Predictive coding in a multisensory path integration task: An fMRI study

Milosz Krala

Department of Neurophysics, University of Marburg, Marburg, Germany
Center for Mind, Brain and Behavior–CMBB, University of Marburg and Justus-Liebig-University Giessen, Germany

✉

Bianca van Kemenade

Center for Mind, Brain and Behavior–CMBB, University of Marburg and Justus-Liebig-University Giessen, Germany
Translational Neuroimaging Marburg, Department of Psychiatry and Psychotherapy, University of Marburg, Marburg, Germany

Benjamin Straube

Center for Mind, Brain and Behavior–CMBB, University of Marburg and Justus-Liebig-University Giessen, Germany
Translational Neuroimaging Marburg, Department of Psychiatry and Psychotherapy, University of Marburg, Marburg, Germany

Tilo Kircher

Center for Mind, Brain and Behavior–CMBB, University of Marburg and Justus-Liebig-University Giessen, Germany
Translational Neuroimaging Marburg, Department of Psychiatry and Psychotherapy, University of Marburg, Marburg, Germany

Frank Bremmer

Department of Neurophysics, University of Marburg, Marburg, Germany
Center for Mind, Brain and Behavior–CMBB, University of Marburg and Justus-Liebig-University Giessen, Germany

During self-motion through an environment, our sensory systems are confronted with a constant flow of information from different modalities. To successfully navigate, self-induced sensory signals have to be dissociated from externally induced sensory signals. Previous studies have suggested that the processing of self-induced sensory information is modulated by means of predictive coding mechanisms. However, the neural correlates of processing self-induced sensory information from different modalities during self-motion are largely unknown. Here, we asked if and how the processing of visually simulated self-motion and/or associated auditory stimuli is modulated by self-controlled action. Participants were asked to actively reproduce a previously observed simulated self-displacement (path integration). Blood oxygen level–dependent (BOLD) activation during this path integration was compared with BOLD activation during a condition in which we passively replayed the exact sensory stimulus that had been produced by the participants in previous trials. We found supramodal BOLD suppression in parietal and frontal regions. Remarkably, BOLD contrast in sensory

areas was enhanced in a modality-specific manner. We conclude that the effect of action on sensory processing is strictly dependent on the respective behavioral task and its relevance.

Introduction

As we navigate in our environment, we experience a constant stream of ever-changing inputs from all of our senses. Although some sensory information might be externally generated, other sensory signals are direct consequences of our own actions and consequently more predictable. Movements can produce stimulus characteristics that are not distinguishable from those produced by other agents, therefore introducing ambiguity to pure sensory information. Additional, nonsensory information induced by our own action is required to resolve this ambiguity. Action indeed has been shown to modulate neural activity in sensory cortices. As an example, it was shown that the rate of sound-inducing actions was directly connected to the strength of the blood oxygen level–dependent (BOLD) response in the primary auditory cortex (Reznik, Ossmy, & Mukamel, 2015). The BOLD response in the somatosensory cortex was shown to be reduced (Blakemore, Wolpert, & Frith, 1998) or enhanced (Simões-Franklin, Whitaker, & Newell, 2011; Ackerley et al., 2012) for self-induced touch versus externally produced stimulation.

Influential frameworks of predictive coding in perception state that sensory information, induced by one's own action, is canceled out by prediction, leading to suppression of both neural sensory computation and perception (Jordan & Rumelhart, 1992; Wolpert, Ghahramani, & Jordan, 1995; Wolpert & Miall, 1996; Rao & Ballard, 1999; Friston, 2005). More recently, these ideas were challenged by an increasing number of observations in which enhancement instead of suppression was detected from neural responses (Doherty, 2005; Chaumon, Drouet, & Tallon-Baudry, 2008; Hughes & Waszak, 2011; Simões-Franklin et al., 2011; Ackerley et al., 2012; Reznik et al., 2015). Recent developments in optimal inference theory attempted to resolve these conflicting observations by proposing an interactive effect of prediction and task relevance (Friston, 2009; Feldman & Friston, 2010; Kok, Rahnev, Jehee, Lau, & De Lange, 2012). It was argued that task relevance and prediction interact synergistically, aiming to improve perceptual inference. In this gist, task relevance increases the precision of predictions by amplifying the gain on sensory evidence, which in turn leads to enhanced neural response. In opposing cases, in which sensory information is irrelevant, weights on sensory information are unmodified, leading to prediction suppression (Rao, 2005; Kok et al., 2012).

In complex goal-directed sensorimotor tasks, such as self-motion, sensory stimuli are typically the consequence of one's own action. As an example, the visual optical flow arriving at the retina is the direct consequence of moving forward and hence is predictable (Kaminiarz et al., 2014). Accordingly, and in line with the predictive coding framework, optic flow responses during self-induced (visually simulated) self-motion might be suppressed. Yet the speed of the optic flow over time could in principal be used to support path integration (i.e., the ability to determine the traveled path length). In such a case, an enhanced response could be expected. Our study aimed to resolve this issue.

At the same time, we have shown before that not only visual optical flow (Bremmer & Lappe, 1999) but also associated tactile and auditory stimuli can be used to solve a path integration task (visual-auditory: von Hopffgarten & Bremmer, 2010; visual-tactile: Churan, Paul, Klingenhoefer, & Bremmer, 2017). In these purely behavioral experiments, subjects employed sensory information to reproduce a previously perceived self-displacement (path integration). Path integration, in general, is the continuous updating of self-motion information during navigation. It has been investigated, for example, by asking subjects to determine the distance and angle to the starting point of their self-motion (e.g., Chrastil & Warren, 2013). The hippocampus, retrosplenial cortex, and parahippocampal cortex have been shown to be involved in such a task. Activity in the same areas increased during both translation and rotation encoding (Chrastil, Sherrill, Hasselmo, & Stern, 2016). The authors suggested that these regions track self-motion information during path integration. How sensory signals from different modalities are processed to support distance reproduction tasks, however, remained unclear. Likewise, in our current study, we aimed to determine if action-induced BOLD modulation acts in a modality-specific (examples illustrated, e.g., in Stein & Stanford, 2008) or in a supramodal manner (e.g., Straube et al., 2017; van Kemenade, Arikan, Kircher, & Straube, 2017). Accordingly, here we asked participants in the functional magnetic resonance imaging (fMRI) to replicate (active condition) perceived distances that were presented (encoding condition) either purely visually (via radial optic flow simulating self-motion across a ground plane or bimodally, with an additional sinusoidal tone with a pitch that scaled linearly with the speed of the visual stimulus) or purely auditorily. The self-produced distance replications were recorded and played back (replay condition) to our participants as control stimuli with identical stimulus characteristics but without self-induced action.

We expected the three conditions (encoding, active, and replay) to activate respective sensory cortices and
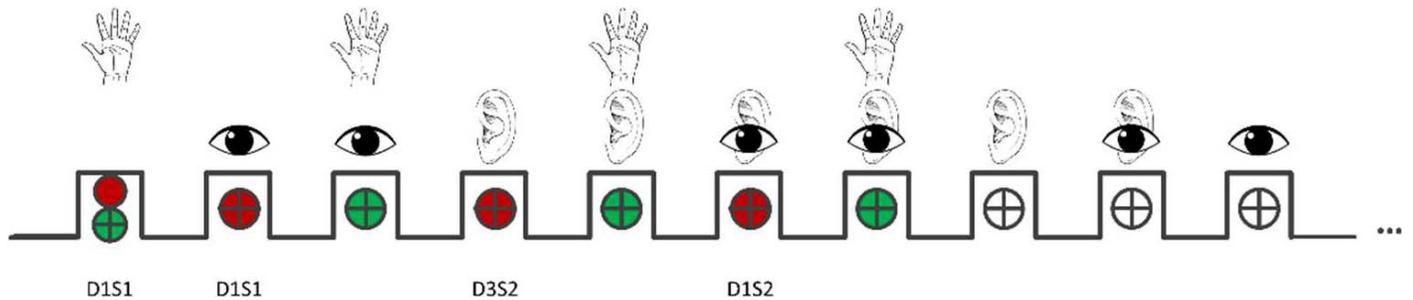
Figure 1. Exemplary stimulation protocol. The color of the fixation dot indicates the condition (encoding: red; active: green; replay: white) for the participant. The upper captions indicate the stimulus modality (eye: visual; ear: auditory; eye and ear: bimodal, hand: motor control and active condition). The lower captions indicate the travel distance (D1, D2, D3) that subjects were instructed to replicate and the travel speed (S1, S2) that was presented in the encoding condition. The replays were presented in a mini-block but in pseudorandomized order. Motor control tasks were presented in a mini-block of three at the beginning and end of each run.

activation of higher-level networks known to be involved in the processing of self-motion (Bremmer et al., 2001; Huk, Dougherty, & Heeger, 2002) and navigation (Wiener, Berthoz, & Wolbers, 2011; Chrastil et al., 2016). Furthermore, we hypothesized that BOLD responses would be modulated in the active condition, relative to the replay, in the respective sensory cortices (i.e., modality specific). Finally, based on the previous literature, we expected modality-independent BOLD suppression in higher-level cortices involved in the processing of action outcome.

Methods

Participants

Twenty-one healthy, right-handed (Edinburgh Handedness Inventory) participants (12 women, 9 men; age range = 20–33 years, mean = 26.5 years) consented to take part in the experiment according to the guidelines of the local ethics committee and the Declaration of Helsinki. All participants had normal or corrected-to-normal vision and were naïve as to the purpose of the study. Three participants were excluded from further analysis because of either (a) anatomical variance that could not be well normalized into the standard brain template (one participant) or (b) their inability to solve the path integration task in this experimental context (two subjects). From the remaining 18 participants, in one subject a single run out of five had to be excluded because of excessive head motion.

Stimulation and procedure

Combinations of uni- and bimodal stimuli were presented in pseudorandomized order, counterbalanced

over each of the five consecutive runs (Figure 1). All stimuli were presented on a computer screen (with a refresh rate of 60 Hz) positioned behind the scanner. Stimuli were presented using Octave (3.8.1) and the Psychtoolbox (Brainard, 1997). The screen was viewed by the participants in an appropriately angled mirror and covered a field of view of $21.7° \times 12.3°$.

In all conditions, a fixation target (ABC-Target from Thaler, Schütz, Goodale, & Gegenfurtner, 2013; outer circle covered 1,1° field of view, inner target covered 0.28° field of view; target luminance for the encoding condition [red] was 31 cd/m^2, in the active condition [green] 58 cd/m^2, and replay [white] 89 cd/m^2) was presented in the center of the screen, and participants were instructed to fixate the target at all times throughout each trial. We recorded eye movements in 16 of our 21 participants with an fMRI-compatible video-based eye-tracking system (EyeLink 1000, SR Research, Kanata, Ontario, Canada). For the remaining five participants, we could not monitor eye movements, mostly because of issues in detection of the corneal reflex. Analysis of the available data confirmed that participants maintained fixation during all conditions highly reliably.

In the unimodal visual modality, a horizontal two-dimensional (2D) cloud of random dots simulated self-motion across a ground plane. In the bimodal modality, the visual stimulus was accompanied by a pure sine tone, which was presented over headphones (MR-Confon Optimel, Magdeburg, Germany). The pitch level of the tone was linearly scaled with the speed of the visually simulated self-motion and ranged from 20 to 400 Hz at maximum visual travel speed, serving as an auditory speed cue. In the unimodal auditory modality, only the auditory stimulus was presented.

To enable investigation into the processing of the sensory consequences of self-induced self-motion, each trial consisted of a sequence of conditions (i.e., encoding, active reproduction, and replay). In the encoding condition, self-motion over a given distance

at one of two distinctive speeds was presented to the subject in one of the three randomly chosen stimulus modalities (encoding). This encoding phase was followed by active reproduction, in which subjects were instructed to reproduce the previously perceived travel distance in the same modality (active). They were enabled to do so using a commercially available gamepad that was customized to the MRI environment (following Trees et al., 2014). By forward deflection of the left analog stick of the gamepad, subjects were able to control their own simulated travel speed using their left thumb. Because encoding and active reproduction would never result in the exact same stimulus characteristics concerning duration and speed, stimulus parameters of the active reproduction trials (i.e., speed of the optic flow field and tone frequency, both over time) were recorded and played back later in time to the subjects as a passive control condition with no task relevance (replay). Conversely, the encoding condition was task relevant but did not require any action. This manipulation allowed us to keep the visual and/or auditory input constant (active vs. replay) and at the same time manipulate task relevance and action-related predictive mechanisms.

The color of the fixation cross indicated the type of task to the subject: red, encoding; green, active reproduction; and white, replay. Participants conducted five runs, each containing 18 trials. The presentation sequence of a given trial was fixed over each run to: (3*[encoding, active], 3*[replay])*6; Figure 1). In other words, we first asked subjects three times to reproduce a previously perceived self-displacement and then presented the three resulting replays in a row. We chose this procedure (i.e., with the replay condition presented as a mini-block toward the end of a run) to weaken the participant's sense of agency in the replay condition and mitigate possible BOLD adaptation effects. This temporal separation of active and replay trials was supposed to make the replay condition more valid as a passive control condition.

Three different distances (9, 12, and 15 arbitrary units [au]) were presented in the encoding condition. To prevent participants from solving this path integration task by simply reproducing the time of the previously perceived passive displacement, one of two different presentation speeds (37.5% or 75% of maximum travel speed in active trials) was used in each encoding condition, decoupling the correlation of travel distance and travel duration.

To control for brain activation caused by the deflection of the gamepad stick and fixating a central target, each run began and ended with a motor control task. Here, a green fixation cross was presented in the center of a screen with a red dot above. Within a random interval, ranging from 1,250 ms to 2,250 ms after fixation onset, the red dot vanished from the

screen. Participants were instructed to deflect the analog stick as long as the red dot was absent. The durations of its absence were matched to the six different durations of travel in the encoding condition. After that duration, the red dot reappeared, prompting the participants to release the analog stick. In total, we recorded the BOLD signal during 30 motor control tasks and 30 distance replications per modality.

fMRI acquisition parameters

Functional MRI data were collected in a Siemens 3 Tesla MR Magnetom Trio Tim scanner. Participants' heads were fixed with foam pads to minimize head motion artifacts. In each experimental run, a total of 350 transversal functional images were recorded. Specific parameters were as follows: echo-planar images, 64×64 matrix, 30 slices descending, field of view = 192 mm, repetition time = 1,450 ms, echo time = 25 ms, flip angle = 70°, slice thickness = 4 mm, gap size: 15%, and voxel resolution = $3 \times 3 \times 4.6$ mm, which covered the whole brain (including the cerebellum) and were positioned parallel to the intercommissural line (anterior commissure–posterior commissure).

fMRI data analysis

The obtained magnetic resonance images were analyzed using Statistical Parametric Mapping (SPM12, www.fil.ion.ucl.ac.uk) on MATLAB 8.5 (MathWorks, Sherborn, MA). Images were preprocessed with standard SPM12 routines, in the following order: realignment of functional data, co-registration between functional and structural scans, segmentation into gray and white matter and cerebrospinal fluid, normalization into the Montreal Neurological Institute (MNI) standard space template and therefore resampled to $2 \times 2 \times 2$ mm, and in a last step smoothed (full width at half maximum 8 mm). In the general linear models at the single-subject level, movement artifacts were accounted for by inclusion of realignment parameters as regressors of no interest. Low-frequency drifts were removed, employing a high-pass filter with a cutoff period of 128 s.

We modeled each condition (encoding, active, replay) separately for every stimulus modality (visual, auditory, bimodal) with a canonical hemodynamic response function. Furthermore, the motor control condition and stimulus segments that had no motion information (static episodes) were modeled in separate regressors. As a regressor of no interest, but to account for memory-related activity, the intertrial intervals between the encoding and the active conditions were modeled as well. For our random effects group

analysis, a flexible factorial design contained all T-contrasts generated on the first level. Reported whole-brain results met a threshold for family-wise error (FWE) correction at $p < 0.05$ as implemented in the standard routines of SPM12. Results showing the effects of implicit attention were thresholded at $p = 0.001$ uncorrected with a minimal cluster size of 68 voxels (determined by a 10.000 iteration Monte Carlo simulation to correct for multiple comparisons at the cluster level, $p < 0.05$). Coordinates of activation peaks are reported in MNI space. The Anatomy toolbox for SPM12 was used to co-localize functional activations with anatomical structures (Eickhoff et al., 2005).

Contrasts of interest

To determine whether active control enhanced the processing of self-motion signals in sensory cortices, we computed the contrast active > replay for each stimulus modality. We accounted for motor-related activation by computing the contrast active > motor control. The conjunction active > replay ∩ active > motor control yielded selectively action-related, but not motor-related, BOLD enhancement. BOLD suppression, on the other hand, was investigated using the inverted contrast replay > active in the respective stimulus modalities. Furthermore, we aimed to determine supramodal BOLD modulation via conjunctions across all modalities, that is, for enhancement: visual(active > replay) ∩ auditory(active > replay) ∩ bimodal(active > replay) ∩ all modalities(active > motor); and suppression: visual(replay > active) ∩ auditory(replay>active) ∩ bimodal(replay > active). We also investigated whether bimodal stimulation recruited additional supramodal cortical areas (bimodal[active] > unimodal[active]).

Results

Behavioral results

A repeated-measures analysis of variance performed on the averaged distance responses, using the factors distance (9, 12, 15 au) and modality (visual, auditory, bimodal), yielded a significant main effect of distance, $F(2, 153) = 19.1$, $p < 0.01$, but not for modality, $F(2, 153) = 1.67$, $p = 0.1919$, indicating that participants' performance in the bimodal condition was not significantly improved as compared with the visual-only or auditory-only condition. The interaction of distance × modality was not significant, $F(2, 153) = 0.03$, $p = 0.9979$.

fMRI results

A conjunction between active > replay and active > motor for each stimulus modality allowed us to determine activations that were enhanced in the active condition. Because stimulus parameters (duration, visual speed, etc.) were matched in the active and replay conditions, differences in BOLD response reflect only extrasensory signals that were consequences of one's own action and/or task relevance. The following section is descriptive. More details can be retrieved from Tables 1 and 2.

We found enhancement during unimodal visual stimulation in the left middle occipital gyrus and bilateral fusiform gyrus. The left middle occipital gyrus and bilateral fusiform gyri were also enhanced in the bimodal modality in roughly the same location with larger clusters. A conjunction of visual and bimodal enhancement confirmed the overlap in both stimulus modalities (cf. Table 1). Auditory stimulation resulted in enhancement in the bilateral superior temporal gyri and a small cluster in the right supplementary motor area. Highly spatially congruent to the unimodal auditory modality, the bilateral superior temporal gyri were also enhanced during bimodal stimulation. A conjunction analysis over all three stimulus modalities revealed no clusters when corrected for motor activity. The conjunctions over visual ∩ bimodal and auditory ∩ bimodal confirmed the overlap in the enhanced regions (cf. Table 1 and Figure 2).

Brain regions revealing an action-related suppression (Figure 3) in the visual modality were bilateral angular gyri, the midline precuneus, a cluster in the right middle temporal gyrus, the right superior temporal gyrus, two clusters in the left middle frontal gyrus, and two small clusters in the medial frontal cortex. Self-induced auditory stimulation led to suppression in the medial frontal cortex, the left angular gyrus, the right angular gyrus at a more liberal threshold of 0.001 unc. (cluster level $p_{FWE-corr} = 0.003$, 453 voxels); a small cluster in the inferior parietal lobule, the middle frontal gyrus, the superior parietal lobule, and the superior frontal gyrus; and a small cluster in the middle temporal gyrus. Self-induced bimodal stimulation exhibited suppression in the bilateral angular gyri, the medial frontal cortex, and the precuneus. Common supramodal suppression was detected in the angular gyri and the medial frontal cortex as shown by conjunction analyses over all three stimulus modalities (cf. Table 2).

Role of attention

Previous studies have investigated the relationship of prediction and attention and their mutual influence on neural activation (e.g., Kok et al., 2012). In our study, the attentional state of our subjects was not controlled

Anatomical region	Extent (anatomy toolbox)	Coordinates (peak of sign, cluster, MNI)					No. of voxels
		Hemisphere	x	y	z	t-value	
Visual							
Middle occipital gyrus	hOc4lp, hOc4la	L	−36	−88	10	5.75	232
	V3	L	−18	−92	−4	4.89	4
Fusiform gyrus	FG1, FG3, V4	L	−28	−62	−10	6.63	180
	FG1, FG3, V4	R	28	−63	−11	5.35	72
Auditory							
Superior temporal gyrus	TE 1.0, TE 1.1, TE 1.2, OP4	L	−51	−22	10	7.27	601
	TE 1.0, TE 1.1, TE 3, OP4	R	55	−22	10	6.43	417
Supplemental motor area		R	4	20	62	4.98	7
Temporal pole	TE 3	R	59	5	−4	4.91	3
Bimodal							
Superior temporal gyrus	TE 1.0, TE 1.1, TE 1.2, TE 3	L	−50	−23	9	6.51	504
	TE 1.0, TE 1.1, TE 1.2, OP4	R	64	−14	10	5.91	308
Fusiform gyrus	FG1, FG3, V3, V4,	L	−26	−65	−9	6.20	332
	FG1, FG3, V3, V4,	R	27	−63	−9	6.40	241
Middle occipital gyrus	hOc4lp, hOc4la	L	−36	−85	13	5.33	57
Conjunction (visual ∩ bimodal)							
Middle occipital gyrus	hOc4lp, hOc4la	L	−36	−84	8	5.33	55
Fusiform gyrus	FG1, FG3, V3, V4	L	−26	−62	−10	6.20	172
	FG1, FG3, V3, V4,	R	28	−64	−14	5.35	70
Conjunction (auditory ∩ bimodal)							
Superior temporal gyrus	TE 1.0, TE 1.1, TE 1.2, TE 3	L	−56	−18	10	6.51	429
	TE 1.0, TE 1.1, TE 1.2, OP4	R	64	−16	10	5.79	274

Table 1. Enhanced cluster activations from the conjunction: active > motor ∩ active > replay, $p = 0.05$ FWE.

Anatomical region	Extent (anatomy toolbox)	Coordinates (peak of sign, cluster, MNI)					No. of voxels
		Hemisphere	x	y	z	t-value	
Visual							
Angular gyrus	PGp, PGa, PFm	L	−45	−69	33	6.58	582
	PGp, PGa	R	50	−66	35	6.64	268
Precuneus	7a	LR	−1	−61	46	5.07	34
Middle temporal gyrus	TE 3	R	62	−7	−17	5.13	28
Superior frontal gyrus		R	23	38	54	5.09	20
Middle frontal gyrus		L	−25	37	50	4.97	17
		L	−34	14	62	5.20	13
Middle orbital gyrus	Fp2	L	−4	59	−9	4.90	6
	Fp2	LR	0	50	−13	4.77	5
Auditory							
Middle orbital gyrus	Fp2, Fo1	LR	0	52	−12	5.74	199
Angular gyrus	PGp, PGa, PFm	L	−43	−71	32	5.61	180
Inferior parietal lobule	PGa	L	−51	−55	27	4.74	3
Middle frontal gyrus		L	−39	13	55	5.76	70
Superior parietal lobule	7A, 7P	L	−26	−77	53	5.03	14
Superior frontal gyrus		L	−24	40	49	4.97	9
Middle temporal gyrus		L	−58	−46	−9	4.92	6
Bimodal							
Angular gyrus	PGa, PGp, PFm, 7A	L	−44	−69	37	6.35	383
		R	50	−67	38	5.55	73
Middle orbital gyrus	Fp2	L	−2	58	−8	5.10	35
Precuneus	7M	LR	0	−66	38	4.80	6
Conjunction (visual ∩ auditory ∩ bimodal)							
Angular gyrus	PGa, PGp, PFm, 7A	L	−42	−74	34	5.35	90
Middle orbital gyrus	Fp2	L	−4	60	−10	4.87	5

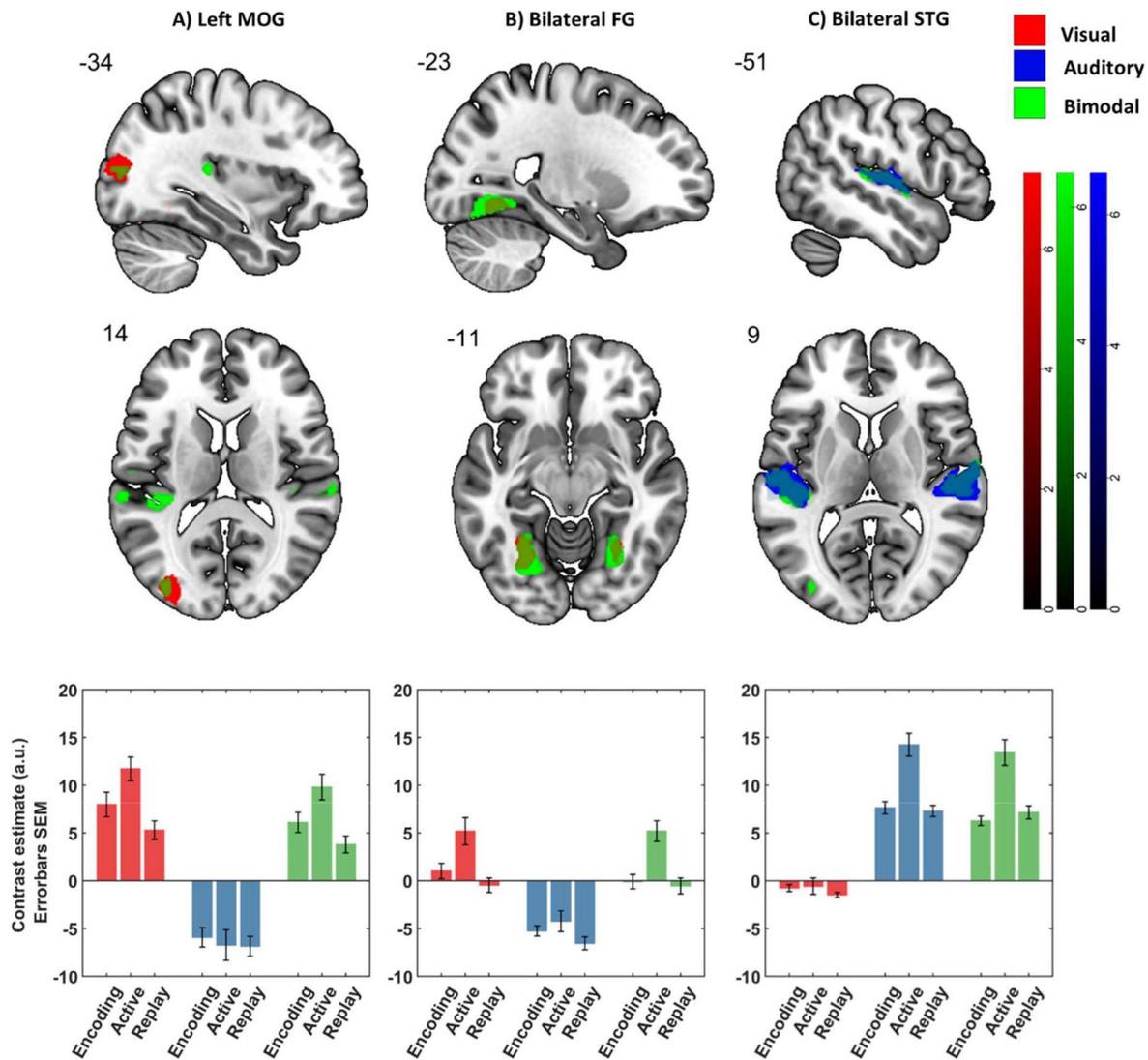Table 2. Contrast results from replay > active, $p = 0.05$ FWE (suppression).

Figure 2. BOLD-enhanced areas from the conjunction contrast active > replay ∩ active > motor control. The panels above depict BOLD enhancement for the different stimulus modalities and corresponding contrast estimates as retrieved from the respective clusters. (A) An area in the left middle occipital gyrus displayed enhancement, both in visual and bimodal stimulation. Contrast estimates reveal an increased activity for the encoding condition and a highly elevated enhancement for the active condition relative to the replay condition in the visual and bimodal modalities. No action-related modulation was observed in the auditory modality. (B) Bilateral enhancement in the fusiform gyrus (FG) for visual and bimodal stimulation. The encoding and replay conditions did not induce a BOLD enhancement in the FG cluster as the active condition. (C) The superior temporal gyrus was enhanced specifically in the auditory and bimodal condition. Top row, sagittal view; bottom row, axial view. Visual stimulation in red, auditory in blue, and bimodal in green. Color bars indicate T-scores. Threshold at 0.05 FWE.

explicitly. Nevertheless, in the encoding and active condition, subjects had to observe the visual and auditory self-motion information attentively, to solve the task of distance reproduction. On the other hand, in the replay condition, the only task of the subjects was to fixate a central target. To quantify the role of (implicit) attention, we computed a conjunction of active > replay and encoding > replay. These contrasts yield common substrates of (a) passively but attentively observed and (b) attentive active traveling in contrast to passively and nonattentively observed self-motion.

Using a correction level for FWE of $p = 0.05$, we could not find significant activation from this conjunction but could at a more liberal threshold of $p = 0.001$ unc. with a minimal cluster size of 68 voxels (determined by a 10.000 iteration Monte Carlo simulation). To further avoid false-positive reports, we list only substrates (see also Table 3) that were commonly activated in visual and bimodal or auditory and bimodal conjunctions. In the visual and bimodal conditions, we found activation bilaterally in the intraparietal sulcus (right: 40, −36, 40, k = 73; left: −52,
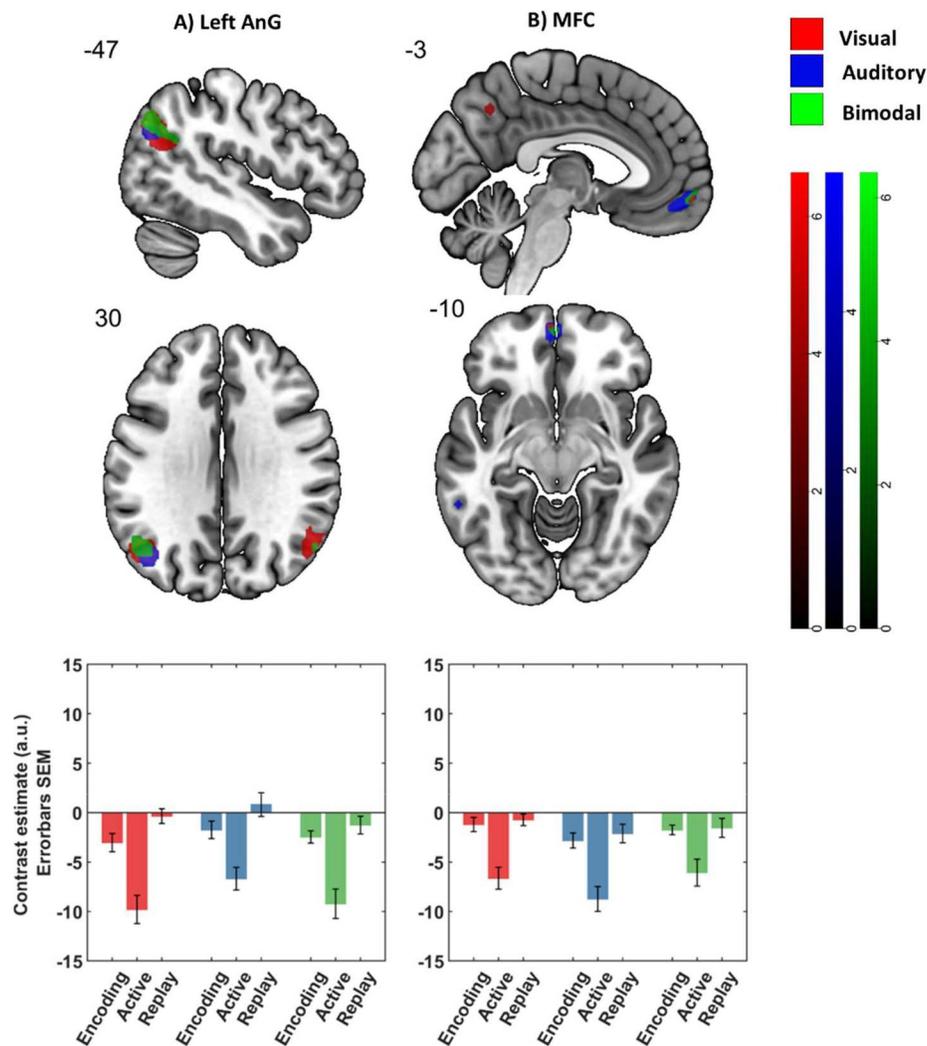
Figure 3. BOLD suppression and corresponding cluster contrast estimates. (A) Overlapping clusters for all stimulus modalities (visual, auditory, and bimodal) appeared in the left angular gyrus. The right angular gyrus displayed a similar pattern at a more liberal threshold of 0.001 unc. Cluster contrast estimates showed a strong BOLD suppression in the active condition across all stimulus modalities. The replay condition was not suppressed, whereas the encoding condition displayed medium suppression. (B) The BOLD signal in the medial frontal cortex was suppressed in all three stimulus modalities, whereas auditory stimulation exhibited the largest cluster extent. Visual stimulation led to a small cluster, whereas the extent of the bimodal cluster was intermediate (details are denoted in Table 2). Cluster estimates showed a strong suppression of the active condition in all modalities, whereas the encoding and replay conditions were less suppressed. Top row, sagittal view; bottom row, axial view. Visual stimulation in red, auditory in blue, and bimodal in green. Color bars indicate T-scores. Threshold at 0.05 FWE.

−38, 44, k = 96), in right medial frontal gyrus (34, 46, 30, k = 71), and in the left inferior occipital gyrus (−44, 80, 0, k = 70).

In the auditory and bimodal conditions, we found activation bilaterally in the medial frontal gyrus, (left: −36, 38, 28, k = 129; right: 34, 46, 30, k = 76), in the left supplementary motor area (−6, 8, 48, k = 185), in the left ventral dienencephalon (−8, −14, −12, k = 102), in the left superior frontal gyrus (−24, −2, 68, k = 87), in the right cerebellum (32, −62, −20, k = 82), and in left intraparietal sulcus (−54, −38, 46, k = 83).

The activations in the intraparietal sulcus correspond to the location of the human functional equivalent of the ventral intraparietal area (hVIP) in the macaque monkey (Bremmer et al., 2001). Inverting these contrasts (e.g., replay > encoding) did not yield any activations.

Discussion

In previous studies, we showed that visual and auditory sensory information can be used to replicate the distance of a previously observed simulated self-motion (path integration; Bremmer & Lappe, 1999; von

Anatomical region	Extent (anatomy toolbox)	Coordinates (peak of sign, cluster, MNI)					No. of voxels
		Hemisphere	x	y	z	*t*-value	
Visual and bimodal							
Superior parietal lobe/intraparietal sulcus	PFt, hIP2, PF	R	40	−36	40	4.14	73
	hIP2, 7PC	L	−52	−38	44	4.09	96
Medial frontal gyrus	Not assigned	R	34	46	30	3.88	71
Inferior occipital gyrus	hOc4la, hOc4lp	L	−44	−82	0	3.62	70
Auditory and bimodal							
Middle frontal cortex	Not assigned	R	34	46	30	3.88	76
	Not assigned	L	−36	38	28	4.65	129
Supplemental motor area	Not assigned	L	−6	8	48	4.39	185
Ventral dienencephalon	Not assigned	L	−8	−14	−12	4.36	102
Superior frontal gyrus	Not assigned	L	−24	−2	68	4.10	87
Cerebellum		R	32	−62	−20	3.87	82
Superior parietal lobe/intraparietal sulcus	PF, PFt, hIP2	L	−54	−38	46	3.75	83

Table 3. Results for implicit attentional effects. *Notes:* Conjunction over active > replay ∩ encoding > replay. Threshold set at 0.001 unc. with a minimum cluster size of 68 voxels.

Hopffgarten & Bremmer, 2011). Here, we show that both sensory-specific neural enhancement and supramodal neural suppression are related to the signal processing during such a path integration task.

We examined the behavioral responses of our participants (see Figure 4) toward a bimodal advantage in path integration performance. Although the response variance in the bimodal condition was smaller than in the auditory or visual condition, it was not as small as predicted. We conclude that visual and auditory stimuli were integrated in the bimodal condition but not in the Bayes-optimal condition (Ernst & Banks, 2002). Indeed, there have been numerous reports in the literature about nonoptimal multisensory or cue integration (as recently reviewed in Rahnev & Denison, 2019). Our task combined a visually simulated self-motion stimulus with a pure sinusoidal tone, whose frequency was scaled with the simulated speed. We had shown before (von Hopffgarten & Bremmer, 2011) that subjects can learn the

mapping of tone frequency on the simulated speed: Changing this mapping in catch trials, which went unnoticed by the participants, induced predictable behavioral errors. Nevertheless, the combination of tone frequency with simulated forward speed, as in this previous study and in our current study, was somewhat arbitrary. In such a case, optimal behavior has to be learned. This has been shown before by Ernst (2007). In this study, *participants* had to integrate luminance information with information concerning the stiffness of an object. Only after learning did participants reveal close to Bayes-optimal performance. Our results are in line with this previous finding. Accordingly, the duration of our task might have simply not been long enough to establish a stronger or even (close to) Bayes-optimal integration.

BOLD responses were enhanced in the active condition relative to a passive control condition in sensory cortices corresponding to the presented stimulus modality. Conversely, supramodal suppression
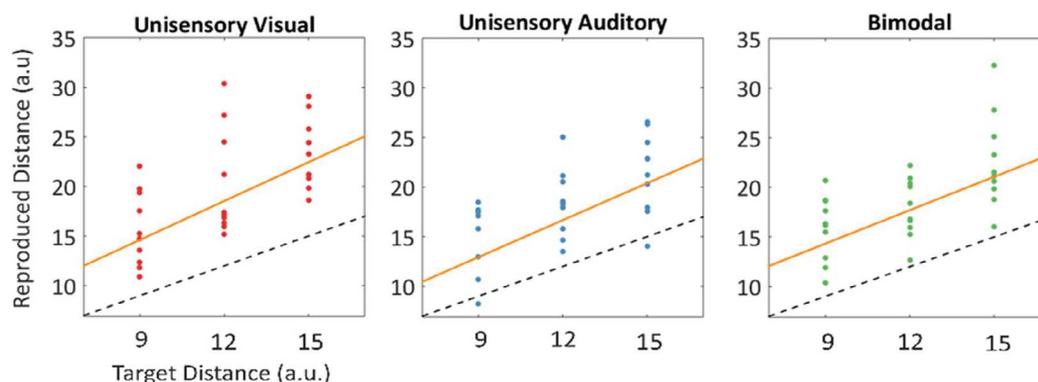

Figure 4. Exemplary behavioral responses of a single subject. The black dotted line indicates the veridical response. Linear regression (orange lines), approximated to raw responses (colored dots), indicated the general performance in the path integration task.

was observed in higher-level cortices irrespective of stimulus modality. Both BOLD suppression and enhancement were reported in earlier studies investigating the neural processing of self-induced versus externally induced sensory events. Most interestingly, and different from the outcomes in this study, previous reports reported either enhancement or suppression in their results but not both (e.g., Blakemore et al., 1998; Flanagin, Wutte, Glasauer, & Jahn, 2009; Hughes & Waszak, 2011; Ackerley et al., 2012; Shergill et al., 2013; Straube et al., 2017). Here, we observed enhancement in respective sensory areas for the corresponding visual, auditory, and bimodal stimulation and suppression for higher-level cortical regions for all stimulus modalities. This corroborates previous findings that prediction shapes information processing in the brain in a ubiquitous fashion, which ranges from the earliest sensory cortices to the highest-level frontal areas. Furthermore, our observations substantiate the notion not only that respective sensory areas are influenced by predictions within their corresponding stimulus modality but also that such predictions can pervade higher levels in a supramodal manner (cf. van Kemenade et al., 2017).

Predictive coding models have become increasingly influential in perception research because of their success in explaining a variety of behavioral phenomena (Clark, 2013). A well-described account of predictive coding in the brain is suppression due to prediction. According to this idea, internal prediction models of the outside world are represented in the brain and are tested against external information. If the prediction matches the actual incoming information, the two cancel out and neural processing is suppressed (Rao & Ballard, 1999; Friston, 2005, 2009). This claim of predictive coding models is in line with empirical findings (e.g., Blakemore et al., 1998; Martikainen, Kaneko, & Hari, 2005; Aliu, Houde, & Nagarajan, 2009; den Ouden, Friston, Daw, McIntosh, & Stephan, 2009). Conversely, signatures of neural enhancement have been reported more recently (Hughes & Waszak, 2011; Simões-Franklin et al., 2011; Ackerley et al., 2012; Huang, Chen, & Sereno, 2015; Reznik et al., 2015), triggering further research about the underlying mechanisms (Reznik et al., 2015). Progress in the development of predictive coding models offers an exciting approach to resolve this ambiguity of neural response suppression and enhancement by accounting for the relevance of sensory information to the task at hand (Friston, 2009; Feldman & Friston, 2010). To this end, Kok and colleagues (2012) manipulated spatial attention (i.e., task relevance) and prediction and probed its respective effects on the BOLD responses. In their study, the authors manipulated the effect of prediction and attention in a 2×2 factorial design. The paradigm allowed them to dissociate between two

different models on how prediction and attention are related. One model suggested that prediction and attention have opposing effects. The second model suggested that spatial attention can be thought of as highlighting a visual feature, thereby increasing the precision of information concerning this feature. The results were in favor of the second model. More specifically, their data revealed an interaction between prediction and attention in the early visual cortex: Predicted stimuli induced reduced activity compared with unpredicted stimuli when they were unattended and task irrelevant. This pattern reversed when the stimuli were attended and task relevant, as in our study. Kok and colleagues (2012) concluded that attention and prediction operate synergistically to improve the precision of perceptual inference, as also hypothesized by Friston and colleagues (Friston, 2009; Feldman & Friston, 2010). Remarkably, this result was found only in area V1 but not in areas V2 and V3. Also, our results were in line with their finding: In our study, enhancement occurred only in the early visual and auditory cortices but not at higher-order processing stages (i.e., angular gyrus and frontal cortex).

Based on these previous findings, our results suggest that in complex task sequences, such as path integration, attending to sensory information across the self-controlled (and thus predictable) sequence is key to successfully solving the task at hand. If the consequences of self-induced sensory stimulation have to be monitored to control further behavior, response enhancement indeed has been described. As an example, Reznik and colleagues (2015) asked participants to press a button at either 1 or 2 Hz. Pressing the button elicited a tone, resulting in a sequence of tones, which was played back as a passive control. A comparison of the active and passive conditions revealed enhancement in bilateral superior temporal gyri (Reznik et al., 2015). To successfully generate a rhythmic sequence of tones, participants were required to monitor their action carefully to precisely time their button press and accumulate temporal information in between tones. In other words, participants had to continuously make use of the self-induced stimulus to generate a proper prediction to provide a correct behavioral response. In a further study by Huang and colleagues (2015), participants were moved along a virtual corridor that had doors on either side that opened and became obstacles during simulated self-motion. Subjects either viewed passively a sideways dodge or actively had to dodge themselves by holding down a button until they felt that they would miss the obstacle. Although the authors did not directly contrast the active and passive conditions, they reported higher activity in the parietoinsular vestibular cortex (PIVC) for the active condition in comparison with the passive one. The authors argued that the PIVC was involved in

computing the timing when lateral dodging reached a sufficient margin to miss the obstruction. Potentially, such enhanced activation mirrors optimization of prediction through active inference by increasing the weight on sensory information.

Our findings are in line with the results from these previous studies. In our study, information from the whole self-motion sequence had to be accumulated to successfully solve the path integration task. When replicating the encoded distance, our participants had to continuously monitor their current speed and elapsed time to infer their traveled distance, until a point of subjective equality between perceived encoded and self-controlled travel distance was reached. Because the path was actively traveled by deflection of the gamepad analog stick in the active condition, self-motion speed was predictable to the participants. The parameter of speed over time was relevant to solve the required task, thus the gain on available sensory information might have been enhanced to improve predictions for the next action within the travel sequence (i.e., keep speed constant or correct speed). We therefore suggest that in our navigational task, the interaction of task relevance and prediction induced enhancement in sensory areas, namely, the superior temporal gyri, the middle occipital gyrus, and fusiform gyri.

Conversely, apart from enhancement in cortices driven by stimulus features, we also observed suppression in areas attributed with higher cognitive function in the medial frontal cortex and angular gyrus across all stimulus modalities. Research about the roles of the angular gyrus has shown involvement in visuospatial navigation (Spreng, Mar, & Kim, 2009); agency attribution (Sperduti, Delaveau, Fossati, & Nadel, 2011); integration of spatial information and conceptual categories, as, for example, the semantic concepts of left and right (Hirnstein, Bayer, Ellison, & Hausmann, 2011; reviewed by Seghier, 2013); and as a comparator area across different modalities (Sack, 2009; Straube et al., 2017; van Kemenade et al., 2017). From the current state of research, we can readily infer that this cortical region processes spatial information in a high-level conceptual and categorical frame of reference, which might be invariant to sensory modality. It was previously shown that the accuracy of the spatial localization of objects was improved through categorization that does not involve spatial metrics but uses, for example, category boundaries (Huttenlocher, Hedges, & Duncan, 1991; Huttenlocher, Hedges, Corrigan, & Crawford, 2004; Crawford & Duffy, 2010). Based on these results, it was argued that in certain behavioral contexts, aspects of spatial navigation may follow categorical (distance achieved vs. distance still not achieved, or this is more likely distance A than distance B or C) rather than metric rules (Wolbers &

Wiener, 2014; reviewed in Ekstrom & Isham, 2017). For cognitive control, the medial frontal cortex takes a central role in the executive control of goal-oriented behavior and prediction in complex tasks (Holroyd & Coles, 2002; Maurer, Maurer, & Müller, 2015; Alexander & Brown, 2018). Because of the characteristics of these cortical areas, no enhancement would be expected during our experimental task. Although we argued that sensory evidence was increased in weighting because of its task relevance, conceptually driven predictions on higher levels were not necessarily affected. Our subjects were able to distinguish well between distances, as shown from our analysis of behavioral data, and they were reasonably capable of solving the path integration task. Hence, a good match between predicted (encoded) and produced distance would have led to an overall reduction in activity, by cancellation of incoming information due to prediction, reflecting conformity of prediction, and information about traveled path length. Importantly, the stimulus modality did not appear to be relevant for the prediction of high-level concepts, as can be inferred from the fact that these cortical areas were suppressed in a modality-independent manner.

Contrasting the encoding and active condition could potentially have introduced systematic errors, because of the very different sensory stimulations (e.g., concerning the durations and speeds of self-motion). Nevertheless, the activation in the encoding condition allowed determination of the role of implicit attention on the observed activation. Kok and colleagues (2012) previously showed an interaction of attention and prediction on cortical activation. Although we had not directed our subjects' attention explicitly toward a specific feature of our visual, auditory, or bimodal stimuli, subjects had to pay attention to these stimuli in the encoding and active conditions. A conjunction analysis revealed, at a reduced statistical threshold, activation in a number of regions, among them in the functional equivalent of the macaque area VIP (Bremmer et al., 2001). We conclude that the area VIP plays a critical role in the attentive processing of unimodal and multisensory self-motion information, which is in line with results in humans (e.g., Wall & Smith, 2008) and monkeys (Britten, 2008; Chen, Deangelis, & Angelaki, 2013).

In the angular gyrus and the medial frontal cortex, we found a reduced activation in the encoding condition. In this encoding condition, subjects had to pay attention to the stimulus to extract the information concerning traveled distance. Yet they could not infer that information a priori. As shown by Kok and colleagues (2012), in such a context, prediction and attention interact, and attention to an unpredictable stimulus induces a reduced activation as compared with a predictable stimulus.

Limitations

Different from our initial hypothesis, motion-sensitive areas MT+/V5 (Huk et al., 2002), cingulate sulcus visual area (CSv; Wall & Smith, 2008; Furlan, Wann, & Smith, 2014) or hVIP (Bremmer et al., 2001) did not appear in our contrasts for either enhancement or suppression. A detailed analysis of our data revealed that all three areas showed enhanced activation in the active > replay contrast but were also strongly activated in our motor control task, in which visual, auditory, or bimodal stimulation was absent but the same action as in the navigation task was involved. Thus, the observed enhancement in the active > replay contrast was masked by our conjunction active > motor and active > replay. Although contrast estimates did show a significant difference between active and replay in all three areas, we were unable to rule out the possibility that the observed activation of area VIP was due to an activation of the nearby somatosensory cortex or of the human functional equivalent of the anterior intraparietal area (AIP) that was shown to be involved in object manipulation in humans and nonhuman primates (Grefkes & Fink, 2005). Recent investigations into the connectivity of CSv demonstrated a strong connection to the motor system, possibly explaining its activation in our motor-control task (Smith, Beer, Furlan, & Mars, 2018). Furthermore, the activation in MT+/V5 that was present in our motor control task could have arisen from the tactile stimulation of the thumb while handling the gamepad. Somatosensory responses during similar tasks were shown to be present in the human medial superior temporal area (hMST; Beauchamp, Yasar, Kishan, & Ro, 2007).

Like for areas MT+/MST and hVIP, we had expected the hippocampus, retrosplenial cortex, and entorhinal cortex to be activated in our task. Indeed, we found that activity in all three regions was smaller in the visual (hippocampus and retrosplenial cortex) or bimodal (entorhinal) condition in the active as compared with the replay condition. Yet this difference did not pass the FWE correction, which was caused by the relatively small number of trials involved for a single condition and by the fact that we only tested a one-dimensional self-motion task, whereas, for example, the entorhinal cortex represents 2D movements with a gridlike structure of activation (Doeller et al., 2010).

Because of the necessary ordering of the experimental conditions (i.e., the active condition had to follow the encoding condition, and the replay conditions could be presented only after the active conditions), we cannot fully exclude ordering effects in our data.

Although we argue that task relevance and prediction shape BOLD responses interactively in our task conditions, from our experimental design, manipulations of task relevance were not conducted in a fully systematic way. Rather, we relied on the idea that participants did not employ sensory information from the replay condition over the course of the experiment because of the lack of informational value regarding the task. Finally, participants might have been able to detect the replay as their own former actions due to the unique character of each speed profile. Although we consider this highly unlikely, we report this potential influence for the sake of completeness.

Conclusion

In our study, we investigated how multisensory self-motion information is processed in a path integration task. We highlight the fact that in this complex task, processing at a modality-specific sensory level was enhanced, whereas the BOLD response was suppressed in high-level cortical areas independently of sensory modality. We interpret the observed enhancement as an interactive effect of task relevance and prediction, in which modality-specific sensory signals are enhanced to improve perceptual inference and facilitate prediction for upcoming action. Conversely, suppression in high-level cortical regions expresses conformity of prediction and incoming information irrespective of the sensory modality involved.

Keywords: self-motion, prediction, action, perceptual inference, task relevance

Acknowledgments

Commercial relationships: none.
Corresponding author: Milosz Krala.
Email: milosz.krala@physik.uni-marburg.de.
Address: Department of Neurophysics, University of Marburg, Marburg, Germany.

References

Ackerley, R., Hassan, E., Curran, A., Wessberg, J., Olausson, H., & McGlone, F. (2012). An fMRI study on cortical responses during active self-touch and passive touch from others. *Frontiers in Behavioral Neuroscience*, 6, 1–9.

Alexander, W. H., & Brown, J. W. (2018). Frontal cortex function as derived from hierarchical predictive coding. *Scientific Reports*, 8, 1–11.

Aliu, S. O., Houde, J. F., & Nagarajan, S. S. (2009). Motor-induced suppression of the auditory cortex. *Journal of Cognitive Neuroscience*, 21, 791–802.

Beauchamp, M. S., Yasar, N. E., Kishan, N., & Ro, T. (2007). Human MST but not MT responds to tactile stimulation. *Journal of Neuroscience*, 27, 8261–8267.

Blakemore, S. J., Wolpert, D. M., & Frith, C. D. (1998). Central cancellation of self-produced tickle sensation. *Nature Neuroscience*, 1, 635–640.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10, 433–436.

Bremmer, F., & Lappe, M. (1999). The use of optical velocities for distance discrimination and reproduction during visually simulated self motion. *Experimental Brain Research*, 127, 33–42.

Bremmer, F., Schlack, A., Shah, N. J., Zafiris, O., Kubischik, M., Hoffmann, K.-P., . . . Fink, G. R. (2001). Polymodal motion processing in posterior parietal and premotor cortex: A human fMRI study strongly implies equivalencies between humans and monkeys. *Neuron*, 29, 287–296.

Britten, K. H. (2008). Mechanisms of self-motion perception. *Annual Reviews*, 31, 389–410.

Chaumon, M., Drouet, V., & Tallon-Baudry, C. (2008). Unconscious associative memory affects visual processing before 100 ms. *Journal of Vision*, 8(3):10, 1–10, https://doi.org/10.1167/8.3.10. [PubMed] [Article]

Chen, A., Deangelis, G. C., & Angelaki, D. E. (2013). Functional specializations of the ventral intraparietal area for multisensory heading discrimination. *Journal of Neuroscience*, 33, 3567–3581.

Chrastil, E. R., Sherrill, K. R., Hasselmo, M. E., & Stern, C. E. (2016). Which way and how far? Tracking of translation and rotation information for human path integration. *Human Brain Mapping*, 37, 3636–3655.

Chrastil, E. R., & Warren, W. H. (2013). Does the human odometer use an extrinsic or intrisic metric? *Attention, Perception, & Psychophysics*, 76, 230–246.

Churan, J., Paul, J., Klingenhoefer, S., & Bremmer, F. (2017). Integration of visual and tactile information in reproduction of traveled distance. *Journal of Neurophysiology*, 118, 1650–1663.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral Brain Sciences*, 36, 181–204.

Crawford, L. E., & Duffy, S. (2010). Sequence effects in estimating spatial location. *Psychonomic Bulletin & Review*, 17, 725–730.

den Ouden, H. E. M., Friston, K. J., Daw, N. D., McIntosh, A. R., & Stephan, K. E. (2009). A dual role for prediction error in associative learning. *Cerebral Cortex*, 19, 1175–1185.

Doeller, C. F., Barry, C., & Burgess, N. (2010, February 4). Evidence for grid cells in a human memory network. *Nature*, 463, 657–661.

Doherty, J. R. (2005). Synergistic effect of combined temporal and spatial expectations on visual attention. *Journal of Neuroscience*, 25, 8259–8266.

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, 1325–1335.

Ekstrom, A. D., & Isham, E. A. (2017). Human spatial navigation: Representations across dimensions and scales. *Current Opinion in Behavioral Sciences*, 17, 84–89.

Ernst, M. O. (2007). Learning to integrate arbitrary signals from vision and touch. *Journal of Vision*, 7(5):7, 1–14, https://doi.org/10.1167/7.5.7. [PubMed] [Article]

Ernst, M. O., & Banks, M. S. (2002, January 24). Humans integrate visual and haptic information in a statistically optimal fashions. *Nature*, 415(6870), 429–433.

Feldman, H., & Friston, K. J. (2010). Attention, uncertainty, and free-energy. *Frontiers in Human Neuroscience*, 4, 1–23.

Flanagin, V. L., Wutte, M., Glasauer, S., & Jahn, K. (2009). Driving dreams: Cortical activations during imagined passive and active whole body movement. *Annals of the New York Academy of Sciences*, 1164, 372–375.

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360, 815–836.

Friston, K. (2009). The free-energy principle: A rough guide to the brain? *Trends in Cognitive Sciences*, 13, 293–301.

Furlan, M., Wann, J. P., & Smith, A. T. (2014). A

representation of changing heading direction in human cortical areas pVIP and CSv. *Cerebral Cortex*, *24*, 2848–2858.

Grefkes, C., & Fink, G. R. (2005). The functional organization of the intraparietal sulcus in humans and monkeys. *Journal of Anatomy*, *207*, 3–17.

Hirnstein, M., Bayer, U., Ellison, A., & Hausmann, M. (2011). TMS over the left angular gyrus impairs the ability to discriminate left from right. *Neuropsychologia*, *49*, 29–33.

Holroyd, C. B., & Coles, M. G. H. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, *109*, 679–709.

Huang, R. S., Chen, C. F., & Sereno, M. I. (2015). Neural substrates underlying the passive observation and active control of translational egomotion. *Journal of Neuroscience*, *35*, 4258–4267.

Hughes, G., & Waszak, F. (2011). ERP correlates of action effect prediction and visual sensory attenuation in voluntary action. *Neuroimage*, *56*, 1632–1640.

Huk, A. C., Dougherty, R. F., & Heeger, D. J. (2002). Retinotopy and functional subdivision of human areas MT and MST. *Journal of Neuroscience*, *22*, 7195–7205.

Huttenlocher, J., Hedges, L. V., Corrigan, B., & Crawford, L. E. (2004). Spatial categories and the estimation of location. *Cognition*, *93*, 75–97.

Huttenlocher, J., Hedges, L. V., & Duncan, S. (1991). Categories and particulars: Pprototype effects in estimating spatial location. *Psychological Review*, *98*, 352–376.

Jordan, M. I., & Rumelhart, D. E. (1992). Forward models: Supervised learning with a distal teacher. *Cognitive Science*, *16*, 307–354.

Kaminiarz, A., Schlack, A., Hoffmann, K.-P., Lappe, M., & Bremmer, F. (2014). Visual selectivity for heading in the macaque ventral intraparietal area. *Journal of Neurophysiology*, *112*, 2470–2480.

Kok, P., Rahnev, D., Jehee, J. F. M., Lau, H. C., & De Lange, F. P. (2012). Attention reverses the effect of prediction in silencing sensory signals. *Cerebral Cortex*, *22*, 2197–2206.

Martikainen, M. H., Kaneko, K., & Hari, R. (2005). Suppressed responses to self-triggered sounds in the human auditory cortex. *Cerebral Cortex*, *15*, 299–302.

Maurer, L. K., Maurer, H., & Müller, H. (2015). Neural correlates of error prediction in a complex motor task. *Frontiers in Behavioral Neuroscience*, *9*, 209.

Rahnev, D., & Denison, R. (2018). Suboptimality in perceptual decision making. *Behavioral and Brain Sciences*, *41*, E223.

Rao, R. P. N. (2005). Bayesian inference and attentional modulation in the visual cortex. *Neuroreport*, *16*, 1843–1848.

Rao, R. P. N., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, *2*, 79–87.

Reznik, D., Ossmy, O., Mukamel, R. (2015). Enhanced auditory evoked activity to self-generated sounds is mediated by primary and supplementary motor cortices. *Journal of Neuroscience*, *35*, 2173–2180.

Sack, A. T. (2009). Parietal cortex and spatial cognition. *Behavioral Brain Research*, *202*, 153–161.

Seghier, M. L. (2013). The angular gyrus: Multiple functions and multiple subdivisions. *Neuroscientist*, *19*, 43–61.

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.

Simões-Franklin, C., Whitaker, T. A., & Newell, F. N. (2011). Active and passive touch differentially activate somatosensory cortex in texture perception. *Human Brain Mapping*, *32*, 1067–1080.

Smith, A. T., Beer, A. L., Furlan, M., & Mars, R. B. (2018). Connectivity of the cingulate sulcus visual area (CSv) in the human cerebral cortex. *Cerebral Cortex*, *28*, 713–725.

Sperduti, M., Delaveau, P., Fossati, P., & Nadel, J. (2011). Different brain structures related to self- and external-agency attribution: A brief review and meta-analysis. *Brain Structure and Function*, *216*, 151–157.

Spreng, R. N., Mar, R. A., & Kim, S. N. A. (2009). The common neural basis of autobiographical memory, prospection, naviation, theory of mind an the default mode: A quantitative meta-analysis. *Journal of Cognitive Neuroscience*, *21*, 489–510.

Stein, B. E., & Stanford, T. R. (2008). Multisensory integration: current issues from the perspective of the single neuron. *Nature Reviews Neuroscience*, *9*, 255–266.

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*, e0169131.

Thaler, L., Schütz, A. C., Goodale, M. A., & Gegenfurtner, K. R. (2013). What is the best

fixation target? The effect of target shape on stability of fixational eye movements. *Vision Research*, *76*, 31–42.

Trees, J., Snider, J., Falahpour, M., Guo, N., Lu, K., Johnson, D. C., ... Liu, T. T. (2014). Game controller modification for fMRI hyperscanning experiments in a cooperative virtual reality environment. *MethodsX*, *1*, 292–299.

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*, 3691–3703.

von Hopffgarten, A., & Bremmer, F. (2011). Self-motion reproduction can be affected by associated auditory cues. *Seeing and Perceiving*, *24*, 203–222.

Wall, M. B., & Smith, A. T. (2008). The representation of egomotion in the human brain. *Current Biology*, *18*, 191–194.

Wiener, J. M., Berthoz, A., & Wolbers, T. (2011). Dissociable cognitive mechanisms underlying human path integration. *Experimental Brain Research*, *208*, 61–71.

Wolbers, T., & Wiener, J. M. (2014). Challenges for identifying the neural mechanisms that support spatial navigation: The impact of spatial scale. *Frontiers in Human Neuroscience*, *8*, 571.

Wolpert, D., Ghahramani, Z., & Jordan, M. (1995, September 29). An internal model for sensorimotor integration. *Science*, *269*(5232), 1880–1882.

Wolpert, D. M., & Miall, R. C. (1996). Forward models for physiological motor control. *Neural Networks*, *9*, 1265–1279.