



## The fraction of an action is more than a movement: Neural signatures of event segmentation in fMRI

Ricarda I. Schubotz<sup>a,b,c,\*</sup>, Franziska M. Korb<sup>a,d</sup>, Anne-Marike Schiffer<sup>b</sup>,  
Waltraud Stadler<sup>a,e</sup>, D. Yves von Cramon<sup>a,b</sup>

<sup>a</sup> Max Planck Institute for Human Cognitive and Brain Sciences, Department of Cognitive Neurology, 04103 Leipzig, Germany

<sup>b</sup> Max Planck Institute for Neurological Research, Motor Cognition Group, 50931 Cologne, Germany

<sup>c</sup> University of Münster, Institute for Psychology, 48149 Münster, Germany

<sup>d</sup> Duke University, Center for Cognitive Neuroscience, Durham, NC, USA

<sup>e</sup> Lehrstuhl für Bewegungswissenschaft, Fakultät für Sport- und Gesundheitswissenschaft, Technische Universität München, 80809 Munich, Germany

### ARTICLE INFO

#### Article history:

Accepted 3 April 2012

Available online 13 April 2012

#### Keywords:

Event segmentation

Action

Superior frontal sulcus

Angular gyrus

Parahippocampal cortex

fMRI

### ABSTRACT

When we observe an action, we recognize meaningful action steps that help us to predict probable upcoming action steps. This segmentation of observed actions, or more generally events, has been proposed to rely in part on changes in motion features. However, segmentation of actions, in contrast to meaningless movements, may exploit additional information such as action knowledge. The present fMRI study sought to tear apart the neural signatures of processing two sources of information that observers may exploit at action boundaries: change in motion dynamics and action knowledge. To this end, subjects performed a segmentation task on both actions (that can be segmented based on motion and action knowledge) as well as tai chi movements (that can be segmented only based on motion) and two further control conditions that implemented point-light walker like displays of the same videos. Behavioral tests showed that motion features played a critical role in boundary detection in all conditions. Consistent with this finding, activity in area MT was enhanced during boundary detection in all conditions, but importantly, this effect was not stronger for actions. In contrast, only action boundary detection was reflected by specific activation in the superior frontal sulcus, parietal angular gyrus and the parahippocampal cortex. Based on these findings, we propose that during action observation, motion features trigger a top-down modulation of the attentional focus and the incitement of retrieving long-term memory place-action associations. While action perception entails activity common to processing of all motion stimuli, it is at the same time unique as it allows long-term memory based predictions of succeeding steps.

© 2012 Elsevier Inc. All rights reserved.

### Introduction

A considerable amount of what we perceive consists of dynamic change. Motion, language, music or actions can all be conceived of as providing an ongoing stream of structured perceivable change. From a biological perspective, events entail change, and catch our particular attention because they often imply the need to adapt one's behavior; therefore, perceiving events leads to the build-up of implicit or explicit expectations (Kurby and Zacks, 2008; Schiffer et al., 2012; Schubotz, 2007; Schütz-Bosbach and Prinz, 2007).

However, some fractions of such events are more predictive of their respective next instance than others, and thus, to an observer, prediction of the event undergoes fluctuations of more certain and less certain phases. For actions, this has been demonstrated in a series of behavioral studies using the *unit marking procedure* developed by Newton (1973). In this paradigm, subjects were asked to press a response button when they judged one meaningful action part to end and the next to start. The time points indicated by these judgments are called event boundaries. Using this approach, Newton and Engquist (1976) showed (a) that perceivers are better in detecting manipulations of the stimulus material that occur at boundaries, (b) that they rely on boundaries when reconstructing an action's meaning, and (c) that recognition memory is enhanced for event boundaries compared to the remaining components of an event. The authors argued that the objective quality of boundaries is that within the continuous sequence boundaries convey a higher amount of information than the remainder of the event, called non-boundaries

\* Corresponding author at: University of Münster, Institute for Psychology, Flieönerstr. 21, 48149 Münster, Germany. Fax: +49 251 83 34110.

E-mail addresses: [rschubotz@uni-muenster.de](mailto:rschubotz@uni-muenster.de) (R.I. Schubotz), [franziska.korb@duke.edu](mailto:franziska.korb@duke.edu) (F.M. Korb), [schiffer@nf.mpg.de](mailto:schiffer@nf.mpg.de) (A.-M. Schiffer), [waltraud.stadler@tum.de](mailto:waltraud.stadler@tum.de) (W. Stadler), [cramon@nf.mpg.de](mailto:cramon@nf.mpg.de) (D.Y. von Cramon).

hereafter (Newtson and Engquist, 1976). In this regard, action boundaries compare to chunk boundaries in general. Both can be described as the first elements to sequentially grouped bits of information, which are considered to be charged with information about all upcoming (chunk) elements (Koch and Hoffmann, 2000; Restle and Brown, 1970; Rosenbaum et al., 1983).

While action observation has attracted considerable attention in human imaging research (Decety and Grezes, 1999; Rizzolatti et al., 2001), most studies have focused on very short actions of limited complexity, i.e., single segments or chunks. Thus, action has been treated as one homogeneous bit of information, not as a structured event. Hence, we are largely ignorant of the neural processes that occur at or around action boundaries. Recently, Zacks and co-workers (Zacks and Sargent, 2009; Zacks et al., 2007) have put forward the *Event Segmentation Theory*, proposing that a set of event models bias processing in the perceptual stream. The authors propose that prediction undergoes fluctuations of certainty: when an event boundary is approached, prediction certainty declines leading to a gating mechanism and thus an update of the current event model (Zacks et al., 2007). Zacks and co-workers (Speer et al., 2003; Zacks et al., 2001) report motion selective area MT and the junction of the left inferior frontal and precentral sulcus to show enhanced activity when subjects perceive boundaries in everyday action. The finding was interpreted as indicating that event segmentation could be driven by prominent features of the observed trajectories, i.e., by perceptual rather than top-down cognitive processing (Zacks et al., 2001). Greater motion changes at event boundaries, as an example of prominent motion features, may thus directly trigger MT activation, which could in turn propagate to prefrontal areas (Speer et al., 2003).

While it is conceivable that the sight of objects or characteristic movement patterns of the body or the manipulating hands triggers the recognition of the upcoming action step more or less directly, there is usually a set of possible next action steps, and hence competition among these options has to be resolved. Since probabilities of these different action options differ, the selection process may be realized by a Bayesian function of probabilistically weighted forward models (Körding and Wolpert, 2006). Moreover, purely stimulus-triggered recognition implies a process following the stimulus, whereas the notion of predictive perception (Bubic et al., 2009; Grush, 2004; Schiffer and Schubotz, 2011; Zacks et al., 2011) means preparation for early selection among available cues. Here, a frontal top-down signal of transiently enhanced control or selection mechanisms would be expected (Miller and Cohen, 2001; O'Reilly et al., 2002; Ridderinkhof et al., 2004a,b). The capacity to predict ongoing action relies on learnt event models (Friston et al., 2011; Neal and Kilner, 2010; Schiffer et al., 2012), therefore activity would be expected in long-term memory-related areas. Further evidence for this long-term memory influence lies in differences between adults' and children's segmentation performances (Baird and Baldwin, 2001).

We agree that action boundaries are relevant nodes for an observer of a continuous action stream, and at these nodes, an update of the internal or event model occurs, probably triggered in part by changes in motion features (Zacks et al., 2001). However, we took a somewhat different view in that we explicitly wished to control for the influence or usage of movement features as cues for action boundaries. While movement information probably serves as a good cue for action segmentation we propose that action script knowledge has to be a further source of information that is exploited here. Changes in movement features may also occur in other kinds of events, including biological and non-biological motion, and we aimed to find out whether area MT is activated specifically for boundaries in goal-directed action or not.

Moreover, we expected the segmentation of action, but not segmentation in a non-action baseline condition (see below), to draw on prefrontal sites. This hypothesis was based mainly on patient

studies showing that chunking action segments into longer, more complex actions relies predominantly on the anterior frontal lobes (Allain et al., 1999; Sirigu et al., 1995; Zalla et al., 2001). Thus, the prefrontal cortex should contribute to the selection among competing alternative action steps on the basis of action scripts. Moreover, prefrontal cortex has been associated with the suppression of dominant but currently unwarranted decisions/actions (Frank and Claus, 2006; Ghahremani et al., 2010; Kuhl et al., 2007; Miller and Cohen, 2001; Ridderinkhof et al., 2004b), a function that would be required when preceding action steps bias or restrict the expectation of the upcoming action steps.

Participants were presented with movie clips showing an actress performing everyday actions such as ironing a shirt or hanging out the laundry. Actions were of a duration of about 1 min and comprised of several action steps; thus, ironing a shirt included for example carrying the washing basket to the ironing board, placing the shirt on the board, switching on the iron, starting to move the iron over the shirt's sleeves and so forth. We used the unit marking procedure developed by Newtson (1973). That is, participants were asked to press a response button whenever they felt that the action proceeded to its next step. No feedback was provided for this purely subjective judgment.

In order to identify action-specific neural correlates of event segmentation, we implemented a control condition where subjects were asked to perform the same segmentation task as in the action observation while watching movie clips that showed an actor or actress performing whole-body tai chi movements. Action-specific boundary detection was tackled by calculating the interaction contrast [(action boundary vs. action non-boundary) vs. (tai chi boundary vs. tai chi non-boundary)]. Importantly, the tai chi condition allowed us to control for the perception of a human being in motion, as well as for button presses. As the segmentation of tai chi movements based on changes in movement patterns at event boundaries, this condition was moreover used to control for those portions of boundary judgments that merely relied on changes of spatiotemporal dynamics of the body.

To investigate whether area MT is particularly sensitive towards action boundaries as compared to movement boundaries, we implemented a region of interest (ROI) analysis that was based on the area MT coordinates reported by Zacks et al. (2001). We tested whether we could replicate the finding that action boundaries yielded a higher response than non-boundaries in area MT, as reported by Zacks et al. (2001), corresponding to a main effect of factor PART (boundary > non-boundary). Moreover, we wanted to test whether this effect was disproportionately large for actions, corresponding to an interaction between the factors CONDITION (action, tai chi) and PART (boundary, non-boundary) area MT.

Regarding behavioral analyses, we aimed to assess that subjects' response behavior was meaningful and consistent, not random button pressing. This was particularly important because the task we implemented was a subjective judgment; therefore, the recorded responses could be neither right nor wrong. Accordingly, we assessed test-retest reliability of response patterns and comparability of conditions on the single subject as well as on the group level.

## Methods

The present study consisted of a pilot study that included two sessions and one fMRI session. The task (as will be described below) was identical in all sessions. The main idea of the pilot study was to assess *comparability* between the four conditions that were implemented (see below) and test-retest-reliability. As we employed a subjective judgment task, the pilot studies' tests for comparability and retest reliability were important to establish whether participants responded according to an internal and stable criterion.

## Subjects

16 participants took part in the pilot study (8 women, 19–30 yrs, mean 23.8 yrs old). They were recruited from the Max Planck Institute for Human Cognitive and Brain Science's volunteer database and were paid 4€ per 30 min participation.

17 Participants (8 women, 20–31 yrs, mean 24.5 yrs old) partook in the fMRI-study. They were all right handed as assessed with the Edinburgh Handedness Inventory. fMRI participants were recruited from the same database and remunerated with 10€/h. None of the participants reported a record of neurological or psychiatric illnesses, and all of them had normal or corrected to normal vision. The experiment accorded to the standards of the ethics committee of the University of Leipzig and all participants gave informed written consent.

## Stimuli and task

The first half of the stimulus material consisted of 12 movies displaying an actor performing every day actions (action movies hereafter) and 12 movies displaying an actor performing tai chi movements (tai chi movies hereafter). All action movies showed the same (female) actor. The movies were shot from a third person perspective that showed the entire body and face of the actor, as well as the room (and furniture) she was moving in as a backdrop. Actions were, e.g. doing the dishes, taking a picture, or ironing a shirt, and were performed in naturalistic settings, for example the doing the dishes movie was shot in the kitchen. The videos' lengths varied between 43.8 and 143.6 s. Average length was 81.1 s, median 65.7 s. Six of the tai chi movies showed a male, six a female actor. All videos also showed the actor and the room they were moving in as a backdrop. Tai chi movies, however, did not take place in household settings. They were shot in a large seminar room and displayed a dark curtain and no furniture as a backdrop to the movements. Each tai chi movie's length was matched to that of one action clip, lengths accordingly varied between 43.8 and 143.7 s. Average length was 81.1 s, median 65.8.

Apart from these naturalistic action and tai chi movies, the experiment contained point cloud versions of both, the action and tai chi movies. Using the PFTrack software (The Pixel Farm®) we marked the head, left and right shoulder, left and right hand and hip of the actor within each movie clip. Over the length of the clip, the coordinates of each marker were tracked from frame to frame (40 ms length) and saved in an array. During the experiments, which were all run on Presentation (Neurobehavioral Systems, Inc.), the program read out the coordinates from frame to frame, setting one rectangle in place of each marker against a gray background. These rectangles changed position from frame to frame, according to the coordinates that had previously been saved. Thus they moved along the same trajectories as the limbs they corresponded to (Fig. 1). In order to avoid recognition of the corresponding action or tai chi movie, and therefore adding a confound to the experimental design, the point trajectories were flipped along the horizontal screen axis, i.e., the point cloud movies were presented upside down. The choice and order of the four conditions was randomized with an individual order for each participant. In the pilot study, participants viewed 48 movies, 12 of each condition, i.e., action, tai chi, action point cloud and tai chi point cloud movies. In the fMRI session, the participants encountered only six movies of each condition. The movies were selected in a "counter-matched" manner: first, six action movies were chosen randomly. Then the six tai chi movies that did not correspond in length were selected. Finally, all the point clouds were picked that corresponded to the remaining 6 action and 6 tai chi movies not presented. To ensure that all movies were presented equally often throughout the study, for each movie list generated based on the criteria above, a second one was created including all remaining movies. Hence, the two resulting movie lists combined included all action, tai chi and point cloud movies,

which were randomly presented during the scanning sessions. In doing so, it was ensured that a) the experiment was of the same length for each participant, and b) the association of a point cloud movie to the underlying action or tai chi movie was made impossible. Therefore, all analyses that rely on averaging over movies built on twice as many cases in the analysis of the pilot study than in the analysis of the fMRI study.

Participants had to press a button whenever they judged a new action or movement step to begin, i.e., a boundary to have occurred. Their instruction stressed that action and movement steps (and thus boundaries) relate to how a description of the display would sound like ('as if you would tell someone what happens'). The same instruction guided the participants' boundary detection for all four conditions.

## Behavioral data analysis

In the pilot study that consisted of two within subject measures (sessions on average 15 days apart; minimum 14 days, maximum 18 days), participants performed the boundary detection task. The task we implemented was a subjective judgment task, therefore, the recorded responses can be neither right nor wrong. It is therefore particularly important to establish and assess that response patterns in the judgment task are meaningful and consistent instead of random button pressing. We assessed test-retest reliability of response patterns and comparability of conditions on the single subject as well as on the group-aggregated level.

Response rates were calculated as the number of responses for each movie divided by its length. We took two measures to analyze reliability of the response patterns. To deal with variability on the subject level, we employed an individual criterion for each participant that took response rate into account to assess reliability. We assessed the position of all responses for each video in the first session for each participant and tested how often the same participant responded in the second session within 0.25 standard deviations of personal condition-specific mean response rate of the responses in the first pilot session. The 0.25 standard deviations criterion was employed to include only those responses that were significantly close, i.e. concurred to a 10% confidence interval.

As a second measure of retest-reliability, we ran a repeated-measure analysis of variance (ANOVA) between the response rates of the first and the second pilot session, with SESSION as a two-level factor and CONDITION as a four-level factor. Moreover we calculated two univariate ANOVAs on response rates. The first univariate ANOVA modeled the first pilot session and the fMRI session as levels of the two-level factor SESSION and the four conditions as levels of the factor CONDITION. The second univariate ANOVA consisted of the same factors, but with the second pilot session and the fMRI session as levels of the SESSION factor.

In the next step, we aimed to establish whether participants agreed on boundaries, that is whether their responses inter-individually co-occurred. For all subsequent group-based analyses, we employed time-windows, called 'bins' hereafter. We aggregated for each frame (40 ms length) how many responses had occurred within the next bin, consisting of 25 frames (one second) of the respective movie. This aggregation of responses pooled the data of all participants. The first bin consisted of the first 25 frames; the second bin encompassed the second to the 26th frame of the movie and so forth. Only one response of each participant was taken into the bin, so that the highest number of responses that could occur in one bin equaled the number of participants. This number of responses per bin was assigned to the frame that was initial to the 25 frames that contained the responses (frame-value hereafter). Another way of explaining this analysis would be to say that the frame-by-frame data were smoothed with a rectangular kernel of width 1 s.



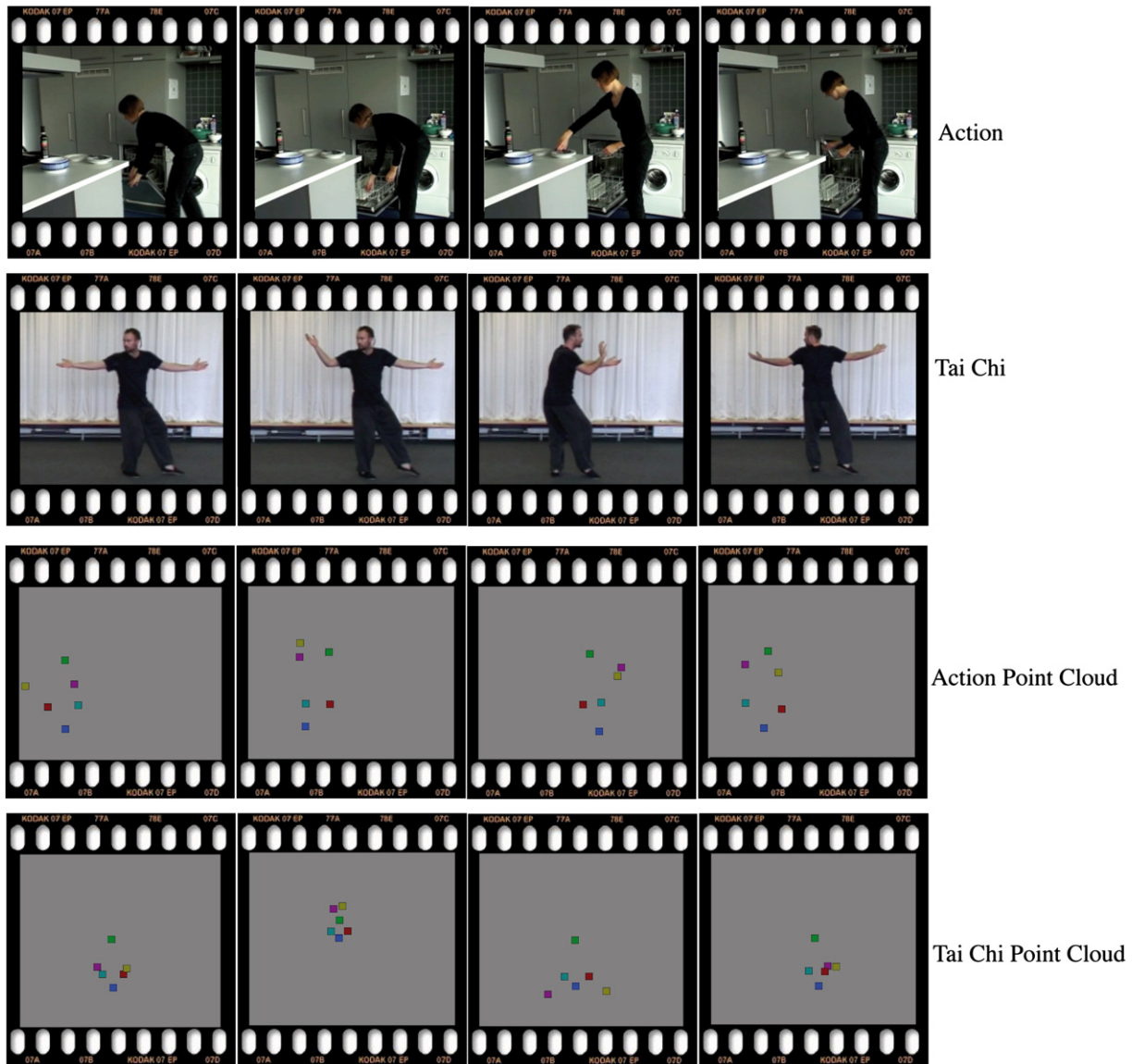


Fig. 1. Examples of the four conditions: action movies, tai chi movies, action point cloud movies and tai chi point cloud movies.

If, for example, the first bin, i.e., frame one to 25 contained eight responses, the bin value was eight and this bin value was assigned to the first frame (frame value = 8). If one response had been within frame one and there was further no response in frame 26, then the second bin, encompassing frames two to 26 carried the value seven and the frame value of the second frame was seven. If then the 27th frame had been responded to three times, the according bin value of the third bin was 10 and the frame value of frame three was accordingly also ten (Fig. 2). Frames were assigned the value zero (“empty frames” hereafter) if no response had occurred within the next bin. The mean number of responses for each movie was calculated as the sum of frame values divided by the number of frames that were not empty, respectively for each specific movie. These mean responses will be referred to as average bin values hereafter, to prevent confusion with individual mean response rates. If an initial frame had a value higher than two standard deviations above the mean for the respective movie it was called a boundary. For a discussion of related methods, see Christoffersen et al. (2007).

As a qualitative measure of comparability between conditions, we first correlated the time series of frame values between associated movies (e.g., the dish-washing action movie with either the tai chi

movie of the same length or with the dish-washing action point cloud movie). In a next step, the *r*-values of all movie correlations were Fisher *z*-transformed and averaged, as a qualitative correlation measure for conditions. This was done separately for all experiments (1st, 2nd pilot; fMRI)

The within-group (as opposed to within-subject) retest-reliability for each movie across sessions was calculated by correlation of frame values of the first and second pilot session. Since there was no retest-

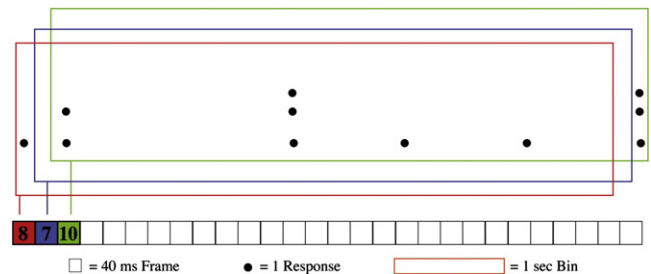


Fig. 2. Frame value assessment. Black dots signify responses. The first frame of a 25-frame bin (1 s) was assigned the number corresponding to all participants' responses within the next bin.

session to the fMRI, we calculated, in an analogy to the retest reliability, a correlation of frame values between fMRI and 1st pilot session.

### MRI data acquisition

Imaging was carried out on a 3T Siemens (Erlangen, Germany) Trio system equipped with a standard birdcage head coil. Participants were placed on the scanner bed in a supine position with their right index finger positioned on the appropriate response button of a response box. Form-fitting cushions were utilized to prevent head, arm, and hand movements. Participants were provided with earplugs to attenuate scanner noise.

Twenty-four axial slices (192 mm field of view;  $64 \times 64$  pixel matrix; 4 mm thickness; 1 mm spacing; in-plane resolution of  $3 \times 3$  mm) parallel to the bicommissural line (AC–PC) covering the whole brain were acquired using a single-shot gradient EPI sequence (2 s repetition time; 30 ms echo time;  $90^\circ$  flip angle; 116 kHz acquisition bandwidth) sensitive to BOLD contrast. Prior to the functional imaging, 26 anatomical T1-weighted MDEFT images (Norris, 2000; Ugurbil et al., 1993) with the same spatial orientation as the functional data were acquired (160 slices of 1 mm thickness).

### fMRI data analysis

After motion correction using rigid-body registration to the central volume, fMRI data was processed using the software package LIPSIA (Lohmann et al., 2001). To correct for the temporal offset between the slices acquired in one image, a cubic-spline interpolation was employed. Low-frequency signal changes and baseline drifts were removed using a temporal high-pass filter with a cut-off frequency of 1/130 Hz. Spatial smoothing was performed with a Gaussian filter of 5.65 mm FWHM. To align the functional data slices with a 3-D stereotactic coordinate reference system, a rigid linear registration with six degrees of freedom (three rotational, three translational) was performed. The rotational and translational parameters were acquired on the basis of the MDEFT slices to achieve an optimal match between these slices and the individual 3-D reference dataset. The MDEFT volume dataset with 160 slices and 1-mm slice thickness was standardized to the Talairach stereotactic space (Talairach and Tournoux, 1988). The rotational and translational parameters were subsequently transformed by linear scaling to the same standard size. The resulting parameters were then used to transform the functional slices employing a trilinear interpolation, to align the resulting functional slices with the stereotactic coordinate system. Resulting data were resampled to a spatial resolution of  $2 \times 2 \times 2$  mm ( $8 \text{ mm}^3$ ).

The statistical evaluation was based on a least-squares estimation using the general linear model for serially auto-correlated observations (Worsley and Friston, 1995). The design matrices were generated with a delta function, convolved with the hemodynamic response function (gamma function). This analysis was based on the events that were time-locked to the participants' button presses and to non-boundaries. Non-boundaries were set at the mean distance between two indicated boundaries. Boundaries that were less than four seconds apart were excluded from the analysis to account for BOLD signal inertia.

The analysis was based on a mixed model design-matrix containing fourteen entries: one for each condition (four epochs), time-locked to movie onset and modeled as block for the entire movie length; one event for the boundaries (time-locked to responses) in each condition (four), modeled as events; one event for the non-boundaries in each condition (four) (time locked to the middle between two boundaries that were at least four seconds apart), modeled as events; and finally, empty trials (so-called null events), and movies that had not been responded to.

The model equation, including the observation data, the design matrices, and the error term, were convolved with a Gaussian kernel of dispersion of 4 s FWHM to account for the temporal autocorrelation (Worsley and Friston, 1995). In the following, contrast images, that is, beta value estimates of the raw-score differences between specified conditions were generated for each participant. The individual functional datasets were aligned to the same stereotactic reference space and the single-subject contrast images were entered into a second-level random effects analysis for each of the contrasts. One-sample t-tests were employed for the group analyses across the contrast images of all subjects that indicate whether observed differences between conditions were significantly different from zero. The t-values were subsequently transformed into z-scores. To correct for false-positive results, in a first step, an initial voxelwise z-threshold was set to  $z = 2.576$  ( $p < .005$ , uncorrected). 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  $p = .05$ , that is, the reported activations were significantly activated at  $p < .05$ .

## Results

### Behavioral data, first pilot session

As described in the **Methods**, boundaries were determined as frames with a frame value two standard deviations above average bin value. As explained above (Fig. 2), frame value is an index of the sum of all participants' responses (aggregated responses) within the next 25-frames bin. An example of a typical movie can be seen in Table 1 and Fig. 3.

### Differences and similarities between conditions

There was a significant effect of the factor CONDITION in any ANOVA. Moreover, the Fisher z-transformed correlations for response rates in action with action point cloud and tai chi with tai chi point cloud movies were higher than all other comparisons. In the 2nd pilot session the Fisher-z transformed correlation averages, calculated in analogy to the analysis for the 1st pilot session, for the naturalistic conditions and their respective counterparts were again higher than all other comparisons. However, the correlation of the time series of frame values between corresponding movies of different conditions reached significance ( $p < 0.05$ ) between action and tai chi movies in seven out of 12 comparisons. Thus, conditions seem comparable, but differences especially between action and tai chi movies exist.

### Reliability of detection judgments

We used the second pilot session to analyze whether participants were consistent in their detection, or whether there was intra-personal variability. In the time-window analysis on the single subject level, a mean of 24% of responses in the second pilot session fell within the participants' individual response window. A standard distribution would have 10% of responses fall within this time frame. This result was found in all conditions with 28% of responses in the action, 21% of responses in the tai chi, 21% of responses in the action point cloud and 26% of responses in the tai chi point cloud fulfilling this criterion of proximity to the response in the first session. Intra-individual detection behavior was thus rather consistent.

This intra-individual reliability was also reflected in group-averaged statistics. The correlation of the time series of frame values was substantial ( $z'_r > 0.5$ ) for all conditions, indicating sound re-test-reliability. Action condition reached a re-test reliability of  $z'_r = 0.7$ , tai chi condition reached  $z'_r = 0.5$ , action point cloud condition reached  $z'_r = 0.7$  and tai chi point cloud test-re-test reliability was  $z'_r = 0.5$ .

**Table 1**  
Bin-analysis measures for one example movie (the making coffee action movie), the time-matched tai chi movie as well as the corresponding action point cloud and tai chi point cloud movies. Labels and abbreviations: Movie name = content related movie name; Condition = stimulus condition; Act = action movie; Tai = tai chi movie; Act cloud = action point cloud movie; Tai cloud = tai chi point cloud movie; N° frames = number of 40 ms frames; Responses = number of responses for the movie pooled across all participants; Initial frames = number of frames that initiated a bin that contained a response; Bin min = minimum of responses per bin; Bin max = maximum of responses per bin; Bin mean = sum of values of initial frames divided by number of initial frames; Bin STD = standard deviation of Bin mean; Boundary = number of bins that contain a number of responses > Bin mean + 2 Bin STD; Dist (s) = average time (seconds) between two boundaries.

Movie name	Condition	N° frames	Responses	Initial frames	Bin min	Bin max	Bin mean	Bin STD	Boundary	Dist (s)
Making coffee	Act	1117	186	999	1	12	4.59	2.63	8	5.59
	Tai	1117	104	950	1	7	2.73	1.41	11	4.06
	Act cloud	1117	122	906	1	12	3.33	2.33	12	3.72
	Tai cloud	1117	92	872	1	8	2.63	1.64	10	4.47

*Behavioral data, fMRI session*

In the target interaction contrast and two control calculations, the onset vectors for the hemodynamic modeling were not derived from group statistics. However, to guarantee comparability of the results from the pilot study and the fMRI experiment, we calculated an analog analysis.

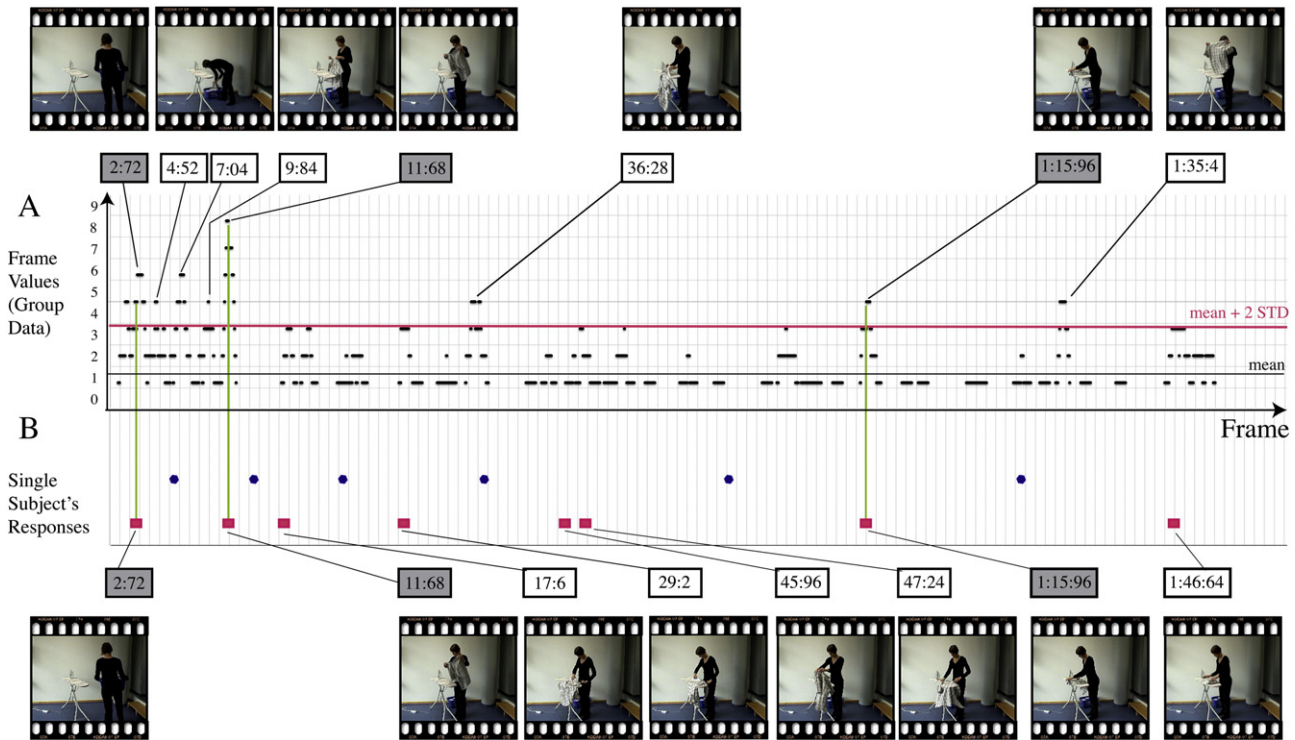
As previously mentioned, the first univariate ANOVA between first pilot and fMRI session showed a significant main effect of CONDITION ( $F_{3,1} = 4.56$ ;  $p < 0.05$ ). There was no main effect of SESSION ( $F_{1,1} = 4.4$ ;  $p > 0.05$ ) and no interaction between the two factors ( $F_{3,1} = 0.15$ ;  $p > 0.05$ ). The univariate ANOVA between the second pilot session and fMRI session yielded the same results, a significant main effect of CONDITION ( $F_{3,7} = 4.91$ ,  $p < 0.05$ ), no main effect of SESSION ( $F_{1,7} = 4.88$ ;  $p > 0.05$ ) and no interaction ( $F_{3,7} = 0.19$ ;  $p > 0.05$ ) (Fig. 4).

In addition to the analysis of individually set boundaries, aggregated bin value-based boundaries were also consistent between the first pilot session and the fMRI session. In a correlation analysis of frame values, a measure carrying information on position in time and participant agreement on boundaries, six out of 12 action movies correlated

significantly ( $p < 0.05$ ). The same significant correlation was found for three out of 12 tai chi movies, 10 out of 12 action point cloud movies and eight out of 12 tai chi point cloud movies.

*Synopsis of the behavioral data analyses*

Subjects reliably detected the same boundaries on re-test measures. Group-aggregated boundaries also showed good re-test reliability. Conditions differed in response rates, especially comparing the action and action point cloud with the tai chi and tai chi point cloud conditions. These differences were stable and no condition seemed to be subject to learning between sessions. The similarity between action and action point cloud and tai chi and tai chi point cloud could point towards the importance of motion features in detection, while the differences between action and tai chi indicates that boundary detection in action vs. tai chi does entail differences. Thus, motion seems important for boundary detection. Boundary detection in action seems to be moreover influenced by something else, apart from motion, accounting for the differences between action and tai chi boundary judgments.



**Fig. 3.** Example of response recording. A: Boundaries derived from group-aggregated bin-analysis. If frame values surpassed the mean frame value by two standard deviations, this frame was designated boundary. Times and images refer to the actual boundaries of the ironing-movie. B: A single subject's responses during the ironing-movie and the according images of the indicated boundaries. Blue dots refer to non-boundaries, placed at the mean distance between two boundaries that were more than 4 s apart. Green lines mark the single subject's responses that coincided with group-aggregated boundaries.



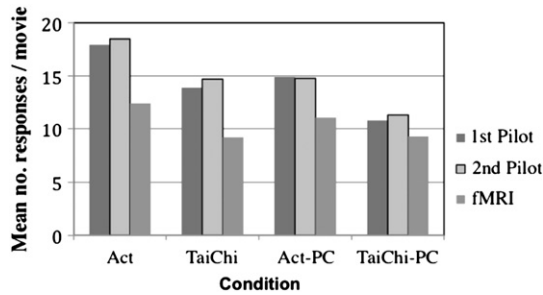


Fig. 4. The responses per condition averaged over participants and movies, displayed for each condition in each session.

Table 2

Target Action boundary interaction contrast. Anatomical specification, Talairach coordinates (x, y, z) and maximal Z-scores of significantly activated voxels for ((ACT-BO > TAI-BO) > (ACT-NOBO > TAI-NOBO)).

Localization	Talairach coordinates			z-Values, local maxima
	x	y	z	
Superior frontal sulcus	-14	36	46	3.33
	-28	16	48	4.04
Angular gyrus	-32	-72	36	4.81
Parahippocampal gyrus	26	-44	0	3.92
	-28	-46	-6	4.82

Functional MRI data analysis

In order to tap boundary-detection related activity specific for action (as in contrast to tai chi movements), we calculated the interaction contrast ((ACT-BO > TAI-BO) > (ACT-NOBO > TAI-NOBO)). We found significant activity in the left superior frontal sulcus (SFS), the left angular gyrus (AG) and bilaterally in the parahippocampal gyri (PHG) (Table 2, Fig. 5).

In order to address the question as to how MT area's response to boundary detection is action-specific or disproportionately strong for actions, we set up a ROI analysis based on MT-coordinates from Zacks et al. (2001) to avoid double dipping (Kriegeskorte et al., 2009). This analysis yielded a main effect of PART (boundary, non-boundary), but neither a significant effect of CONDITION (action, tai chi), nor a significant interaction PART by CONDITION. Further analysis revealed that the main effect of PART was driven by a higher hemodynamic response in boundaries as compared to non-boundaries (Table 3).

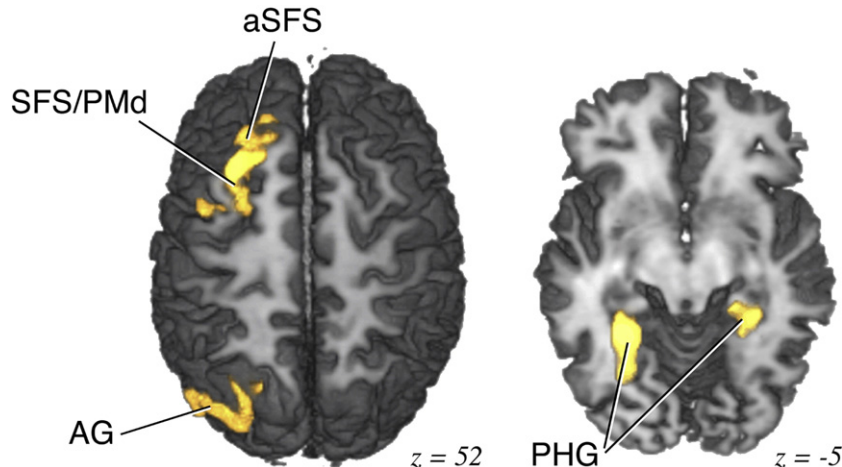


Fig. 5. Activity in the target contrast ((ACT-BO > TAI-BO) > (ACT-NOBO > TAI-NOBO)), corrected for multiple comparisons at Z < 0.05. aSFS: anterior superior frontal sulcus; SFS: superior frontal sulcus; PMd: dorsal premotor cortex; AG: angular gyrus; PHG: parahippocampal gyrus.

Table 3

F-values and p-values from the area MT ROI-analysis. Talairach coordinates from Zacks et al., 2001.

	Left MT		Right MT	
	(x = -34; y = -73; z = 7)		(x = 34; y = -70; z = 6)	
CONDITION main: (Action vs tai chi)	F <sub>1,15</sub> = .447	p = 0.513	F <sub>1,15</sub> = 1.49	p = 0.240
PART main:	F <sub>1,15</sub> = 9.382	p = 0.007	F <sub>1,15</sub> = 11.028	p = 0.004
Boundary vs. non-boundary				
Interaction:	F <sub>1,15</sub> = 0.299	p = 0.529	F <sub>1,15</sub> = 0.313	p = 0.584
CONDITION * PART				

We finally calculated the conjunction ((ACT > null) ∩ (TAI > null) ∩ (ACT-Cloud > null) ∩ (TAI-Cloud > null)) in order to test whether we could identify significant activity in area MT using a whole-brain analysis as well. Null events were displays of a gray screen dispersed between the movies. This contrast yielded in fact significant activity in area MT (Fig. 6). Further active spots were localized in the dorsal premotor cortex/frontal eye field (PMd/FEF), the anterior intraparietal sulcus (IPS), the horizontal segment of the IPS, the left descending segment of the IPS and in the anterior fusiform gyrus (Table 4).

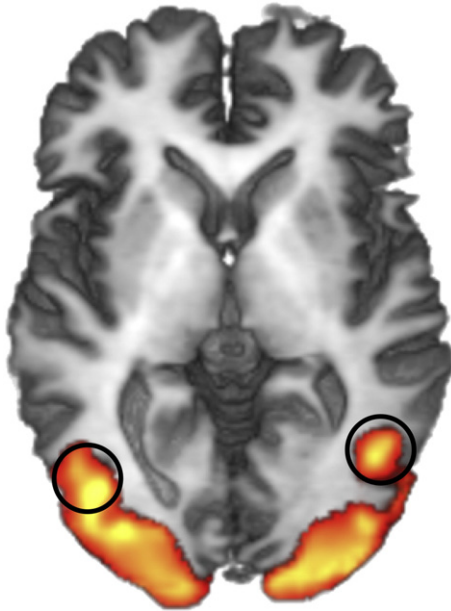
To summarize the fMRI results, the targeted interaction contrast yielded activation in a clearly circumscribed network. Moreover, the presence of enhanced area MT activity in all conditions (conjunction contrast) as well as the results from the ROI analysis create a coherent picture: activity in area MT was specifically elevated at segmentation boundaries, but it was not particularly enhanced in the action as compared to the tai chi condition.

Discussion

The current fMRI study investigated whether the detection of meaningful steps in actions relies on more than the detection of change in motion, i.e., dynamic information. In particular, we tested whether motion area (MT) enhancement at action boundaries is specific to action segmentation as compared to segmentation of movements, and whether selection- and long-term-memory related brain areas are particularly elevated at action boundaries, as opposed to movement boundaries.

The behavioral results revealed that boundary judgments were a reliable measure, and comparable between different action, tai chi and point cloud motion conditions. Inferential statistics showed that the segmentation behavior was similar between the action and the action-point-cloud movies. The latter condition contained the same dynamic features as the former, the action condition. The same similarity of segmentation behavior was found for tai chi movies and the

## Area MT



**Fig. 6.** Activity in the control conjunction  $[(ACT > null) \cap (TAI > null) \cap (ACT-Cloud > null) \cap (TAI-Cloud > null)]$ , corrected for multiple comparisons at  $z = 2.56$  and  $z$ -map optically thresholded at  $z = 3.0$ .

tai chi point cloud movies. These two results indicate that the analysis of motion was of great importance in both action as well as the tai chi condition. Moreover, the boundary detection in the action movies differed significantly from boundary detection in the tai chi movies, with regard to response rates, suggesting the existence of differences between these conditions. The fact that dynamic features seem to be a common aspect to boundary detection falls in place with the established activity of area MT in all conditions if compared against the rest. However, the existence of a circumscribed network specifically for action boundary detection, as revealed in the interaction contrast, points out differences between segmentation in action as compared to movements. This network comprised of the left superior frontal sulcus including aSFS and pSFS/PMd, the left angular gyrus (AG) and both parahippocampal gyri (PHG).

Of the frontal areas, the mid-dorsolateral prefrontal cortex (BA 46, 9/46, 9) has access to the hippocampal memory system, as it gives rise to fibers that travel as part of the cingulum bundle via retrosplenium to the hippocampal region (see Petrides, 2005, for review). The aSFS and PMd are also directly connected to the angular gyrus, which in turn is connected to the parahippocampal gyrus (Makris et al.,

**Table 4**

Control conjunction. Anatomical specification, Talairach coordinates ( $x, y, z$ ) and maximal  $Z$ -scores of significantly activated voxels for  $[(ACT > null) \cap (TAI > null) \cap (ACT-Cloud > null) \cap (TAI-Cloud > null)]$ .

Localization	Talairach coordinates			z-Values, local maxima
	x	y	z	
Area MT	42	-63	6	4.53
	-38	-73	6	5.21
Dorsal premotor cortex	24	-3	48	3.34
	-22	-5	48	3.62
Intraparietal sulcus	36	-37	48	3.79
	18	-59	56	4.57
	24	-71	46	4.22
	-26	-49	54	5.13
	-22	-67	56	4.04
Fusiform gyrus	42	-15	-36	3.67
Medial occipital gyrus	32	-83	12	5.56
	-28	-95	16	5.31

2004; Olson and Berryhill, 2009; Rushworth et al., 2006; Schmahmann et al., 2007). Thus, input from the dorsolateral prefrontal cortex reaches the posterior inferior parietal lobule. This dorsal fronto-parietal connection, is realized via SLF I and II (Makris et al., 2004). The SLF II network has been suggested to regulate attention within different parts of space, whereas the SLF I network is related to spatial motor control. Finally, dense connections exist between the parahippocampal gyrus and the posterior inferior parietal lobule (Cavada and Goldmanrakis, 1989; Seltzer and Pandya, 1984; Suzuki and Amaral, 1994).

### Predictive attentional control

Shifts of visual or mnemonic attention are reported to be associated with transient control signals in a fronto-parietal network (for references, see Tamber-Rosenau et al., 2011). Both the middle frontal gyrus and SFS are suggested to play a part in shifts of perceptual and mnemonic attention, but their respective specific functional role remains undescribed so far. Perceptual and mnemonic attention is relevant at event boundaries: Kurby and Zacks (2008) have argued that at event boundaries, a prediction error occurs because the currently valid event model has come to its end and provides no further cues for prediction. These transient increases in prediction errors leads to an updating of the models. Event boundaries can be thought of as nodes where several upcoming action steps are possible, hence conflict increases when a boundary is approached. A predictive account of perception proposes that competing models are retrieved and held ready before a boundary is reached. At the time when subjects press the response button to indicate that a new action step has begun, conflict is already resolved or is vanishing quickly. Against this backdrop, it is particularly interesting that the left anterior SFS has been identified in a study that investigated post-error behavioral adjustments (King et al., 2010). The left aSFS was enhanced in activity when participants showed a post-error reduction of interference. As the authors also found an increase in task-relevant perceptual areas (here: fusiform face area), the post-error interference reduction was suggested to reflect an enhancement of relevant representations rather than a suppression of irrelevant ones. Transferring these functional suggestions to our findings, aSFS activity increase at action boundaries reflects a top-down attentional enhancement of those parts of the *action scenario* that are becoming relevant at the very moment. A slightly different interpretation holds that aSFS exerts a top-down enhancement of the parts of the *action script* that are currently held in working memory (Tamber-Rosenau et al., 2011). Accordingly, aSFS may enhance attention towards (the memory or the percept of) a specific location or item in space. For instance, after the actress has put on her shoes and slipped into her coat, the keyboard next to the door may become relevant because the overarching action – preparing to leave the apartment – will very probably encompass the actress' grabbing of the keys. This item or location is prioritized on the basis of action expertise, i.e. knowledge retrieved from long-term memory, and this retrieval is probably reflected by concurrent activity in the parahippocampal gyrus (as will be discussed below). In close keeping with the interpretation put forward by King and co-workers, aSFS may also help to hold the relevant representations (e.g. keyboard) online and thereby enable fast and efficient stimulus processing (e.g. to perceive the actress grabbing the keys).

Bledowski and co-workers (Bledowski et al., 2009) noted that a core feature of working memory is that some of the maintained items become transiently more important than others. The authors report left aSFS and posterior cingulate cortex for the selection operation that retrieves the most relevant item from memory; in contrast, bilateral dorsal premotor cortex (PMd; caudal SFS) and posterior parietal cortex were found to increase activity either for an updating operation that changes the focus of attention onto the retrieved item, or for both updating and selecting. In our study, both the selection of the



relevant item as well as updating attention to the new item may be reflected in the SFS activity, as it extended from the crossing with superior precentral sulcus (PMd) to anterior prefrontal cortex (aSFS, corresponding to  $y = 40$  of the Talairach space). The notion expressed by Bledowski and coworkers of an updating process subserved by PMd is corroborated by a recent fMRI study where left PMd was found to be the only area selectively activated by predictive bridging of occluded actions (Stadler et al., 2011). It is important to see these findings in the context of other research in human (and animal) findings on PMd, showing that PMd is engaged in preparatory and anticipatory processes, particularly in the spatial domain (cf. Rushworth et al., 2003; Schubotz, 2007). In the present experiment as well as in Stadler's study, videos showed an actress moving around in a room while performing the action, and transport of objects and non-distal body movements were prominent cues for segmentation (in our study) and prediction (in Stadler's study). This argues in favor of the relevance of *spatial* attention/processing in our study, a point further strengthened by angular gyrus and hippocampal activation, as will be discussed in the following section.

### Adaptation of spatial attention

In addition to aSFS, we found the left posterior angular gyrus and the parahippocampal gyrus bilaterally to be specifically enhanced by boundary detection in actions. Rushworth and colleagues (Rushworth et al., 2006) showed that the posterior angular gyrus can be distinguished from adjacent parietal fields by its distinct connection (via inferior longitudinal fascicle) with the parahippocampal gyrus. This study strongly corroborates a functional connection between angular gyrus and parahippocampal gyrus, and this connection also provides a hint for the specific functional meaning of angular gyrus activity in our study. In particular, it suggests that attentional adaptations to the perceived changes are informed by long-term memory.

Human angular gyrus is associated with the redirecting of saccades and the reorientation of attention (Kincade et al., 2005; Mort et al., 2003; Rushworth et al., 2001; Thiel et al., 2004, 2005). Importantly, following a functional segregation in parietal fields proposed by Corbetta and Shulman (2002), angular gyrus activity in our study points to a top-down (frontally) mediated attentional modulation rather than to the detection of salient or unexpected stimuli. In addition, angular gyrus may also receive relevant input from area MT (cf. Speer et al., 2003; Zacks et al., 2001). In particular, these MT signals can relate to a change in movement pattern, for instance the release of an object (e.g. the coat) or the turning of the hands or the body towards the next object (e.g. the keyboard). While our findings suggest that motion signals can serve as a boundary detection trigger in moving stimuli, not only in observed action, BOLD effects suggest that only in the action condition, motion signals are mediated to the dorsolateral prefrontal cortex, triggering the selection of the next relevant item in the scene and the consequential updating operation that changes the focus of attention onto the retrieved item (cf. Bledowski et al., 2009). It is plausible to assume that aSFS exerted predictive attentional control onto angular gyrus based on the currently valid action model. Early selection of informative changes in the stimulus and the adaptation of predictive perception can be driven by a set of probabilistically weighted forward models (Summerfield et al., 2006). This operation has to be informed by long-term memory retrieval (cf. Schacter and Addis, 2009), and we consider parahippocampal activation to reflect this latter function.

### Relevant action models supplied by long term memory

The medial temporal lobe, including the hippocampal formation, is established for long-term episodic and semantic memory (Kim

and Baxter, 2001). The hippocampus proper is a core structure for encoding spatial environment and relations in long-term memory (Rosenbaum et al., 2004). However, mental navigation in old environments was found to rather rely on the parahippocampal gyrus (Maguire et al., 1997; Rosenbaum et al., 2004; Shelton and Gabrieli, 2002). Thus, parahippocampal gyrus seems to figure as a long-term memory for topographical orientation in highly familiar environments (Aguirre et al., 1996; Bohbot et al., 1998; Maguire et al., 1997; Rosenbaum et al., 2004).

The parahippocampal activation included the parahippocampal place area or PPA, an area suggested to code for the layout of local space (Epstein and Kanwisher, 1998). Recent findings extend PPA function to automatically re-activate visual context to mediate successful episodic memory retrieval (Hayes et al., 2007). According to this suggestion, in the current study, the shift of attention towards aspects in the scene that become important at event boundaries (e.g. the keyboard) may reflect an automatic reinstatement of visual context (e.g. the vestibule of one's own apartment) to enhance episodic retrieval (e.g. actions and experiences we had at this or prototypically similar locations). Rosenbaum et al. (2000; 2001) proposed that spatial memories in the parahippocampal gyrus might be schematic cognitive maps that retain the gist of the environment without rich contextual details that allow one to re-experience it. Accordingly, parahippocampal gyrus in our study provides "spatial semantics" or environmental gist that associates a currently relevant stimulus part (e.g. the keyboard) with memories of actions that have been frequently experienced in such a context (e.g. grabbing or storing away keys). In the context of action observation, this memory retrieval builds the basis of forward models, and hence expectations and predictions, about upcoming action steps (e.g. leaving the apartment).

The "spatial gist" of parahippocampal function is in line with that of co-activated angular gyrus and left SFS/PMd, as mentioned before. However, a more *action*-related account of parahippocampal activity is plausible as well. The activation of parahippocampal gyrus in our study extended from  $y = -32$  to  $y = -52$  (right) and  $y = -70$  (left) in anterior–posterior direction. Sommer and co-workers (Sommer et al., 2005) found that a comparable segment of the parahippocampal gyrus was associated equally with the retrieval of objects based on location cues and the retrieval of locations based on object cues. Moreover, several neuroimaging studies have put forward a more general role of the parahippocampal gyrus or cortex in associative memories, as it is active during encoding and retrieval of associative information (Düzel et al., 2003; Eldridge et al., 2000; Henke et al., 1997; 1999; 2003; Krause et al., 1999; Rombouts et al., 1997). Hence this area engages during both scene processing and episodic retrieval because both rely on associations (Bar et al., 2008), i.e., spatial as well as non-spatial ones (Aminoff et al., 2007; however, cf. Epstein and Ward, 2010).

As we found this area to show an increased hemodynamic response at event boundaries, i.e., associative nodes in the actions, the associative nature of actions rather than enhanced spatial memory retrieval may have triggered parahippocampal gyrus activity. Related to this view, probabilistic learning studies indicate that hippocampus proper, but also parahippocampal gyrus (Amso et al., 2005; Turke-Brown et al., 2010) is sensitive to the stochastic or probabilistic structure of observed events. More specifically, its activity level announces the expected informational content of an event before it occurs (Strange et al., 2005). Parahippocampal gyrus activation increase at event boundaries hence may also point to a change in the probabilistic structure that we have outlined in the beginning.

### Concluding remarks

Actions and movements provide a continuous stream of input to observers. In this study we found that individuals are very consistent

when judging upon meaningful segments or steps, albeit high inter-individual variability. Imaging results indicate that motion serves as a bottom-up cue for boundary detection in different types of dynamic stimuli, but that only for actions, this information also triggers a long-term memory search that guides further expectations. The emerging picture is that motion feature signals are conveyed to the prefrontal cortex that exerts a top-down modulation of attention on informative features in posterior parietal lobule. This attentional adaptation to a detected change in the action makes use of long-term memory place-action associations encoded in the hippocampal formation. Future studies have to address whether this mechanism depends on expertise, i.e. the existence of internal models in long-term memory. In this case, tai chi experts should reveal a similar network when judging tai chi boundaries.

## Acknowledgments

The authors wish to cordially thank David Horbank, Marcel Mücke, Stephan Liebig, Sarah Schröder, Christiane Ahlheim and Andreas Johnen for their valuable contributions.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.neuroimage.2012.04.008>.

## References

- Aguirre, G.K., Detre, J.A., Alsup, D.C., D'Esposito, M., 1996. The parahippocampus subserves topographical learning in man. *Cereb. Cortex* 6 (6), 823–829.
- Allain, P., Le Gall, D., Etcharry-Bouyx, F., Aubin, G., Emile, J., 1999. Mental representation of knowledge following frontal-lobe lesion: dissociations on tasks using scripts. *J. Clin. Exp. Neuropsychol.* 21 (5), 643–665.
- Aminoff, E., Gronau, N., Bar, M., 2007. The parahippocampal cortex mediates spatial and nonspatial associations. *Cereb. Cortex* 17, 1493–1503.
- Amso, D., Davidson, M.C., Johnson, S.P., Glover, G., Casey, B.J., 2005. Contributions of the hippocampus and the striatum to simple association and frequency-based learning. *NeuroImage* 27 (2), 291–298.
- Baird, J.A., Baldwin, D.A., 2001. Making sense of human behavior: action parsing and intentional inference. In: Malle, B.F., Moses, L.J., Baldwin, D.A. (Eds.), *Intentions and Intentionality: Foundations of Social Cognition*. The MIT Press, Cambridge, MA, US, pp. 193–206.
- Bar, M., Aminoff, E., Schacter, D.L., 2008. Scenes unseen: the parahippocampal cortex intrinsically subserves contextual associations, not scenes or places per se. *J. Neurosci.* 28 (34), 8539–8544.
- Bledowski, C., Rahm, B., Rowe, J.B., 2009. What “works” in working memory? Separate systems for selection and updating of critical information. *J. Neurosci.* 28 (43), 13735–13741.
- Bohbot, V.D., Kalina, M., Stepankova, K., Spackova, N., Petrides, M., Nadel, L., 1998. Spatial memory deficits in patients with lesions to the right hippocampus and to the right parahippocampal cortex. *Neuropsychologia* 36, 1217–1238.
- Bubic, A., von Cramon, D.Y., Schubotz, R.I., 2009. Prediction, cognition and the brain. *Front. Hum. Neurosci.*
- Cavada, C., Goldman-Rakic, P.S., 1989. Posterior parietal cortex in rhesus-monkey. 2. Evidence for segregated corticocortical networks linking sensory and limbic areas with the frontal-lobe. *J. Comp. Neurol.* 287 (4), 422–445.
- Corbetta, M., Shulman, G.L., 2002. Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* 3 (3), 201–215.
- Christoffersen, K., Woods, D.D., Blike, G.T., 2007. Discovering the events expert practitioners extract from dynamic data streams: the mUMP technique. *Cogn. Technol. Work* 9, 81–98.
- Decety, J., Grezes, J., 1999. Neural mechanisms subserving the perception of human actions. *TICS* 3, 172–178.
- Düzel, E., Habib, R., Rotte, M., Guderian, S., Tulving, E., Heinze, H.-J., 2003. Human hippocampal and parahippocampal activity during visual associative recognition memory for spatial and nonspatial stimulus configurations. *J. Neurosci.* 23 (28), 9439–9444.
- Eldridge, L.L., Knowlton, B.J., Furmanski, C.S., Bookheimer, S.Y., Engel, S.A., 2000. Remembering episodes: a selective role for the hippocampus during retrieval. *Nat. Neurosci.* 3 (11), 1149–1152.
- Epstein, R.A., Kanwisher, N., 1998. A cortical representation of the local visual environment. *Nature* 392 (6676), 598–601.
- Epstein, R.A., Ward, E.J., 2010. How reliable are visual context effects in the parahippocampal place area? *Cereb. Cortex* 20, 294–303.
- Frank, M.J., Claus, E.D., 2006. Anatomy of a decision: striato-orbitofrontal interactions in reinforcement learning, decision making, and reversal. *Psychol. Rev.* 113 (2), 300–326.
- Friston, K.J., Mattout, J., Kilner, J.M., 2011. Action understanding and active inference. *Biol. Cybern.* 104 (1–2), 137–160.
- Ghahremani, D.G., Monterosso, J., Jentsch, J.D., Bilder, R.M., Poldrack, R.A., 2010. Neural components underlying behavioral flexibility in human reversal learning. *Cereb. Cortex* 20 (8), 1843–1852.
- Grush, R., 2004. The emulation theory of representation: motor control, imagery, and perception. *Behav. Brain Sci.* 27, 377–396.
- Hayes, S.M., Nadel, L., Ryan, L., Al, H.E.T., 2007. The effect of scene context on episodic object recognition: parahippocampal cortex mediates memory encoding and retrieval success. *Hippocampus* 17, 873–889.
- Henke, K., Buck, A., Weber, B., Wieser, H.G., 1997. Human hippocampus establishes associations in memory. *Hippocampus* 7 (3), 249–256.
- Henke, K., Weber, B., Kneifel, S., Wieser, H.G., Buck, A., 1999. Human hippocampus associates information in memory. *PNAS* 96 (10), 5884–5889.
- Henke, K., Mondadori, C.R., Treyer, V., Nitsch, R.M., Buck, A., Hock, C., 2003. Nonconscious formation and reactivation of semantic associations by way of the medial temporal lobe. *Neuropsychologia* 41 (8), 863–876.
- Kim, J.J., Baxter, M.G., 2001. Multiple brain-memory systems: the whole does not equal the sum of its parts. *Trends Neurosci.* 24 (6), 324–330.
- Kincade, J.M., Abrams, R.A., Astafiev, S.V., Shulman, G.L., Corbetta, M., 2005. An event-related functional magnetic resonance imaging study of voluntary and stimulus-driven orienting of attention. *J. Neurosci.* 25, 4593–4604.
- King, J.A., Korb, F.M., von Cramon, D.Y., Ullsperger, M., 2010. Post-error behavioral adjustments are facilitated by activation and suppression of task-relevant and task-irrelevant information processing. *J. Neurosci.* 30 (38), 12759–12769.
- Koch, I., Hoffmann, J., 2000. Patterns, chunks, and hierarchies in serial reaction-time tasks. *Psychol. Res.* 63 (1), 22–35.
- Körding, K.P., Wolpert, D.M., 2006. Bayesian decision theory in sensorimotor control. *TICS* 10, 319–326.
- Krause, B.J., Horwitz, B., Taylor, J.G., Schmidt, D., Mottaghy, F.M., Herzog, H., Halsband, U., Müller-Gärtner, H.-W., 1999. Network analysis in episodic encoding and retrieval of word-pair associates: a PET study. *Eur. J. Neurosci.* 11 (9), 3293–3301.
- Kriegeskorte, N., Simmons, W.K., Bellgowan, P.S., Baker, C.I., 2009. Circular analysis in systems neuroscience: the dangers of double dipping. *Nat. Neurosci.* 12 (5), 535–540.
- Kuhl, B.A., Dudukovic, N.M., Kahn, I., Wagner, A.D., 2007. Decreased demands on cognitive control reveal the neural processing benefits of forgetting. *Nat. Neurosci.* 10 (7), 908–914.
- Kurby, C.A., Zacks, J.M., 2008. Segmentation in the perception and memory of events. *TICS* 12 (2), 72–79.
- Lohmann, G., Mueller, K., Bosch, V., Mentzel, H., Hessler, S., et al., 2001. Lipsia – a new software system for the evaluation of functional magnetic resonance images of the human brain. *Comput. Med. Imaging Graph.* 25 (6), 449–457.
- Maguire, E.A., Frackowiak, R.S., Frith, C.D., 1997. Recalling routes around London: activation of the right hippocampus in taxi drivers. *J. Neurosci.* 17, 7103–7110.
- Makris, N., Kennedy, D.N., McInerney, S., Sorensen, A.G., Wang, R., Caviness, J.V.S., Pandey, D.N., 2004. Segmentation of subcomponents within the superior longitudinal fascicle in humans: a quantitative, in vivo, DR-MRI study. *Cereb. Cortex* 15, 854–869.
- Miller, E.K., Cohen, J.D., 2001. An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* 24, 167–202.
- Mort, D.J.P.R., Mannan, S.K., Hodgson, T.L., Anderson, E., Quest, R., McRobbie, D., McBride, A., Husain, M., Kennard, C., 2003. Differential cortical activation during voluntary and reflexive saccades in man. *NeuroImage* 18, 231–246.
- Neal, A., Kilner, J.M., 2010. What is simulated in the action observation network when we observe actions? *Eur. J. Neurosci.* 32 (10), 1765–1770.
- Newton, D., 1973. Attribution and the unit of perception of ongoing behavior. *J. Pers. Soc. Psychol.* 28, 28–38.
- Newton, D., Engquist, G., 1976. The perceptual organization of ongoing behavior. *J. Exp. Soc. Psychol.* 12 (5), 436–450.
- Norris, D.G., 2000. Reduced power multislice MDEFT imaging. *J. Magn. Reson.* 11, 445–451.
- Olson, I.R., Berryhill, M., 2009. Some surprising findings on the involvement of the parietal lobe in human memory. *Neurobiol. Learn. Mem.* 91 (2), 155–165.
- O'Reilly, R.C., Noelle, D.C., Braver, T.S., Cohen, J.D., 2002. Prefrontal cortex and dynamic categorization tasks: representational organization and neuromodulatory control. *Cereb. Cortex* 12 (3), 246–257.
- Petrides, M., 2005. Lateral prefrontal cortex: architectonic and functional organization. *Philos. Trans. R. Soc. Lond. B* 360, 781–795.
- Restle, F., Brown, E.R., 1970. Serial pattern learning. *J. Exp. Psychol.* 83 (1), 120–125.
- Ridderinkhof, K.R., Ullsperger, M., Crone, E.A., Nieuwenhuis, S., 2004a. The role of the medial frontal cortex in cognitive control. *Science* 306 (5695), 443.
- Ridderinkhof, K.R., van den Wildenberg, W.P.M., Segalowitz, S.J., Carter, C.S., 2004b. Neurocognitive mechanisms of cognitive control: the role of prefrontal cortex in action selection, response inhibition, performance monitoring, and reward-based learning. *Brain Cogn.* 56 (2), 129–140.
- Rizzolatti, G., Fogassi, L., Gallese, V., 2001. Neurophysiological mechanisms underlying the understanding and imitation of action. *Nat. Rev. Neurosci.* 2 (9), 661–670.
- Rombouts, S.A.R.B., Machielsen, W.C., Witter, M.P., Barkhof, F., Lindeboom, J., Scheltens, P., 1997. Visual association encoding activates the medial temporal lobe: a functional magnetic resonance imaging study. *Hippocampus* 7 (6), 594–601.
- Rosenbaum, D.A., Kenny, S.B., Derr, M.A., 1983. Hierarchical control of rapid movement sequences. *J. Exp. Psychol. Hum. Percept. Perform.* 9 (1), 86–102.
- Rosenbaum, R.S., Priselac, S., Köhler, S., Black, S.E., Gao, F., Nadel, L., et al., 2000. Remote spatial memory in an amnesic person with extensive bilateral hippocampal lesions. *Nat. Neurosci.* 3, 1044–1048.

- Rosenbaum, R.S., Winocur, G., Moscovitch, M., 2001. New views on old memories: re-evaluating the role of the hippocampal complex. *Behav. Brain Res.* 127, 183–197.
- Rosenbaum, R.S., Ziegler, M., Winocur, G., Grady, C.L., Moscovitch, M., 2004. "I have often walked down this street before": fMRI studies on the hippocampus and other structures during mental navigation of an old environment. *Hippocampus* 14 (7), 826–835.
- Rushworth, M.F.S., Ellison, A., Walsh, V., 2001. Complementary localization and lateralization of orienting and motor attention. *Nat. Neurosci.* 4, 656–661.
- Rushworth, M.F.S., Johansen-Berg, H., Goebel, S.M., Devlin, J.T., 2003. The left parietal and premotor cortices: motor attention and selection. *Neuroimage* 20, S89–S100.
- Rushworth, M.F.S., Behrens, T.E.J., Johansen-Berg, H., 2006. Connection patterns distinguish 3 regions of human parietal cortex. *Cereb. Cortex* 16, 1418–1430.
- Schacter, D.L., Addis, D.R., 2009. On the nature of medial temporal lobe contributions to the constructive simulation of future events. *Philos T Roy Soc B* 364 (521), 1245–1253.
- Schiffer, A.-M., Schubotz, R.I., 2011. Caudate nucleus signals for breaches of expectation in a movement observation paradigm. *Front. Hum. Neurosci.* 5.
- Schiffer, A.M., Ahlheim, C., Ulrichs, K., Schubotz, R.I., 2012. Neural changes when actions change: adaptation of strong and weak expectations. *Hum. Brain Mapp.*, <http://dx.doi.org/10.1002/hbm.22023> (Electronic publication ahead of print, Mar 16).
- Schmahmann, J.D., Pandya, D.N., Wang, R., Dai, G., D'Arceuil, H.E., de Crespigny, A.J., Weeden, J.V., 2007. Association fibre pathways of the brain: parallel observations from diffusion spectrum imaging and autoradiography. *Brain* 130, 630–653.
- Schubotz, R.I., 2007. Prediction of external events with our motor system: towards a new framework. *TICS* 11 (5), 211–218.
- Schütz-Bosbach, S., Prinz, W., 2007. Prospective coding in event representation. *Cogn. Process.* 8 (2), 93–102.
- Seltzer, B., Pandya, D.N., 1984. Further observations on parieto-temporal connections in the rhesus monkey. *Exp. Brain Res.* 55 (2), 301–312.
- Shelton, A.L., Gabrieli, J.D.E., 2002. Neural correlates of encoding space from route and survey perspectives. *J. Neurosci.* 22, 2711–2717.
- Sirigu, A., Zalla, T., Pillon, B., Grafman, J., Agid, Y., Dubois, B., 1995. Selective impairments in managerial knowledge following pre-frontal cortex damage. *Cortex* 31 (2), 301–316.
- Sommer, T., Rose, M., Gläscher, J., Wolbers, T., Büchel, C., 2005. Dissociable contributions within the medial temporal lobe to encoding of object-location associations. *Learn. Mem.* 343–351.
- Speer, N.K., Swallow, K.M., Zacks, J.M., 2003. Activation of human motion processing areas during event perception. *Cogn. Affect. Behav. Neurosci.* 3 (4), 335–345.
- Stadler, W., Schubotz, R.I., von Cramon, D.Y., Springer, A., Graf, M., Prinz, W., 2011. Predicting and memorizing observed action: differential premotor cortex involvement. *HBM* 32 (5), 677–687.
- Strange, B.A., Duggins, A., Penny, W., Dolan, R.J., Friston, K.J., 2005. Information theory, novelty and hippocampal responses: unpredicted or unpredictable? *Neural Netw.* 18 (3), 225–230.
- Summerfield, C., Egner, T., Greene, M., Koehlin, E., Mangels, J., Hirsch, J., 2006. Predictive codes for forthcoming perception in the frontal cortex. *Science* 314 (5803), 1311–1314.
- Suzuki, W.A., Amaral, D.G., 1994. Perirhinal and parahippocampal cortices of the macaque monkey – cortical afferents. *J. Comp. Neurol.* 350 (4), 497–533.
- Talairach, J., Tournoux, P., 1988. *Coplanar Stereotaxic Atlas of the Human Brain*. Thieme, New York.
- Tamber-Rosenau, B.J., Esterman, M., Chiu, Y.C., Yantis, S., 2011. Cortical mechanisms of cognitive control for shifting attention in vision and working memory. *J. Cogn. Neurosci.*, <http://dx.doi.org/10.1162/jocn.2011.21608>
- Thiel, C.M., Zilles, K., Fink, G.R., 2004. Cerebral correlates of alerting, orienting and reorienting of visuospatial attention: an event-related fMRI study. *Neuroimage* 21, 318–328.
- Thiel, C.M., Zilles, K., Fink, G.R., 2005. Nicotine modulates reorienting of visuospatial attention and neural activity in human parietal cortex. *Neuropsychopharmacology* 30, 810–820.
- Turke-Brown, N.B., Scholl, B.J., Johnson, M.K., Chun, M., 2010. Implicit perceptual anticipation triggered by statistical learning. *J. Neurosci.* 30 (33), 11177–11187.
- Ugurbil, K., Garwood, M., Ellermann, J., Hendrich, K., Hinke, R., Hu, X.P., Kim, S.G., Menon, R., Merkle, H., Ogawa, S., Salmi, R., 1993. Imaging at high magnetic-fields – initial experiences at 4-T. *Magn. Reson. Q.* 9, 259–277.
- Worsley, K.J., Friston, K.J., 1995. Analysis of fMRI time series revisited – again. *Neuroimage* 2 (3), 173–181.
- Zacks, J.M., Sargent, J.Q., 2009. Event perception: a theory and its application to clinical neuroscience. In: Ross, B.H. (Ed.), *Psychol Learn Motiv* 53, 1–45.
- Zacks, J.M., Braver, T.S., Sheridan, M.A., Donaldson, D.I., Snyder, A.Z., Ollinger, J.M., Buckner, R.L., Raichle, M.E., 2001. Human brain activity time-locked to perceptual event boundaries. *Nat. Neurosci.* 4 (6), 651–655.
- Zacks, J.M., Speer, N.K., Swallow, K.M., Braver, T.S., Reynolds, J.R., 2007. Event perception: a mind-brain perspective. *Psychol. Bull.* 133 (2), 273–293.
- Zacks, J.M., Kurby, C.A., Eisenberg, M.L., Haroutunian, N., 2011. Prediction error associated with perceptual segmentation of naturalistic events. *J. Cogn. Neurosci.* 23 (12), 4057–4066.
- Zalla, T., Plassiard, C., Pillon, B., Grafman, J., Sirigu, A., 2001. Action planning in a virtual context after prefrontal cortex damage. *Neuropsychologia* 39 (8), 759–770.