

Research report

Predicting goals in action episodes attenuates BOLD response in inferior frontal and occipitotemporal cortex



Moritz F. Wurm^{a,c,*}, Mari Hrkáč^{b,c}, Yuka Morikawa^d, Ricarda I. Schubotz^{b,c}

^a University of Trento, Center for Mind/Brain Sciences, Via delle Regole, 101, 38100 Mattarello, TN, Italy

^b University of Münster, Institute of Psychology, 48149 Münster, Germany

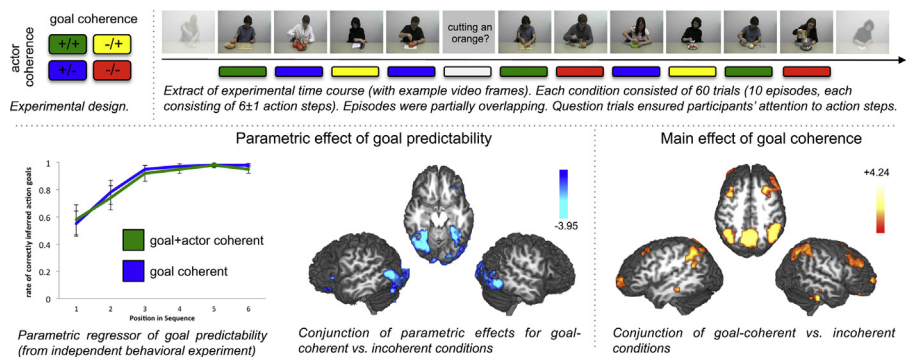
^c Max Planck Institute for Neurological Research, 50931 Cologne, Germany

^d University Hospital Münster, Department of Clinical Radiology, 48149 Münster, Germany

HIGHLIGHTS

- Actions consist of several action steps directed towards a common overarching goal.
- We investigate the change of BOLD during observation of unfolding action episodes.
- Neural activity in IFG and OTC decreases as a function of goal predictability.
- IFG integrates action steps to episodes enabling anticipation of forthcoming actions.
- Detection of goal coherence draws on episodic memory network.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:

Received 25 July 2014

Accepted 30 July 2014

Available online 6 August 2014

Keywords:

fMRI

Action observation

Coherence

Goal

Prediction

Inferior frontal gyrus

ABSTRACT

Actions are usually made of several action steps gearing towards an overarching goal. During observation of such action episodes the overarching action goal becomes more and more clear and upcoming action steps can be predicted with increasing precision. To tap this process, the present fMRI study investigated the dynamic changes of neural activity during the observation of distinct action steps that cohere by an overarching goal. Our hypotheses specifically addressed the role of the inferior frontal gyrus (IFG), a region assumed to be a key hub for integration functions during action processing, as well as the role of regions involved in action perception (often referred to as action observation network or AON) that should benefit from the predictability of forthcoming action steps.

Participants watched separate action steps that formed a coherent action goal or not (factor goal coherence) and were performed by a single actor or not (factor actor coherence). Independent of actor coherence, neural activity in IFG and occipitotemporal cortex decreased as a function of goal predictability during the unfolding of goal-coherent episodes. In addition, we identified a network (precuneus, dorsolateral prefrontal and orbitofrontal cortex, angular gyrus, and middle temporal gyrus) that showed increased activity for goal coherence.

We conclude that IFG fosters the integration of action steps to build overarching goals. Identifying the unifying goal of an action episode allows anticipation, and thus efficient processing, of forthcoming action steps. To this end, past action steps of the action episode are buffered and recollected with recourse to episodic memory.

© 2014 Elsevier B.V. All rights reserved.

* Corresponding author at: Center for Mind/Brain Sciences, University of Trento, Via delle Regole, 101, 38100 Mattarello, TN, Italy. Tel.: +39 0461 28 2783.
E-mail addresses: moritz.wurm@unitn.it, moritz.f.wurm@gmail.com (M.F. Wurm).

1. Introduction

Goal-directed behaviour is usually organized in a sequential and hierarchically structured manner: Single action steps, e.g. buttering bread, slicing cheese, etc., make up a meaningful action to achieve an overarching goal, e.g., preparing a sandwich [1–4]. A recent study suggests that both the presence of a common overarching goal and the continuity of the actor can serve as cues for detection of meaningful associations between action steps [5]. Exploiting this coherence can be considered beneficial for the anticipation of upcoming associated action steps. So far, it has not been investigated whether brain activity reflects coherence in observed action, and if so, in which way.

We employed fMRI to investigate if coherence impinges on action perception in a spontaneous fashion. That is, coherence was neither part of an instructed detection task nor was it particularly promoted by the stimulus. Nonetheless, coherence should have a spontaneous impact on action observation as a predictor of upcoming actions. In a companion study [6], we focused on the effect of coherence induced by continuity of the actor; in the present study, we investigated coherence between single action steps (e.g., buttering bread, slicing cheese) established by overarching goals (e.g. preparing a sandwich). We tested two hypotheses:

First, during the unfolding of a goal-coherent action sequence (episode hereafter) BOLD activity should decrease in several brain areas as a function of predictability. Attenuation should be observed in the inferior frontal gyrus (IFG) as this region was found to be negatively correlated with the ease of semantic integration in general [7–9] and specifically of observed action [6,10,11]. In addition, IFG is activated during the selection of overarching action plans that consist of separate motor responses [12], supporting the view that IFG is relevant for integration of action steps to form higher order action goals. As a consequence of semantic integration in IFG, activation should also decrease in the premotor-parietal network that is reported for execution, planning, imagery and observation of action [13]. The recognition of actions, i.e., the matching of observed action and action memory, should be promoted by predictability of goal-coherent actions, thereby reducing the cognitive load in premotor-parietal regions [14,15]. Finally, we expected a decrease of activity in occipitotemporal cortex (OTC) due to top-down suppression of anticipated visual input and overall facilitation of action perception [16–18].

Second, integrating a currently observed action with recently observed actions as being performed in pursuit of an overarching goal draws on episodic memory. Hence, episodic memory networks should be activated more when goal coherence becomes operative in action observation than when it does not. The retrieval and maintenance of episodic memory is associated with activation of dorsolateral prefrontal cortex (dlPFC) and lateral parietal cortex [19–21]. In addition, we expected to find activation linked to the presence of coherence. Coherence has not been investigated in the context of action observation so far. However, a meta-analysis on text comprehension identified two dominant networks, consisting of precuneus and medial prefrontal cortex (mPFC), and anterior temporal lobe and mid middle temporal gyrus (MTG), that are activated for coherent vs. incoherent sentences or words [22].

We tested these hypotheses by presenting participants with actions that either formed a common overarching goal or not (factor GOAL COHERENCE). Effects of goal coherence should not only manifest in goal-coherent episodes that are performed by a single actor but also in goal-coherent episodes that are carried out by different actresses and actors. Therefore, episodes were either performed by the same actor or by different actresses and actors (factor ACTOR COHERENCE). We thereby obtained a 2×2 design with four conditions: goal-coherent and actor-coherent=both coherent, BC; goal-coherent and actor-incoherent, GC; goal-incoherent and

Table 1

2×2 Factorial experimental design. In the goal-coherent conditions (BC, GC) 10 goals, each consisting of 6 ± 1 actions, were used. Correspondingly, in the goal-incoherent conditions $10 \times 6 \pm 1$ unrelated actions were used. Each condition employed 10 actors.

		GOAL COHERENCE	
		+	–
ACTOR COHERENCE	+	BC 1 Goal (6 ± 1 actions) per episode performed by 1 actor 10 Goals/actors → 60 trials	AC 6 ± 1 Independent actions per episode performed by 1 actor 10 Actors → 60 trials
	–	GC 1 Goal (6 ± 1 actions) per episode performed by 6 ± 1 actors 10 Goals/actors → 60 trials	NC 6 ± 1 Independent actions per episode performed by 6 ± 1 actors 10 Actors → 60 trials

actor-coherent, AC; goal-incoherent and actor-incoherent=non coherent, NC (Fig. 1A; Table 1). There were 6 ± 1 actions that belonged to an episode. The actions of one episode were not presented in direct temporal succession but interleaved with actions of the other conditions (Fig. 1B). Participants' attention was drawn to each single action by occasionally asking them whether a presented action description, e.g. "buttering bread", matched the preceding action or not. The task thus did not orient the participants' attention to the overarching goal but toward the actual action step.

We reasoned that activity in IFG, as well as in the premotor-parietal network and in OTC, attenuates as the episode unfolds and the action goal becomes more and more evident. In a simple model, IFG attenuation could be modulated in a linear way with every action step that adds to the episode. However, it is more likely that the predictability of the overarching action goal increases in a nonlinear way because action steps at the beginning of the episode contribute more to goal predictability than action steps at the end of the episode. Such a model resembles more a logarithmic function. To test putative nonlinear effects of goal predictability, we therefore modeled the BOLD response with data from a behavioral experiment, in which participants were asked to predict the overarching goal of each of the actions of the episodes (parametric factor GOAL PREDICTABILITY). Finally, main effects of goal coherence were investigated by contrasting goal-coherent vs. goal-incoherent episodes.

2. Methods

2.1. Participants

Twenty-two right-handed, healthy volunteers (11 females, mean = 25.89 years, range 22–30 years) participated in the study. All participants were right-handed according to the Edinburgh Inventory Manual Preference [23], had normal or corrected-to-normal vision, and were native German speakers. No subject had a history of neurological or psychiatric disorder. Four participants had to be excluded from the analysis due to technical problems. All participants were screened by a physician and gave informed written consent to participate in this study. The study was approved by the local ethics committee of the Medical Faculty, University of Cologne, Germany.

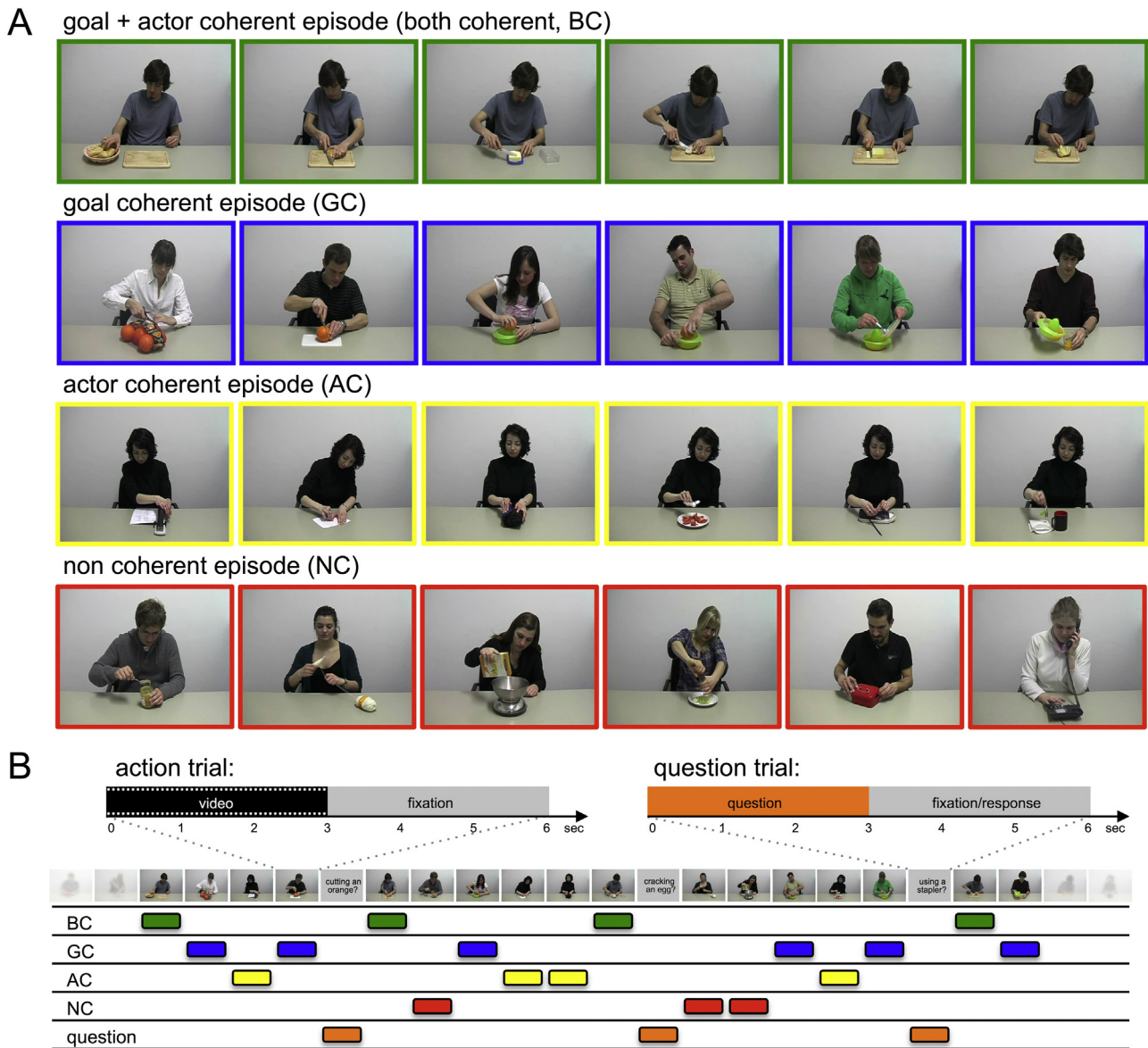


Fig. 1. Experimental design. (A) Example video frames of episodes for each condition. 6 ± 1 separate action steps added up to one episode. Action steps pertained to an overarching goal or not (goal coherence) and were performed by either one or by varying actors (actor coherence). BC: both goal and actor coherent, GC: goal coherent, AC: actor coherent, NC: non coherent. (B) Video and question trials were presented in an event-related fashion. Question trials required participants to confirm or reject a verbal action description with regard to the preceding trial. Episodes were partially overlapping. Hence, videos of one episode were mostly not shown in direct succession.

2.2. Stimuli

Stimuli consisted of 3 s long video clips of 240 actions performed by 40 different actors. These actions were either steps of a common action with an overarching goal or not (factor GOAL COHERENCE). In addition, actions were performed by either the same or by varying actors (factor ACTOR COHERENCE). This 2×2 design resulted in four experimental conditions (BC, GC, AC, NC; see Table 1). In the goal-coherent conditions (BC, GC), there were 10 episodes per condition with overarching goals. In a similar vein, in the goal-incoherent conditions actions were grouped to episodes that had no overarching goal.

To provide an example, an episode in condition BC (both goal and actor coherent) was organized as follows (compare Fig. 1A): an actor takes a bun, cuts it, butters it, cuts cheese, and puts the cheese on the bun. In condition GC (goal coherent), the first actress takes an orange, the second actor halves it, the third actress squeezes the first half, the fourth actor squeezes the second half, the fifth actress

removes the pulp from the squeezer, and the sixth actor pours the juice into a glass. In condition AC (actor coherent), an actress staples papers in the first action, writes a letter in the second, opens a wallet in the third, salts a tomato in the fourth, laces a shoe in the fifth, and makes a tea in the sixth. In condition NC (non coherent), varying actresses and actors take honey from a glass, knit, pour flour in a beaker, pepper a meal, and open a cash box.

In each condition, episodes consisted of six actions on average (two episodes of five actions, six episodes of six actions, and two episodes of seven actions). There were 60 actions per condition in total.

The occurrence of actors was balanced across conditions: Every actor was assigned to one condition and did not appear in any other. There were hence 10 different actors per condition (40 in total). Every actor was assigned to one condition and did not appear in any other. Overall, each actor occurred 6 ± 1 times (depending on the length of the episode, see above) during the experiment. In the conditions GC and NC, where episodes were performed by different

actors, the actors occurred evenly spread over the course of the experiment with a gap length of 23 trials on average, while in the conditions BC and AC the gap was four trials at most (on average 1.82 ± 1.2). Objects involved in the actions appeared only in actions of the same episodes. For example, the squeezer in the example in Fig. 1A appeared only in this episode and not in any other episode of the same or other conditions.

The episodes of each condition were partially overlapping (see Fig. 1B). The overlap was balanced across conditions to avoid confounding due to different working memory loads. In addition, the gaps between actions of an episode had a maximum length of four actions of other episodes (1.82 ± 0.09 trials on average). The gap length was balanced across conditions to avoid confounding effects by virtue of different episode inference difficulties. In addition, gap length was included in the design matrix as nuisance regressor (see Section 2.5). The mean number of actions shown for the completion of an episode was 13 (12.9 ± 0.36), including the overlapping actions of other episodes.

Taken together, this experimental design and set up guaranteed a full balance of the conditions with respect to number of actions, number of different actors, number of episodes, overlap of episodes, and gap length between actions of the same episode.

Actions were embedded into 6 s long trials (action trials) that were presented in an event-related fashion intermixed with 48 written action descriptions referring to these actions (question trials). Each trial started with a video or a question (3 s) followed by a fixation phase (3 s). Every second question trial was followed by a null event (6 s) requiring participants to fixate a fixation dot. Question trials and null events followed and preceded each of the four experimental conditions equally often (12 ± 2 times).

2.3. Goal predictability estimation

To estimate how well the overarching goals of the BC and GC episodes could be predicted from a given action step, we performed a separate behavioral experiment. Ten participants (not from the fMRI sample) watched the trials of the 10 BC and 10 GC episodes in a blocked way, i.e., all 6 ± 1 trials of one episode in direct succession. The order of the 20 episode blocks was randomized for each participant. After each trial the participants had to name the overarching goal they suggested for the presented action (e.g. “preparing orange juice” for the trial “halving an orange”). We then computed the mean correct response rate for each trial position in the sequences (see Fig. 2B). Each value hence represents the likelihood of a correct prediction of the action goal after observation of a given action step. To estimate how well the goal of a given trial could be predicted from the preceding trials of the episode, we assigned the values of trial n to trial $n + 1$. From the first trial, the action goal was unpredictable and therefore had the value 0. These values were used as parametric factor GOAL PREDICTABILITY.

2.4. Task

Participants were instructed to watch the presented videos attentively. They were informed that some action videos were followed by an action description that either corresponded to the content of the preceding video or not. Half of the action descriptions matched and half did not. Participants responded to the description on a two-button response box, using the right index finger to accept and the right middle finger to reject. The action descriptions used in the question trials referred to the subgoals of actions (e.g., “halving an orange”), which were directly derivable from the shown videos. Phrasings like “preparing orange juice” were avoided in order not to point the participants to the episodic structure of the actions.

2.5. Nuisance regressors

We included four nuisance regressors to control for potential effects that might go along with the unfolding of episodes but that are not directly related to goal coherence.

First, although we used an implicit task that did not require the participants to follow the episode, we could not rule out that participants actively kept information of the unfolding episode in working memory. We therefore controlled for potential working memory effects. Each action step adds information to the episode that might be kept in working memory. Thus, working memory load should increase parametrically with each action step. We therefore included POSITION IN SEQUENCE (first action step = 1, second action step = 2, etc.) as nuisance regressor in the design matrix.

Second, intermittent/interspersed action steps might increase the load to keep the previous action step of the current episode in working memory. This load should vary as a function of the number of unrelated action steps between the current action and the previous related action step of the same episode. We therefore also included GAP LENGTH as nuisance regressor in the design matrix. For the non-coherent episodes, both POSITION IN SEQUENCE and GAP LENGTH were set to “1”.

Third, we sought to rule out putative confounding influences due to differences in peculiarity and attractiveness of the actors between the four conditions. In a post fMRI session survey, we therefore asked the participants to rate the attractiveness and peculiarity of the actors on a four point Likert scale (from 1 = “not attractive” and “not peculiar” to 4 = “very attractive” and “very peculiar”). An ANOVA for the attractiveness rating revealed no significant differences between the conditions. The ANOVA for the peculiarity rating revealed that actors in the NC condition were significantly more peculiar (mean Likert rating 2.48) than in GC (2.05; $p < 0.001$) and AC (2.06; $p < 0.001$). We therefore included PECULIARITY in the fMRI analysis as nuisance regressor.

Fourth, action steps that serve the achievement of an overarching goal often, but not always, involve a common set of objects. In the present experiment, too, goal-coherent actions involved a varying number of objects that were manipulated in different parts of the respective action sequence (see Fig. 1A). To rule out potential effects due to recognition of objects that appeared already in previous action steps of the same episode, we constructed a regressor that modeled the OBJECT RECENCY: For each trial we counted how often each object in that trial was shown already in previous trials of the same episode. In the “preparing sandwich” example in Fig. 1A, the first trial has “0”, the second trial has “2” (bread = 1, cutting board = 1), the third trial has “0”, the fourth trial has “5” (bread = 2, cutting board = 2, butter knife = 1), the fifth trial has “3” (cutting board = 3; note that the knife is different from the other knives used in the episode), the sixth trial has “8” (bread = 3, cutting board = 4, cheese = 1). This parameter was included as nuisance regressor in the design matrix. For goal-incoherent episodes the parameter was set to “0”, since they never involved objects that appeared in other actions of the experiment.

2.6. MRI data acquisition

Imaging was performed on a 3-T Scanner (Siemens Magnetom TRIO, Erlangen, Germany) equipped with a standard birdcage head coil. Participants lay supine on the scanner bed with their right index and middle fingers positioned on the appropriate response buttons of a response box. To prevent head, arm, and hand movements, form-fitting cushions were used. Participants were provided earplugs to attenuate scanner noise. Twenty-eight axial slices (4 mm thickness; 1 mm spacing; 200 mm field of view; 64×64 pixel matrix; in-plane resolution of 3×3 mm) covering the whole brain were acquired using a single shot gradient EPI

sequence (2 s repetition time; 30 ms echo time; 90° flip angle; 116 kHz acquisition bandwidth) sensitive to BOLD contrast. Prior to functional imaging, 28 anatomical T1-weighted MDEFT images [24,25] were acquired. In different session, high-resolution whole brain images were acquired from each subject to improve the localization of activation foci using a T1-weighted 3D-segmented MDEFT sequence.

2.7. MRI data analysis

After motion correction using rigid-body registration to the central volume, the fMRI data were processed using the software package LIPSIA [26]. A cubic-spline interpolation was used to correct for the temporal offset between the slices acquired in one image. To remove low-frequency signal changes and baseline drifts a temporal high-pass filter with a cutoff frequency of 1/85 Hz was used. Spatial smoothing with a Gaussian filter of 5.65 mm FWHM was applied. A rigid linear registration with six degrees of freedom (three rotational, three translational) was performed to align the functional data slices with a 3D stereotactic coordinate reference system. The rotational and translational parameters were obtained on the basis of the MDEFT and the EPI-T1 slices to achieve an optimal match between these slices and the individual 3D reference dataset. The MDEFT volume dataset with 128 slices and 1-mm slice thickness was standardized to the Talairach stereotactic space [27]. The rotational and translational parameters were subsequently normalized by linear scaling to a standard size. The resulting parameters were then used to transform the functional slices using trilinear interpolation, so the resulting functional slices were aligned with the stereotactic coordinate system, thus generating isotropic voxels with a spatial resolution of $3 \times 3 \times 3$ mm³.

The statistical evaluation was based on a least-squares estimation using the general linear model for serially autocorrelated observations [28,29]. The design matrix was generated with a delta function, convolved with the hemodynamic response function. Brain activations were analyzed time-locked to onset of the videos, and the analyzed epoch comprised the full duration (3 s) of the presented videos, the duration of the null events (6 s), and the reaction time in question trials (max. 3 s). The design matrix contained 12 regressors: four predictors of the four experimental conditions, two regressors for the goal-coherent conditions with an amplitude modeled by the parameter GOAL PREDICTABILITY, one regressor for question trials, one regressor for null events, and the four nuisance regressors with the parameters POSITION IN SEQUENCE, GAP LENGTH, PECULIARITY, and OBJECT RECENCY (see Section 2.5). All parameters were z-scored. For the analysis of the main effect of goal coherence, the first trial of each episode was analyzed as belonging to the NC condition, as it was not recognizable as belonging to a specific condition. In the parametric analysis all trials were included, as we were interested in the change of activity during the unfolding of the whole episode. A Gaussian kernel of dispersion of 4 s FWHM was applied on the model equation, including the observation data, the design matrix, and the error term, to account for the temporal autocorrelation [29]. Contrast images, i.e., beta value estimates of the raw-score differences between specified conditions, were generated for each subject. As the individual functional datasets were aligned to the same stereotactic reference space, the single-subject contrast images were entered into a second-level random effects analysis for each of the contrasts.

For the group analyses one-sample *t* tests were used across the contrast images of all participants that indicated whether observed differences between conditions were significantly distinct from zero. The *t* values were then transformed into *Z* scores. To correct for false-positive results, in a first step, an initial voxel-wise *z*-threshold was set to $Z=2.576$ ($p=0.005$). In a second step, the

results were corrected for multiple comparisons using cluster-size and cluster-value thresholds obtained by Monte Carlo simulations at a significance level of $p<0.05$, i.e., the reported activations are significantly activated at $p<0.05$, corrected for multiple comparisons at the cluster level. To specifically test the hypothesis that IFG is modulated as function of goal predictability, a small volume correction was performed [30]. Left and right IFG were segmented in the 3D T1-weighted whole brain scans according to anatomical landmarks based on Talairach and Tournoux [27]. The volume of both left and right IFG was used to calculate the minimum cluster size at a significance level of $p=0.05$. Parametric *z*maps were corrected using a voxel-wise threshold of $Z=2.576$ ($p=0.005$), with a minimum activation area of 289 mm³. Conjunctions were calculated by extracting the minimum *Z* value of the two input contrasts for each voxel [31].

To validate if the attenuation observed in the target regions follows a logarithmic function as predicted by the model of goal predictability, we plotted the time course of activation change in left IFG and bilateral OTC from the first to the last action step of the episodes. To this end, we measured the mean % signal changes from each action step of the episodes separately (position in sequence, i.e., 1 = first action step, 2 = second action step, etc). Design matrices containing six regressors for action steps 1–6. Note that some episodes contained 7 action steps. The seventh action step was treated as position in sequence = 6. In addition, design matrices contained one regressor for all remaining action videos (amplitude = 1) and the 4 nuisance regressors (see Section 2.5). Design matrices were generated for BC and GC episodes separately. ROIs were defined by detecting the peak coordinates in left IFG (BA 45, BA 47) and left and right OTC in the group contrast [all action videos (all conditions) vs. baseline]. For each action step, mean % signal changes (average BOLD response 4–9 s after video onset) were extracted from the peak voxel plus six adjacent voxels.

3. Results

3.1. Parametric effects of action goal predictability

To assess neural responses associated with the predictability of the action goals during unfolding episodes, we computed parametric contrasts with the factor GOAL PREDICTABILITY in the goal-coherent conditions BC and GC. For BC episodes, we found a parametric decrease of the BOLD response in bilateral IFG (BA 45), bilateral supramarginal gyrus (SMG), right postcentral gyrus, bilateral OTC, and bilateral parahippocampal cortex (PHC; Fig. Table 2). For GC episodes we found a parametric decrease of BOLD response in bilateral OTC and PHC. A conjunction of these two contrasts revealed overlapping effects of goal predictability in left IFG (BA 45), OTC and PHC (parametric decrease; Fig. 2A). Effects in IFG were corrected for multiple comparisons restricted to left and right IFG, all other effects also survived corrections for multiple comparisons in the whole brain. Plots of mean percent signal changes for each action step illustrate that neural activity in left IFG and bilateral OTC attenuates in a nonlinear, logarithmic way (Fig. 2C). We will discuss attenuation effects that are significant for both BC and GC episodes (i.e., left IFG and bilateral OTC).

3.2. Main effect of goal coherence

To investigate brain regions involved in the presence of goal coherence we contrasted actions that together form an episode with a common overarching goal with actions that were not related to each other with respect to an overarching goal. This was done separately for episodes performed by the same or by varying actors. We then computed the conjunction of these two separate contrasts

Table 2
Areas activated as a function of goal predictability in unfolding goal-coherent action episodes.

Macroanatomical specification	Brodmann area	Talairach coordinates			Z	
		x	y	z		
<i>Parametric decrease for goal predictability in goal + actor coherent sequences (pBC)</i>						
L	IFG	45	–50	24	0	–3.10*
R	IFG	45	49	27	12	–3.18*
L	IFG/OFC	47/11	–41	33	–15	–3.36*
R	prCG/prCS	6	37	3	27	–4.30
R	SMG	40	49	–27	33	–3.51
L	OTC	37/19	–40	–60	–3	–4.75
R	OTC	37/19	50	–63	12	–3.61
L	OC	18	–38	–81	0	–4.15
R	OC	18	31	–75	–3	–4.47
L	PHC	36/37	–32	–51	–15	–4.77
R	PHC	36/37	28	–48	–15	–3.62
<i>Parametric decrease for goal predictability in goal coherent sequences (pGC)</i>						
L	IFG	45	–53	24	3	–2.95*
L	IFG/OFC	47/11	–32	33	–12	–4.91*
L	OTC	37/19	–41	–60	6	–3.91
R	OTC	37/19	49	–57	–6	–3.55
L	OC	18	–38	–81	0	–4.66
R	OC	18	26	–69	–9	–4.33
L	PHC	36/37	–35	–42	–15	–3.80
R	PHC	36/37	28	–48	–12	–3.96
<i>Parametric decrease for goal predictability in goal coherent sequences (conjunction pBC ∩ pGC)</i>						
L	IFG	45	–50	24	3	–2.61*
L	IFG/OFC	47/11	–29	30	–12	–2.77*
L	OTC	37/19	–44	–69	6	–3.53
R	OTC	37/19	49	–60	–6	–3.36
L	OC	18	–38	–81	–3	–4.15
R	OC	18	25	–69	–6	–4.00
L	PHC	36/37	–35	–42	–15	–3.80
R	PHC	36/37	28	–48	–12	–3.56

Hemisphere (left, right), macroanatomical specification, Brodmann area, Talairach coordinates (x, y, z), and maximal Z scores (Z); z-thresholded at 2.576, corrected cluster threshold $p=0.05$ (asterisks indicate small volume correction for left and right IFG). Abbreviations: L=left; R=right; BA=Brodmann area; IFG=inferior frontal gyrus, OC=occipital cortex, OFC=orbitofrontal cortex, OTC=occipitotemporal cortex, PHC=parahippocampal cortex, prCG/prCS=precentral gyrus/sulcus, SMG=supramarginal gyrus.

(BC vs. AC ∩ GC vs. NC) to determine regions specifically active when participants encounter goal coherence.

Goal-coherent vs. goal-incoherent actions performed by a single actor (BC vs. AC) or different actors (GC vs. NC) revealed bilateral activations in precuneus, middle frontal gyrus/superior frontal sulcus (MFG/SFS), and angular gyrus. In addition, we found bilateral activation in the orbitofrontal cortex (OFC), middle temporal gyrus (MTG) and the cerebellum (Table 3). All regions overlapped as revealed in a conjunction contrast (Fig. 3A).

In addition we performed a ROI analysis in the common network to further assess the effects of goal coherence in relation to actor coherence, including putative superadditive effects of both goal and actor coherence (see Fig. 3B). To this end we extracted beta values of all four conditions in the regions activated for goal-coherent vs. incoherent episodes. We ensured independence of data selection by computing the conjunction BC vs. AC ∩ GC vs. NC for odd and even episodes separately [32]. Peak voxels were detected in the conjunction of even BC and GC episodes while beta values were estimated in the odd BC and GC episodes (from the peak voxel plus six adjacent voxels), and vice versa.

An ANOVA with the factors GOAL COHERENCE and ACTOR COHERENCE revealed main effects for goal coherence in all tested ROIs (Table 4). No main effects for actor coherence and no interaction effects were observed. Accordingly, we found no evidence for superadditive effects of goal and actor coherence.

4. Discussion

The recognition of an action may benefit from being embedded in episodes, as episodes often entail overarching goals the actor

pursues and hence foster the observer's preparedness for upcoming actions. The present fMRI study aimed to assess whether episodes impinge on the processing of observed actions and if so in which way. We found that the embedding into episodes, and hence the presence of goal coherence, influenced BOLD activity in two ways.

First, regions involved in action perception, in particular IFG and OTC, showed a decrease in activation during unfolding goal-coherent episodes. This attenuation conforms the expected data pattern reflecting a facilitation of action processing furnished by the presence of an overarching goal.

Second, there was a set of areas that showed increased activation during the observation of goal-coherent vs. incoherent episodes. This bilateral network consisted of dlPFC, OFC, AG, precuneus, and MTG. We will discuss the idea that these regions are involved in the detection of goal coherence.

4.1. Facilitation of action perception as a function of goal predictability

Goal predictability was hypothesized to modulate neural activity in several regions associated with action perception. Based on previous findings [6,11], we expected that IFG activity decreases the better an overarching long-term goal can be anticipated from preceding goal-coherent action steps. Our results confirmed this expectation. We propose that IFG attenuation reflects reduced neurocognitive load in integrating the currently observed action into the episode, whose long-term goal becomes more and more evident in the unfolding episode as reflected by goal predictability. Our finding complements previous studies that probed the opposite effect: IFG shows increased activity when it is difficult to find

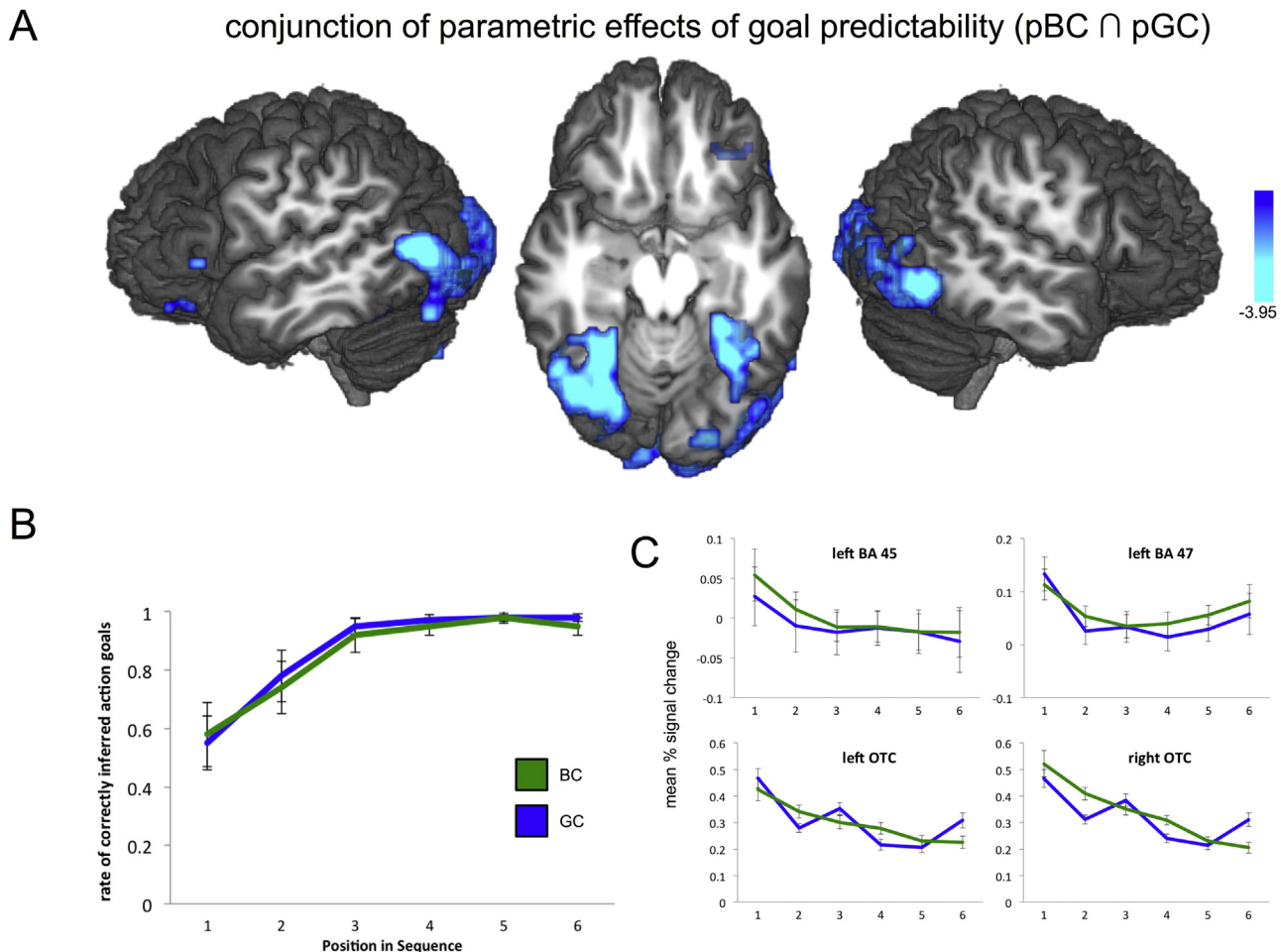


Fig. 2. (A) Parametric effects of goal predictability during unfolding of goal-coherent episodes (conjunction of goal + actor-coherent and goal-coherent). Negative effects indicate attenuation of BOLD response from the first to the last action step of the episode as a function of goal predictability. Statistical maps are corrected for multiple comparisons ($z > 2.576$, corrected cluster threshold $p = 0.05$). (B) Rates of correctly inferred goals used a parametric regressor of goal predictability from separate rating experiment ($N = 10$). Error bars indicate standard error of mean. For details see Section 2.3. (C) Mean % signal change for each action step separately (1 = first action step, 2 = second action step, etc). Error bars indicate standard error of mean.

an overarching goal, e.g., when the observed action is incompatible with action goals implied by the room in which the action takes place [11], or when goal-incoherent actions are expected to belong together because they are performed by the same actor [6]. Together, our findings support the view that IFG's role in action observation is the reconstruction of possible overarching long-term goals via retrieval and selection of semantic information [33] relevant for the integration of an action into an overarching action structure [34]. In a broader view, this interpretation corroborates recent notions of IFG as potential substrate for integration of perceptually distinct elements into hierarchically ordered structures [35–37]. Indeed, there are striking similarities of hierarchical organization in language and complex action sequences such as object manipulation ([36]; but see [38]). Our findings complement the existing evidence from the linguistic domain on IFG's involvement in embedding discrete elements into overarching syntactic structures.

Reconstructing the goal of an action is proposed to subserve the anticipation of the further progress of an action [39]. Following predictive coding accounts [40,41], the anticipation of an action step is considered to amount to a top-down mediated pre-activation of a hierarchically lower region. This anticipation of sensory input supersedes bottom-up processing in the action perception cascade. In our study we found massive attenuation of the BOLD activity in the left and right OTC, as well as in the parahippocampal

cortex, varying as a function of goal predictability. This finding suggests that during the unfolding of action episodes, visual input is effectively filtered as a function of action goal predictability. A similar attenuation effect has been observed in monkey STS during the observation of own movements [42] and in human inferior temporal cortex in response to predictable visual input [43]. The present findings provide evidence for similar attenuation effects in action perception in response to action goal predictability. Notably, predictability effects were observed even when the episode was interrupted by unrelated actions. This suggests that certain aspects of the observed actions were maintained in episodic memory and affected the neural processing of temporally separate actions. It is therefore unlikely that the observed attenuation effects in OTC are the result of mere stimulus-driven repetition suppression.

Notably, attenuation effects were stronger in OTC compared to IFG suggesting different causes for attenuation in OTC (top-down predictive suppression in OTC) and IFG (reduced load in integrating the currently observed action into the episode).

4.2. Presence of goal coherence during action observation

In the present study we assessed the main effect of overarching goals providing coherence to temporally separate actions by contrasting goal-coherent actions vs. goal-incoherent actions. We found that the presence of a common overarching goal was

Table 3
Areas activated for the main effect of goal coherence.

Macroanatomical specification		Brodmann area	Talairach coordinates			Z
			x	y	z	
<i>Goal coherence vs. incoherence, same actor (BC vs. AC)</i>						
L	Precuneus	7	–2	–60	54	5.50
R	Precuneus	7	10	–66	54	5.43
L	AG	39/7	–41	–42	45	4.33
R	AG	39/7	43	–42	54	3.96
L	MFG/SFS	9	–23	27	45	4.64
R	MFG/SFS	9	25	12	48	4.49
L	OFC	11	–23	54	3	4.21
R	OFC	11	22	42	–12	4.12
L	MTG	20/21	–50	–30	–18	4.69
	ITG	20/21	–50	–6	–27	2.98
R	MTG	20/21	46	–30	–18	3.72
L	Lingual gyrus	18	–14	–99	9	4.93
R	Lingual gyrus	18	7	–90	3	3.76
R	Hippocampus		25	–21	–15	3.56
R	Cerebellum		22	–63	–24	3.79
<i>Goal coherence vs. incoherence, different actors (GC vs. NC)</i>						
L	Precuneus	7	14	–63	45	4.57
R	Precuneus	7	4	–69	39	5.14
L	AG	39/7	–38	–48	30	3.80
R	AG	39/7	40	–48	45	4.77
L	MFG/SFS	9	–35	15	48	2.97
R	MFG/SFS	9	43	24	33	3.15
L	OFC	11	–32	60	3	2.82
R	OFC	11	25	54	3	3.05
L	MTG	20/21	–50	–33	–6	3.69
R	MTG	20/21	55	–30	3	4.31
L	aSTG	38	–53	3	0	2.89
L	Cerebellum		–32	–75	–33	5.62
R	Cerebellum		19	–78	–24	3.26
<i>Goal coherence vs. incoherence (conjunction BC vs. AC ∩ GC vs. NC)</i>						
L	Precuneus	7	–14	–63	45	4.38
R	Precuneus	7	7	–57	42	3.97
L	AG	39/7	–35	–60	36	3.49
R	AG	39/7	40	–45	54	3.92
L	MFG/SFS	9	–35	15	48	2.97
R	MFG/SFS	9	43	24	33	3.15
L	OFC	11	–32	60	3	2.82
R	OFC	11	25	54	3	3.05
L	MTG	20/21	–50	–30	–9	3.06
R	MTG	20/21	49	–30	–21	3.13

Hemisphere (left, right), macroanatomical specification, Brodmann area, Talairach coordinates (x, y, z), and maximal Z scores (Z); z-thresholded at 2.576, corrected cluster threshold $p < 0.05$. Abbreviations: L = left; R = right; BA = Brodmann area; AG = angular gyrus, aSTG = anterior superior temporal gyrus, IFG, inferior temporal gyrus, MFG = middle frontal gyrus, MTG = middle temporal gyrus, OFC = orbitofrontal cortex, OTG = occipitotemporal cortex, SFS = superior frontal sulcus.

accompanied with increased neural activity in a number of brain regions in both hemispheres, including dIPFC (MFG/SFS), OFC, AG, precuneus, and MTG. These activations were not modulated by the constancy of actor (see Fig. 2). Thus, activity increase in these regions was only driven by the presence of an overarching goal embedding separate actions into a coherent episode.

To date, neural effects of coherence have been investigated mostly with respect to text comprehension [22,44–46]. Ferstl et al.

[47] argued that establishing coherence between two sentences (e.g., “sort every thing into piles” and “powder is needed”) often requires a global context or *situation model* (e.g., “washing laundry”). In the present study, a situation model was maintained by the presence of an overarching goal. Interestingly, we found activation for goal coherence overlapping with regions activated for coherence in text comprehension. These regions were the left precuneus and the left MTG [22]. This overlap suggests that coherence-related

Table 4
Specificity of goal coherence in comparison to actor coherence.

	Goal coherence		Actor coherence		Interaction	
	$F_{(1,17)}$	p	$F_{(1,17)}$	p	$F_{(1,17)}$	p
Left OFC	15.241	0.001	0.202	0.659	0.531	0.476
Right OFC	18.47	<0.001	0.11	0.744	0.677	0.422
Left MFG/SFS	11.344	0.004	3.202	0.091	1.129	0.303
Right MFG/SFS	13.764	0.002	1.103	0.308	2.25	0.152
Left MTG	18.704	<0.001	0.36	0.556	0.199	0.661
Right MTG	12.403	0.003	1.003	0.331	3.098	0.096
Left AG	17.68	0.001	0.677	0.422	0.133	0.72
Right AG	14.636	0.001	2.876	0.108	0.649	0.432
Left precuneus	55.389	<0.001	0.48	0.829	1.44	0.247
Right precuneus	17.084	0.001	0.518	0.481	2.905	0.107

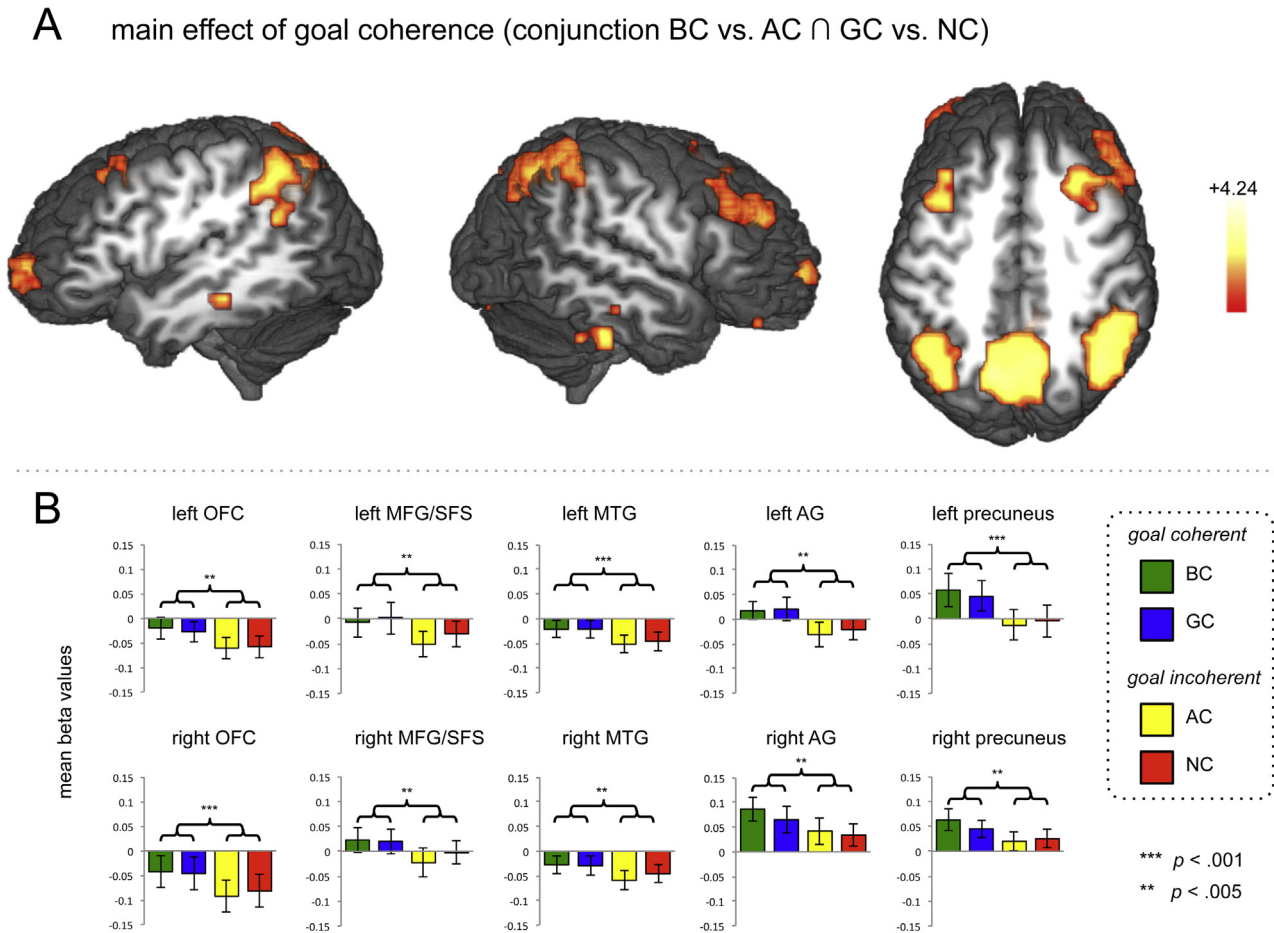


Fig. 3. Main effects of goal coherence (conjunction of goal + actor-coherent and goal-coherent episodes). Statistical maps are corrected for multiple comparisons ($z > 2.576$, corrected cluster threshold $p = 0.05$). Effects were specific for goal coherence in comparison with actor coherence as shown in a separate ROI analysis (B). Error bars indicate standard error of mean.

activity in these regions is modality-independent. In particular, precuneus has been suggested to be the central hub for maintenance of situation models by linking semantic knowledge (here: script knowledge, e.g., how to prepare a sandwich) to episodic knowledge (here: goal-coherent action steps of previous trials) [48]. A possible substrate for semantic action script knowledge is the MTG [49].

Divergent activations for goal coherence in observed action as compared to coherence in text comprehension were found in dlPFC, AG, and OFC. Activation of dlPFC and AG likely reflects the retrieval from episodic memory as required in our study but not so or much less in Ferstl's [19], as actions steps making up one episode were intermixed by action steps from other episodes. Correspondingly, the same set of areas is activated at action boundaries (as compared to non-goal-directed boundaries) in ongoing action sequences [50]. At these points, expectations have to be updated by a recollection of possible upcoming action steps in episodic memory. With reference to priority maps [51], AG, under guidance of dlPFC, functions as a searchlight on the episode model established and maintained by precuneus.

5. Conclusion

Observing action episodes with a common overarching goal comes with widespread changes of BOLD activity both within and outside regions involved in action observation. These changes occurred automatically, i.e., without explicit requirement by the task, and even if goal-coherent action steps were performed by different actors and were interrupted by other unrelated action

steps. We propose that attenuation in IFG reflects that actions of an episode can be successfully integrated under a common goal and OTC attenuates accordingly as following action steps can be anticipated. Increased activity in dlPFC, precuneus and angular gyrus indicate that single action steps have to be buffered and concatenated with recourse to episodic memory.

Acknowledgements

This research was supported by the Center for Mind/Brain Sciences of the University of Trento, Italy, the University of Münster, Germany, and the Max Planck Society.

References

- [1] Botvinick MM. Hierarchical models of behavior and prefrontal function. *Trends Cogn Sci* 2008;12:201–8.
- [2] Byrne RW, Russon AE. Learning by imitation: a hierarchical approach. *Behav Brain Sci* 1998;21:667–84, discussion 684–721.
- [3] Cooper R, Shallice T. Contention scheduling and the control of routine activities. *Cogn Neuropsychol* 2000;17:297–338.
- [4] Cooper R, Shallice T. Hierarchical schemas and goals in the control of sequential behavior. *Psychol Rev* 2006;113:887–916, discussion 917–831.
- [5] Wurm MF, von Cramon DY, Schubotz RI. Do we mind other minds when we mind other minds' actions? A functional magnetic resonance imaging study. *Hum Brain Mapp* 2011;32:2141–50.
- [6] Hrkač M, Wurm MF, Schubotz RI. Action observers implicitly expect actors to act goal-coherently, even if they do not: an fMRI study. *Hum Brain Mapp* 2014;35:2178–90.
- [7] Badre D, Poldrack RA, Pare-Blagoev EJ, Insler RZ, Wagner AD. Dissociable controlled retrieval and generalized selection mechanisms in ventrolateral prefrontal cortex. *Neuron* 2005;47:907–18.

- [8] Gold BT, Balota DA, Jones SJ, Powell DK, Smith CD, Andersen AH. Dissociation of automatic and strategic lexical-semantics: functional magnetic resonance imaging evidence for differing roles of multiple frontotemporal regions. *J Neurosci* 2006;26:6523–32.
- [9] Moss HE, Abdallah S, Fletcher P, Bright P, Pilgrim L, Acres K, et al. Selecting among competing alternatives: selection and retrieval in the left inferior frontal gyrus. *Cereb Cortex* 2005;15:1723–35.
- [10] Schubotz RI, von Cramon DY. The case of pretense: observing actions and inferring goals. *J Cogn Neurosci* 2009;21:642–53.
- [11] Wurm MF, Schubotz RI. Squeezing lemons in the bathroom: contextual information modulates action recognition. *NeuroImage* 2012;59:1551–9.
- [12] Koechlin E, Jubault T. Broca's area and the hierarchical organization of human behavior. *Neuron* 2006;50:963–74.
- [13] Jeannerod M. Neural simulation of action: a unifying mechanism for motor cognition. *NeuroImage* 2001;14:S103–9.
- [14] Cross ES, Liepelt R, Hamilton AF, Parkinson J, Ramsey R, Stadler W, et al. Robotic movement preferentially engages the action observation network. *Hum Brain Mapp* 2012;33:2238–54.
- [15] Schiffer AM, Ahlheim C, Ulrichs K, Schubotz RI. Neural changes when actions change: adaptation of strong and weak expectations. *Hum Brain Mapp* 2013;34:1713–27.
- [16] Keysers C, Perrett DI. Demystifying social cognition: a Hebbian perspective. *Trends Cogn Sci* 2004;8:501–7.
- [17] Kok P, Rahnev D, Jehee JF, Lau HC, de Lange FP. Attention reverses the effect of prediction in silencing sensory signals. *Cereb Cortex* 2012;22:2197–206.
- [18] Summerfield C, Egner T. Expectation (and attention) in visual cognition. *Trends Cogn Sci* 2009;13:403–9.
- [19] Cabeza R, Ciaramelli E, Olson IR, Moscovitch M. The parietal cortex and episodic memory: an attentional account. *Nat Rev Neurosci* 2008;9:613–25.
- [20] Donaldson DI, Wheeler ME, Petersen SE. Remember the source: dissociating frontal and parietal contributions to episodic memory. *J Cogn Neurosci* 2010;22:377–91.
- [21] Rugg MD, Otten LJ, Henson RN. The neural basis of episodic memory: evidence from functional neuroimaging. *Philos Trans R Soc London B Biol Sci* 2002;357:1097–110.
- [22] Ferstl EC, Neumann J, Bogler C, von Cramon DY. The extended language network: a meta-analysis of neuroimaging studies on text comprehension. *Hum Brain Mapp* 2008;29:581–93.
- [23] Oldfield RC. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 1971;9:97–113.
- [24] Norris DG. Reduced power multislice MDEFT imaging. *J Magn Reson Imaging* 2000;11:445–51.
- [25] Ugurbil K, Garwood M, Ellermann J, Hendrich K, Hinke R, Hu X, et al. Imaging at high magnetic fields: initial experiences at 4T. *Magn Reson Q* 1993;9:259–77.
- [26] Lohmann G, Muller K, Bosch V, Mentzel H, Hessler S, Chen L, et al. LIPSIA—a new software system for the evaluation of functional magnetic resonance images of the human brain. *Comput Med Imaging Graph* 2001;25:449–57.
- [27] Talairach J, Tournoux P. Co-planar stereotaxic atlas of the human brain. New York, NY: Thieme; 1988.
- [28] Friston KJ, Holmes AP, Worsley KJ, Poline JB, Frith CD, Frackowiak RS. Statistical parametric maps in functional imaging: a general linear approach. *Hum Brain Mapp* 1995;2:189–210.
- [29] Worsley KJ, Friston KJ. Analysis of fMRI time-series revisited—again. *NeuroImage* 1995;2:173–81.
- [30] Poldrack RA. Region of interest analysis for fMRI. *Soc Cogn Affect Neurosci* 2007;2:67–70.
- [31] Nichols T, Brett M, Andersson J, Wager T, Poline JB. Valid conjunction inference with the minimum statistic. *NeuroImage* 2005;25:653–60.
- [32] Kriegeskorte N, Simmons WK, Bellgowan PS, Baker CI. Circular analysis in systems neuroscience: the dangers of double dipping. *Nat Neurosci* 2009;12:535–40.
- [33] Badre D, Wagner AD. Left ventrolateral prefrontal cortex and the cognitive control of memory. *Neuropsychologia* 2007;45:2883–901.
- [34] Koechlin E, Summerfield C. An information theoretical approach to prefrontal executive function. *Trends Cogn Sci* 2007;11:229–35.
- [35] Fiebach CJ, Schubotz RI. Dynamic anticipatory processing of hierarchical sequential events: a common role for Broca's area and ventral premotor cortex across domains? *Cortex* 2006;42:499–502.
- [36] Tettamanti M, Weniger D. Broca's area: a supramodal hierarchical processor? *Cortex* 2006;42:491–4.
- [37] Vicari G, Adenzato M. Is recursion language-specific? Evidence of recursive mechanisms in the structure of intentional action. *Conscious Cogn* 2014;26:169–88.
- [38] Moro A. On the similarity between syntax and actions. *Trends Cogn Sci* 2014;18:137–60.
- [39] Csibra G. Action mirroring and action interpretation: an alternative account. In: Haggard P, Rosetti Y, Kawato M, editors. *Sensorimotor foundations of higher cognition. Attention and performance XXII*. Oxford: Oxford University Press; 2007. p. 435–59.
- [40] Friston KJ, Mattout J, Kilner JM. Action understanding and active inference. *Biol Cybern* 2011;104:137–60.
- [41] Kilner JM, Friston KJ, Frith CD. Predictive coding: an account of the mirror neuron system. *Cogn Process* 2007;8:159–66.
- [42] Hietanen JK, Perrett DI. Motion sensitive cells in the macaque superior temporal polysensory area: I. Lack of response to the sight of the animal's own limb movement. *Exp Brain Res* 1993;93:117–28.
- [43] Summerfield C, Trittschuh EH, Monti JM, Mesulam MM, Egner T. Neural repetition suppression reflects fulfilled perceptual expectations. *Nat Neurosci* 2008;11:1004–6.
- [44] Ferstl EC, von Cramon DY. Time, space and emotion: fMRI reveals content-specific activation during text comprehension. *Neurosci Lett* 2007;427:159–64.
- [45] Virtue S, Haberman J, Clancy Z, Parrish T, Jung Beeman M. Neural activity of inferences during story comprehension. *Brain Res* 2006;1084:104–14.
- [46] Virtue S, Parrish T, Jung-Beeman M. Inferences during story comprehension: cortical recruitment affected by predictability of events and working memory capacity. *J Cogn Neurosci* 2008;20:2274–84.
- [47] Ferstl EC, von Cramon DY. The role of coherence and cohesion in text comprehension: an event-related fMRI study. *Brain Res Cogn Brain Res* 2001;11:325–40.
- [48] Binder JR, Desai RH. The neurobiology of semantic memory. *Trends Cogn Sci* 2011;15:527–36.
- [49] Watson CE, Cardillo ER, Ianni GR, Chatterjee A. Action concepts in the brain: an activation likelihood estimation meta-analysis. *J Cogn Neurosci* 2013;25:1191–205.
- [50] Schubotz RI, Korb FM, Schiffer AM, Stadler W, von Cramon DY. The fraction of an action is more than a movement: neural signatures of event segmentation in fMRI. *NeuroImage* 2012;61:1195–205.
- [51] Bisley JW, Goldberg ME. Attention, intention, and priority in the parietal lobe. *Annu Rev Neurosci* 2010;33:1–21.