https://doi.org/10.1371/journal.pbio.0050316.

Mifsud, N. G., Oestreich, L. K. L., Jack, B. N., Ford, J. M., Roach, B. J., Mathalon, D. H., & Whitford, T. J. (2016). Self-initiated actions result in suppressed auditory but amplified visual evoked components in healthy participants. *Psychophysiology, 53*(5), 723–732, https://doi.org/10.1111/psyp.12605.

Morein-Zamir, S., Soto-Faraco, S., & Kingstone, A. (2003). Auditory capture of vision: Examining temporal ventriloquism. *Brain Research. Cognitive Brain Research, 17*(1), 154–163.

Mumford, J. A., Poline, J.-B., & Poldrack, R. A. (2015). Orthogonalization of regressors in FMRI models. *PLoS One, 10*(4): e0126255, https://doi.org/10.1371/journal.pone.0126255.

Murray, S. O., Kersten, D., Olshausen, B. A., Schrater, P., & Woods, D. L. (2002). Shape perception reduces activity in human primary visual cortex. *Proceedings of the National Academy of Sciences,*

USA, 99(23), 15164–15169, https://doi.org/10.1073/pnas.192579399.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9*(1), 97–113.

Pernet, C. R. (2014). Misconceptions in the use of the general linear model applied to functional MRI: A tutorial for junior neuro-imagers. *Frontiers in Neuroscience, 8*: 1, https://doi.org/10.3389/fnins.2014.00001.

Pynn, L. K., & DeSouza, J. F. X. (2013). The function of efference copy signals: Implications for symptoms of schizophrenia. *Vision Research, 76*, 124–133, https://doi.org/10.1016/j.visres.2012.10.019.

Reznik, D., Henkin, Y., Levy, O., & Mukamel, R. (2015). Perceived loudness of self-generated sounds is differentially modified by expected sound intensity. *PLoS One, 10*(5), 4–9, https://doi.org/10.1371/journal.pone.0127651.

Reznik, D., Henkin, Y., Schadel, N., & Mukamel, R. (2014). Lateralized enhancement of auditory cortex activity and increased sensitivity to self-generated sounds. *Nature Communications, 5*: 4059, https://doi.org/10.1038/ncomms5059.

Rorden, C., & Brett, M. (2000). Stereotaxic display of brain lesions. *Behavioural Neurology, 12*(4), 191–200.

Roussel, C., Hughes, G., & Waszak, F. (2013). A preactivation account of sensory attenuation. *Neuropsychologia, 51*(5), 922–929, https://doi.org/10.1016/j.neuropsychologia.2013.02.005.

Roussel, C., Hughes, G., & Waszak, F. (2014). Action prediction modulates both neurophysiological and psychophysical indices of sensory attenuation. *Frontiers in Human Neuroscience, 8*: 115, https://doi.org/10.3389/fnhum.2014.00115.

Sato, A. (2008). Action observation modulates auditory perception of the consequence of others' actions. *Consciousness and Cognition, 17*(4), 1219–1227, https://doi.org/10.1016/j.concog.2008.01.003.

Schnell, K., Heekeren, K., Schnitker, R., Daumann, J., Weber, J., Hesselmann, V., . . . Gouzoulis-Mayfrank, E. (2007). An fMRI approach to particularize the frontoparietal network for visuomotor action monitoring: Detection of incongruence between test subjects' actions and resulting perceptions. *NeuroImage, 34*(1), 332–341, https://doi.org/10.1016/j.neuroimage.2006.08.027.

Shergill, S. S., Samson, G., Bays, P. M., Frith, C. D., & Wolpert, D. M. (2005). Evidence for sensory prediction deficits in schizophrenia. *American Journal of Psychiatry, 162*(12), 2384–2386, https://doi.org/10.1176/appi.ajp.162.12.2384.

Shergill, S. S., White, T. P., Joyce, D. W., Bays, P. M., Wolpert, D. M., & Frith, C. D. (2013). Modulation of somatosensory processing by action. *NeuroImage*, *70*, 356–362, https://doi.org/10.1016/j.neuroimage.2012.12.043.

Shimada, S., Hiraki, K., & Oda, I. (2005). The parietal role in the sense of self-ownership with temporal discrepancy between visual and proprioceptive feedbacks. *NeuroImage*, *24*(4), 1225–1232, https://doi.org/10.1016/j.neuroimage.2004.10.039.

Shimojo, S., & Shams, L. (2001). Sensory modalities are not separate modalities: Plasticity and interactions. *Current Opinion in Neurobiology*, *11*(4), 505–509.

Sinnett, S., Soto-Faraco, S., & Spence, C. (2008). The co-occurrence of multisensory competition and facilitation. *Acta Psychologica*, *128*(1), 153–161, https://doi.org/10.1016/j.actpsy.2007.12.002.

Sirigu, A., Daprati, E., Pradat-Diehl, P., Franck, N., & Jeannerod, M. (1999). Perception of self-generated movement following left parietal lesion. *Brain: A Journal of Neurology*, *122*(Pt. 10), 1867–1874.

Slotnick, S. D., Moo, L. R., Segal, J. B., & Hart, J. (2003). Distinct prefrontal cortex activity associated with item memory and source memory for visual shapes. *Brain Research. Cognitive Brain Research*, *17*(1), 75–82.

Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., . . . Matthews, P. M. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*, *23*, S208–S219, https://doi.org/10.1016/j.neuroimage.2004.07.051.

Stevens, M. C., Kiehl, K. A., Pearlson, G., & Calhoun, V. D. (2007). Functional neural circuits for mental timekeeping. *Human Brain Mapping*, *28*(5), 394–408, https://doi.org/10.1002/hbm.20285.

Stoodley, C. J., & Schmahmann, J. D. (2010). Evidence for topographic organization in the cerebellum of motor control versus cognitive and affective processing. *Cortex*, *46*(7), 831–844, https://doi.org/10.1016/j.cortex.2009.11.008.

Straube, B., van Kemenade, B. M., Arikan, B. E., Fiehler, K., Leube, D. T., Harris, L. R., & Kircher, T. (2017). Predicting the multisensory consequences of one's own action: BOLD suppression in auditory and visual cortices. *PLoS One*, *12*(1): e0169131, https://doi.org/10.1371/journal.pone.0169131.

Szameitat, A. J., Shen, S., Conforto, A., & Sterr, A. (2012). Cortical activation during executed, imagined, observed, and passive wrist movements in healthy volunteers and stroke patients. *Neuro-Image*, *62*(1), 266–280, https://doi.org/10.1016/j.neuroimage.2012.05.009.

Teder-Sälejärvi, W. A., Di Russo, F., McDonald, J. J., & Hillyard, S. A. (2005). Effects of spatial congruity on audio-visual multimodal integration. *Journal of Cognitive Neuroscience*, *17*(9), 1396–1409, https://doi.org/10.1162/0898929054985383.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., . . . Joliot, M. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage*, *15*(1), 273–289, https://doi.org/10.1006/nimg.2001.0978.

van Kemenade, B. M., Arikan, B. E., Kircher, T., & Straube, B. (2016). Predicting the sensory consequences of one's own action: First evidence for multisensory facilitation. *Attention, Perception, & Psychophysics*, *78*(8), 2515–2526, https://doi.org/10.3758/s13414-016-1189-1.

van Kemenade, B. M., Arikan, B. E., Kircher, T., & Straube, B. (2017). The angular gyrus is a supramodal comparator area in action-outcome monitoring. *Brain Structure and Function*, *222*(8), 3691–3703, https://doi.org/10.1007/s00429-017-1428-9.

van Kemenade, B. M., Arikan, B. E., Podranski, K., Steinsträter, O., Kircher, T., & Straube, B. (2019). Distinct roles for the cerebellum, angular gyrus, and middle temporal gyrus in action–feedback monitoring. *Cerebral Cortex*, *29*(4), 1520–1531, https://doi.org/10.1093/cercor/bhy048.

Wang, W. Y., Hu, L., Valentini, E., Xie, X. B., Cui, H. Y., & Hu, Y. (2012). Dynamic characteristics of multisensory facilitation and inhibition. *Cognitive Neurodynamics*, *6*(5), 409–419, https://doi.org/10.1007/s11571-012-9197-x.

Waszak, F., Cardoso-Leite, P., & Hughes, G. (2012). Action effect anticipation: Neurophysiological basis and functional consequences. *Neuroscience & Biobehavioral Reviews*, *36*(2), 943–959, https://doi.org/10.1016/j.neubiorev.2011.11.004.

Wearden, J. H., & Jones, L. A. (2007). Is the growth of subjective time in humans a linear or nonlinear function of real time? *Quarterly Journal of Experimental Psychology*, *60*(9), 1289–1302, https://doi.org/10.1080/17470210600971576.

Weiller, C., Jüptner, M., Fellows, S., Rijntjes, M., Leonhardt, G., Kiebel, S., . . . Thilmann, A. F. (1996). Brain representation of active and passive movements. *NeuroImage*, *4*(2), 105–110, https://doi.org/10.1006/nimg.1996.0034.

Weiskrantz, L., Elliott, J., & Darlington, C. (1971,

April 30). Preliminary observations on tickling oneself. *Nature*, *230*(5296), 598–599.

Wilke, M. (2012). An alternative approach towards assessing and accounting for individual motion in fMRI timeseries. *NeuroImage*, *59*(3), 2062–2072, https://doi.org/10.1016/j.neuroimage.2011.10.043.

Wolpert, D. M. (2007). Probabilistic models in human sensorimotor control. *Human Movement Science*, *26*(4), 511–524, https://doi.org/10.1016/j.humov.2007.05.005.

Wolpert, D. M., & Flanagan, J. R. (2001). Motor prediction. *Current Biology: CB*, *11*(18), R729–R732.

Wolpert, D. M., Ghahramani, Z., & Flanagan, J. R. (2001). Perspectives and problems in motor learning. *Trends in Cognitive Sciences*, *5*(11), 487–494.

Wolpert, D. M., Miall, R. C., & Kawato, M. (1998). Internal models in the cerebellum. *Trends in Cognitive Sciences*, *2*(9), 338–347, https://doi.org/10.1016/S1364-6613(98)01221-2.

Worsley, K. J., Evans, A. C., Marrett, S., & Neelin, P. (1992). A three-dimensional statistical analysis for CBF activation studies in human brain. *Journal of Cerebral Blood Flow & Metabolism*, *12*(6), 900–918, https://doi.org/10.1038/jcbfm.1992.127.

Worsley, K. J., & Friston, K. J. (1995). Analysis of fMRI time-series revisited—Again. *NeuroImage*, *2*(3), 173–181, https://doi.org/10.1006/nimg.1995.1023.

Worsley, K. J., Marrett, S., Neelin, P., Vandal, A. C., Friston, K. J., & Evans, A. C. (1996). A unified statistical approach for determining significant signals in images of cerebral activation. *Human Brain Mapping*, *4*(1), 58–73, https://doi.org/10.1002/(SICI)1097-0193(1996)4:1<58::AID-HBM4>3.0.CO;2-O.

Yon, D., & Press, C. (2017). Predicted action consequences are perceptually facilitated before cancellation. *Journal of Experimental Psychology: Human Perception and Performance*, *43*(6), 1073–1083, https://doi.org/10.1037/xhp0000385.

Zeki, S., Watson, J. D., Lueck, C. J., Friston, K. J., Kennard, C., & Frackowiak, R. S. (1991). A direct demonstration of functional specialization in human visual cortex. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *11*(3), 641–649.