Neural changes when actions change: Adaptation of strong and weak expectations

Authors:

Anne-Marike Schiffer*[1], Christiane Ahlheim[1], Kirstin Ulrichs, Ricarda I. Schubotz[1,2]

[1]Motor Cognition Group, Max Planck Institute for Neurological Research, Cologne, Germany

[2] Westfälische Wilhelms-Universität Münster, Institut für Psychologie, Münster, Germany

*Corresponding author:

Anne-Marike Schiffer

Motor Cognition Group

Max Planck Institute for Neurological Research

Gleueler Straße 50

50931 Cologne, Germany

Phone: +49 (0)221 4726 303

Fax: +49 (0)221 4726 298

E-Mail: schiffer@nf.mpg.de

Running Title: Solidity dependent model adaptation

Abstract

Repeated experiences with an event create the expectation that subsequent events will expose an analog structure. These spontaneous expectations rely on an internal model of the event that results from learning. But what happens when events change? Do experience-based internal models get adapted instantaneously, or is model adaptation a function of the solidity of, i.e. familiarity with, the corresponding internal model? The present fMRI study investigated the effects of model solidity on model adaptation in an action observation paradigm. Subjects were made acquainted with a set of action movies that displayed an altered script when encountered again in the scanning session. We found model adaptation to result in an attenuation of the premotor-parietal network for action observation. Model solidity was found to modulate activation in the parahippocampal gyrus and the anterior cerebellar lobules, where increased solidity correlated with activity increase. Finally, the comparison between early and late stages of learning indicated an effect of model solidity on adaptation rate. This contrast revealed the involvement of a fronto-mesial network of Brodmann area 10 and the ACC in those states of learning that were signified by high model solidity, no matter if the memorized original or to the altered action model was the more solid component. Findings suggest that the revision of an internal model is dependent on its familiarity. Unwarranted adaptations, but also perseverations may thus be prevented.

1 Introduction

We don't inspect events without expecting their course. According to the predictive coding account of action observation, action perception triggers an "internal model" (Kilner, Friston, & Frith, 2007; Neal & Kilner, 2010) that is run in real time and consists of predictions on the course of action (Schutz-Bosbach & Prinz, 2007). Evidently, such predictions save resources (Zacks, Speer, Swallow, Braver, & Reynolds, 2007).

However, it is not only of tremendous importance to establish internal models through experience, but also to attune them to persistent changes, and thus maintain valid predictions. Consider being forced to change your well-known way to work because of some indiscernible traffic condition at some point of the route. If this happens once, you

may surely assume that something like a traffic accident has happened. In all probability you would not decide to take another way to work on the next day. This is an example of a well-established and therefore solid internal model being violated. Solidity means that a model has strong connection weights between encompassed events. Events that have through repeated exposure become very well associated with each other elicit implicit prediction of each other. Solidity, i.e. a large strength of association, determines that the deviation is treated as a one-time occurrence of no further importance for future predictions.

Now consider being on holiday and the road to the beach being blocked on your second day in the unfamiliar countryside. You may start wondering whether you have chosen exactly the way you went the day before and try to reverse the mental map you have created of your surroundings. If you find a new way to the beach and follow it on all occasions thereafter, you may quite forget, or begin to doubt that another way has ever been possible. This form of adaptation seems likely in case of low familiarity, i.e. a weak internal model. The weak internal model is questioned and possibly revised after a one-time breach of expectation. However, it remains to be experimentally established how an internal model's solidity influences its revision and hence adaptation of predictions. To our knowledge only a few studies on reversal learning in stimulus-response paradigms (Ghahremani, Monterosso, Jentsch, Bilder, & Poldrack, 2009) have dealt with the influence of model solidity on adaptation; no study has addressed the question in an action observation paradigm.

The present fMRI study was designed to investigate the influence of model solidity on its adaptation during iterations of a divergent script. Internal models of different solidity were established by presenting a number of scripts, i.e. movies showing everyday actions (as will be described below in more detail). The concept of solidity is similar to associative strength (McClelland, et al., 1995) between components of an internal representation. Thus, solidity pertains to an internal model whose constituent events are highly associated with each other. Hence, in a fixed temporal schedule, each constituent elicits prediction of the next. This prediction is a consequence of statistical learning (Turke-Brown, et al., 2010). Statistical learning results from repeated pairing of events, i.e. stimulus familiarity, that has been proposed to be critical in extending the

persistence of memory (Eichenbaum, 2000). Concisely, repeated exposure leads to solidity. In a solid model, each event is highly associated with its neighbor. Solidity was expected to affect adaptation rate to subsequent script change. Within the Bayes' theorem framework, the goal probabilistic learning can be described as the acquisition of appropriate models for inference based on past experience. Events that co-occur persistently shape a solid model. The estimated likelihood of an event is dependent on its base-rate and how reliably it occurred in the past given that an associated event had happened. This likelihood is adapted on each iteration of the predictive and the associated event. (Fiser, et al., 2010). The more often one event has followed another, the closer is the association between them and the more likely seems the succession. Hence, within solid models, the likelihood of the respective next event is very high. This tying of prediction to a conditional probability is proposed to result in slower adaptation of more solid models. It takes longer to rewrite, or rather rewire, strong associations. Lastly, we were interested in "biased" adaptation stages at early and advanced stages of learning. In biased stages, the number of iterations of divergent expositions differed considerably from the number of iterations of the respective original script. These states are of specific interest to the validation of predictions. To resurrect the picture outlined above, only a well-known path blocked/diverged instigates maintenance of the original idea, or 'shielding' predictions from divergent influences. But previous experiences in a new environment should pale in insignificance to repeatedly coming across a divergence for the creation of an internal script and its predictions.

Functional neuroanatomy

As a main effect of the factor ADAPTATION, we expected adaptation of the internal model to the divergent script to lead to BOLD attenuation in a premotor-parietal network. The premotor-parietal network is associated with action observation and prediction of external events (cf. Schubotz, 2007). Its parietal constituent is associated with coding for object pragmatics and space (Fagg & Arbib, 1998). The frontal constituent, the lateral premotor cortex has been suggested to code for transformations underlying both our movements as well as observed events, for example changes in the position of objects (cf. Schubotz, 2007), and hence contributes to both action planning and action prediction. The

concept of prediction refers to 'filtering' of anticipated perception as has been described in motor control theories (Wolpert & Flanagan, 2001; cf. Schubotz 2007). We therefore expected that repeated exposure of the same action would lead to a decrease of activity in the premotor-parietal network, signifying adaptation.

As a main effect of the factor SOLIDITY, we hypothesized higher activity for more solid compared to weaker models in the hippocampal formation. The close proximity of the concept of solidity to associative strength (Eichenbaum, 2000; Kim & Baxter, 2001; McClelland, et al., 1995) and probabilistic learning (Kim & Baxter, 2001; Turke-Brown, et al., 2010) points towards an involvement of the hippocampal cortex, revealed in stronger activity for more solid compared to less solid models (Eichenbaum, 2000; Kim & Baxter, 2001; McClelland, et al., 1995; Turke-Brown, et al., 2010).

Finally, we expected a significant interaction of the factors SOLIDITY and ADAPTATION. This common-sense assumption is supported by the fact that habits (also habits of thought), as an example for solid associations, are particularly difficult to unlearn (see Graybiel, 2008 for a review). Moreover it has been established that stable environments, which by inference allow shaping solid models, are signified by a slow learning rate (Rushworth & Behrens, 2008). However, as the neural correlates of an influence of solidity on adaptation have not been investigated so far, the study was explorative concerning the existence and location of the interaction's neural correlates.

Implementation

To test our hypothesis, we familiarized participants previous to the fMRI session with a number of scripts containing everyday life actions, for example, a movie of making a salad. Each script encompassed a number of action steps for example, taking a bowl, grasping the lettuce, placing it in the bowl, sprinkling vinegar on top, taking salad tongues, tossing the salad. Original scripts were presented either 3, 6 or 9 times in a pre-experimental exposition session. In the fMRI session, participants encountered some scripts in the same version as before. Some scripts, however, the sequence changed from a certain point on. For example, the salad script now contained the sub-events taking the bowl, grasping the lettuce, placing it in the bowl, reaching for the cheese, reaching for a knife, cutting pats of cheese into the bowl. Note that divergent scripts did not contain any

action slips but were actions as valid as the original. Each script was shown nine times during the fMRI, either 9 times in the original or 9 times in the divergent version (no script appeared in two versions during the fMRI). Two main effects and their interaction were calculated:

To investigate the SOLIDITY effect, we contrasted the perception of divergent scripts with a large number (i.e. nine) of pre-experimental expositions (factor level 'solid') with the perception of divergent scripts with a low number (i.e. three) of pre-experimental expositions (factor level 'weak').

To test whether ADAPTATION would occur, we contrasted the first (i.e. first three - factor level 'first') with the last (i.e. seventh to ninth) repetitions (factor level 'last') of the divergent scripts pooled over all pre-experimental exposition frequencies.

Finally, we aimed to establish a neuronal network that would reflect the dependence of ADAPTATION rates on model SOLIDITY. To this end, we calculated the interaction contrast between the two-level factors ADAPTATION and SOLIDITY.

2 Methods

2.1 Subjects:

19 right-handed, healthy participants (seven women, age 22-30 years old. mean age 25.3 years) took part in the study. The participants were right handed as assessed with the Edinburgh Handedness Inventory (Oldfield, 1971). All participants were health screened by a physician and gave written informed consent.

2.2 Stimuli and Task:

The stimulus material contained 37 different movies of eight to 12 seconds length. The movies were shot from the third-person perspective, not showing the actor's face. They contained every-day actions, taking place at a table. Most movie scripts, e.g. making a sandwich, existed in two versions (a & b). These scripts had an identical beginning, but started to diverge at some individual point, where after no commonality existed (Figure 1). Each version of each script was filmed 18 times. Thus, even though the same script appeared repeatedly during the pre-experimental exposition and the experiment, the exact same shot of each script occurred only once. This method was

employed to minimize surface-similarities between the scripts. A subset of 13 scripts was filmed in five different versions.

The experiment consisted of a pre-experimental exposition of the movie material and an fMRI session starting exactly 15 minutes after the end of the pre-exposition. During the pre-experimental exposition session, participants were seated in a sound attenuated chamber facing a computer screen. Distance to the screen was adjusted to ensure that the video displayed on the screen did not extend 5° of visual angle. They watched 27 scripts, a third of which was displayed three times, another third six times and the last third nine times in a randomized fashion over the course of the 28 minutes lasting session. The participants saw one version of each script; but each repetition was another shot of the same script. Questions concerning whether some action or another was part of the script (e.g. "grasping an apple?") were posed on average after every fifth script to ensure ongoing attention to the stimulus material. Participants received visual feedback for 400 ms on whether they had answered correctly, incorrectly, or too late. After the pre-exposition, the participants were transferred directly to the fMRI chamber.

2.3 FMRI session

The fMRI session encompassed display of 36 different scripts. Each script was repeated nine times over the experiment. Nine scripts that had previously been displayed during the pre-exposition returned in the fMRI session in the same version as before ('originals' hereafter). Another nine of the pre-experimentally shown scripts were presented in the fMRI session only in their complementary version ('divergents' hereafter) (Figure 1 & Figure 2). The last nine scripts appeared in five different versions during the fMRI, each being displayed only once ('unpredictables' hereafter). The first third of the originals, the divergents and the unpredictables had previously been displayed three times each, the second third of all three kinds six times each, and the last third nine times each. Additionally, the design encompassed nine scripts that were completely new to the participants ("new originals") when they were displayed during the fMRI. The latter as well as the unpredictables will not be subject of the present paper but discussed in detail in a companion paper (Schiffer, Ahlheim, Wurm, & Schubotz, in prep). However, the likely psychological effect of the unpredictables should be taken into

account. Their presence and the associated experience of constantly changing scripts should decrease the likelihood of a divergent to be accepted as persistent at first encounter. That means that having seen a divergent only once does not allow the prediction that it returns in the same fashion – it could still turn out unpredictable at the second encounter. Only the second encounter of the same divergent delivers evidence that this script, albeit changed, is 'learnable'.

The randomization distributed scripts of the same function, for instance the first presentation of the divergent version, evenly across the session. Thus, the temporal correlation between the function of a script and experiment duration, as well as the accumulation of identical functions during a specific period was minimized.

During the fMRI session, participants lay supine on the scanner bed. Their head and arms were stabilized using form-fitting cushioning and their hands rested on a rubber foam tablet. On the right hand side, a response panel was mounted on the tablet and fixed with tape. With their right hand index and middle finger resting on two response buttons, participants could answer the 32 intermittent questions concerning the content within the same response-contingencies as in the pre-exposition (Figure 3). Participants had three seconds to answer the question. Feedback on whether a response had been registered or not was displayed on the screen for 400ms. The participants wore earplugs and headphones to attenuate scanner noise. Participants saw a reflection of the screen in a mirror, built into the head-coil and adjusted individually to allow for comfortable view of the entire screen. The movies did not extend further than 5° of visual angle in the mirror image of the computer screen. 16 null-events of 10 seconds length were displayed, consisting of display of the grey background on the screen. Participants were instructed to relax during null-events.

2.4 Data Acquisition

The functional imaging session took place in a 3T Siemens Magnetom Trio scanner (Siemens, Erlangen, Germany). In a separate session, prior to the functional MRI, high-resolution 3D T-1 weighted whole-brain MDEFT sequences were recorded for every participant (128 slices, field of view 256mm, 256 by 256 pixel matrix, thickness 1mm, spacing 0.25 mm)

The functional session engaged a single-shot gradient echo-planar imaging (EPI) sequence sensitive to blood oxygen level dependent contrast (28 slices, parallel to the bicommisural plane, echo time 30ms, flip angle 90°; repetition time 2000ms; serial recording). Following the functional session immediately, a set of T1-weighted 2D-FLASH images was acquired for each participant (28 slices, field of view 200mm, 128 by 128 pixel matrix, thickness 4mm, spacing 0.6mm, in-plane resolution 3 by 3 mm).

2.5 FMRI Data Analysis

Functional data were offline motion-corrected using the Siemens motion protocol PACE (Siemens, Erlangen, Germany). Further processing was conducted with the LIPSIA software package (Lohmann, et al., 2001). Cubic-spline interpolation was used to correct for the temporal offset between the slices acquired in one scan. To remove low-frequency signal changes and baseline drifts, a 1/110Hz filter was applied. The matching parameters (6 degrees of freedom, 3 rotational, 3 translational) of the T1-weighted 2D-FLASH data onto the individual 3D MDEFT reference set were used to calculate the transformation matrices for linear registration. These matrices were subsequently normalized to a standardized Talairach brain size (x=135 mm, y=175 mm, z=120 mm; Talairach and Tournoux, 1988) by linear scaling. The normalized transformation matrices were then applied to the functional slices, to transform them using trilinear interpolation and align them with the 3D reference set in the stereotactic coordinate system. The generated output had thus a spatial resolution of 3 by 3 by 3 mm.

The statistical evaluation was based on a least-square estimation using the general linear model (GLM) for serially auto-correlated observations (Worsley & Friston, 1995). Temporal Gaussian smoothing (4 seconds FWHM) was applied to deal with temporal autocorrelation and determine the degrees of freedom (Worsley & Friston, 1995). A spatial Gaussian filter of FWHM 5 mm was applied. The design matrix was generated by hemodynamic modeling using a γ-function and encompassed the first derivate. The onset vectors in the design matrix were modeled in a time-locked event-related fashion.

All contrasts were drawn from one design matrix. The first contrast accounted for the effect of model SOLIDITY. The second contrast accounted for the overall ADAPTATION

effect. The third contrast targeted the INTERACTION between model solidity and adaptation. To ensure that the activation from the interaction contrast was rooted in an orthogonal interaction, we also calculated the conjunction analysis that accounted for the same proposed interaction effect. The onset vectors were modeled to the point in time when the divergent was recognizable as divergent (hereupon 'breach', Figure 1). This breach had previously been visually timed to the moment when movement trajectories revealed that either the manipulation or the reached-for object was different from that in the originals. All divergents as well as the null-events were added as conditions of no-interest into the design matrix.

Main effect SOLIDITY

This effect was calculated as (solid / first ∩ solid / last) > (weak / last ∩ weak / first). Factor level 'solid' refers to models that had been pre-exposed nine times; factor level 'weak' refers to models that had been pre-exposed three times. Factor level 'first' refers to first three presentations of a divergent; factor level 'last' refers to its last three presentations (Figure 4).

Main effect ADAPTATION

This effect was calculated as (solid / first ∩ weak / first) > (solid / last ∩ weak / last). Please refer above for explanation of the factor levels (Figure 5).

Interaction SOLIDITY by ADAPTATION

The interaction contrast signifies the interaction between the two two-level factors SOLIDITY and ADAPTATION, and is thus derived from the crossing of the respective levels. Hence, it was calculated as contrast (solid / first > weak / first) > (solid / last > weak / last). Please refer above for explanation of the factor levels (Figure 6).

In order to enable an interpretation of the significant effects derived from this interaction contrast, it was important to ensure that all significant voxels reflected the same direction of the effect (this rationale applies to all interaction contrasts in fMRI). Therefore, we additionally calculated the conjunction of the contrasts (weak / first > weak / last) and (solid / first > solid / last).

All contrast images were fed into a second-level random effects analysis. The group analysis consisted of one-sample t tests across all contrast images to analyze whether the observed differences between conditions were significantly deviant from zero. Acquired t-values were transformed to z-scores. A two-step correction for false positive results based on a Monte-Carlo simulation was performed. In a first step, an initial z-threshold of 2.33 (p < .05, one-tailed) was applied to the simulated voxels. Afterwards, based on the remaining clusters, statistically thresholds were calculated to correct for false positives at a significance level of p = .05. Cluster size as well as cluster value were taken into account at thresholding in a compensatory matter to prevent neglecting true positive activations in small anatomical structures (Lohmann et al., 2008). Hence, all reported activations were significantly activated at p <= .05, corrected for multiple comparisons at cluster level.

2.6 Pilot study

Previous behavioral results support the validity of the described contrasts. A preceding pilot study in another group of participants had provided behavioral evidence for the influence of solidity on adaptation. In the study, participants viewed each movie first three, six, or nine times in the original version, followed by three, six, or nine divergent displays and eventually one or two original presentations. Meanwhile they had to constantly indicate whether the version that was on display at the moment was identical to the last display, or represented a change in script. We measured reaction times (RT) for the responses and conducted a repeated measures ANOVA on the RTs of all correct responses to repetitions of divergents. The repeated measures ANOVA thus included 2 factors, the 2-level factor original presentations (levels: three original presentations, nine original presentations) and 8-level factor divergent iteration (levels: 2nd ieration, 3rd iteration, …, 9th iteration). The first divergent was not included in the analysis, as it demanded a different response (indication of change) than the ensuing divergents (indication of repetition). The interaction effect between number of original presentations and iteration of the divergent approached significance at p = .07

(Greenhouse-Geisser corrected). To disentangle what effect carried the interaction we correlated the RT for each iteration with the number of previous originals. The correlation between RT of the divergents that had been displayed 3 times as original and their iterations was not significant (r= .081, p = .3). In contrast, the correlation between RT of the divergents that had been displayed 9 times as original and their iterations approached significance (r= -.157, p = .06). This marginal correlation reveals a continuous decrease in reaction times that we take to reflect ongoing adaptation to the divergents that had previously been shown nine times in their original version. Taken together, these results reflect a difference in adaptation rate dependent on the number of pre-expositions.

3 Results

3.1 Behavioural results

The participants answered on average 87% of the 32 questions correctly (< 27 questions). Standard deviation was 7%. In the post-experimental questionnaire participants were asked whether all movies had returned as before and no participant indicated that all movies had. Six of the 19 participants reported spontaneously to the open question whether they wished to report anything whatsoever, that some movies were different than before. This behavioral measures furthers the argument that the participants were aware that some movies were altered versions of what they had seen pre-experimentally, instead of believing that the different movies (divergents) were not related to the initial version.

3.2 FMRI results

The model SOLIDITY CONTRAST (solid / first ∩ solid / last) > (weak / last ∩ weak / first) yielded activity in the right parahippocampal cortex, and also in the right cerebellar lobule III (centralis) and bilaterally in the lobule IV (culmen) of the cerebellum (Table 1) (Figure 4).

The model ADAPTATION contrast (solid / first ∩ weak / first) > (solid / last ∩ weak / last) yielded bilateral activity in the inferior frontal sulcus (IFS), the left premotor cortex (PM), the left superior parietal lobe (SPL) and intraparietal sulcus (IPS), extending into anterior IPS in the left hemisphere. The posterior middle temporal gyrus (MTG) was activated bilaterally (Table 2) (Figure 5).

The SOLIDITY by ADAPTATION interaction contrast (solid / first > weak / first) > (solid / last > weak / last) showed significant activation of the frontopolar cortex comprising mesial Brodmann Area (BA) 10 and right lateral BA10. Further activations were in the anterior cingulate cortex (ACC), right orbitofrontal cortex (OFC), the right striatum, right posterior superior temporal gyrus (pSTS), cuneus and the left fusiform gyrus (Table 3) (Figure 6). The second approach to this analysis, the conjunction analysis (iii-a), i.e. (weak / last > weak / first) ∩ (solid / first > solid / last), yielded activity in the mesial and the lateral BA10, ACC and cuneus, and in the right fusiform gyrus (Table 4).

4 Discussion

Internal models of an action encompass expectations on the development of this action (Bar, 2009; Jeannerod, 1995). Valid predictions make perception more efficient and are beneficial to fast reactions (Wolpert & Flanagan, 2001). The present fMRI study investigated the neural correlates of the influence of the solidity of the original internal model of an action on subsequent adaptation of the internal model to a divergent script. To that end, participants watched movies that familiarized them with the original scripts and thus to establish internal model of them. In the fMRI they were confronted with divergent versions of the previously learnt scripts.

We found a persistent effect of pre-experimental exposition frequency (main effect of solidity) in the right parahippocampal cortex as implied by the concept's proximity to associative strength. There was also an effect of solidity bilaterally in the anterior cerebellum. This result stresses the importance of previous experience to expectation, especially in the face of new information. As hypothesized, divergent experiences incited adaptation in fronto-parietal motor regions, i.e. left PMv, bilateral IFS and IPS. Moreover

the adaptation effect was evident in the posterior MTG and in the left SPL. Finally, the exciting finding of a network dealing with a solidity bias, i.e. stages where solidity of one script surpasses that of another (solidity by adaptation interaction), supports the notion of a lasting influence of possible alternatives. The activity that was found for this interaction, located in the left frontomedian cortex (FMC), i.e. BA 10 and the ACC, as well as right striatum and right OFC, suggests a continuous processing of divergent information in these areas, be it current or past.

4.1 Solidity exerts prolonged influence

Activity in the solidity contrast reflects an ongoing response to divergent scripts that is more pronounced for solid than for weaker original internal models. The cerebellar activity was in a classical motor region (Marvel & Desmond, 2010), in lobules III and IV (Schmahman et al., 1999). Working memory function, proposed for cerebellar lobules VI/crus I (Marvel & Desmond, 2010) is rather an unlikely explanation for this anterior activity. Hence, we take it to reflect continuing mismatch between the internal motor model's expectations and perception, which is increased if the original internal model was highly habituated. The parahippocampal cortex has been associated with topographical learning (Aguirre, Detre, Alsop, & D'Esposito, 1996), scene processing (Epstein & Kanwisher, 1998) and the association of scenes and locations with objects (Bar, Aminoff, & Schacter, 2008; Sommer, Rose, Gläscher, Wolbers, & Büchel, 2005). Here, we propose that parahippocampal activity signifies the revision of associations (Eichenbaum, 2000; McClelland, et al., 1995) between scenes and actions or action-relevant objects. The present data allow no decision between these alternatives as the divergent script sometimes included the use of a different object than the original script did, but sometimes only entailed an altered manipulation of the same object.

4.2 Adaptation in the cortical motor network

The adaptation contrast (ii) yielded activity in the left PM(v), the bilateral IPS and the left posterior MTG, a network that is not only relevant for action execution, but also prominent in action observation (Jeannerod, 1995). The adaptation contrast tested

whether the hypothesized fronto-parietal motor regions would be sensitive to violated expectations and show an adaptation to the new action script.

During the first encounters of the divergent script, perception was assumed to deviate from the internal model. An increase of neuronal activity at this stage reflects a breach of expectation signal that incites learning (Summerfield, Trittschuh, Monti, Mesulam, & Egner, 2008). This signal can also be understood as a correlate of the processing of unexpected (salient) objects or manipulations (Keysers & Perret, 2004). These functions can be seen as two sides of the same coin: Accordingly, the original script acts like a filter that minimizes processing demands of all according perceptions. Divergent perceptions, however, are not filtered, rendering them more salient than pre-filtered perceptions. The resulting increased activation is a 'breach of expectation signal' and incites learning. As soon as the divergent script has been learnt, it can serve as a filter for all according perceptions again.

Adapting the internal model to account for the divergent script is a learning or re-learning process, and in a stable environment, strong evidence should be required to motivate learning (Rushworth & Behrens, 2008). Otherwise, assembling and memorizing experiences would be pointless, as they would loose their capacity to guide successful behavior as soon as a one-time breach of expectation had occurred. Hence, the divergent perception should not cause instantaneous adaptation of the internal model; accordingly, a process of adaptation is revealed by diminution of the neural correlate of divergence over a large number of iterations (Friston, Kilner, & Harrison, 2006; Grill-Spector, Henson, & Martin, 2006; Majdanžić, Bekkering, van Schie, & Toni, 2009) as targeted in the adaptation contrast (ii). It has previously been established that the cortical motor network is capable of predicting the ongoing course of action (Jeannerod, 1995). The current study furthers our understanding thereof, suggesting that the network is sensitive to salient violations of its predictions and shows appropriately slow adaptation. A detailed account of the proposed functions of the constituents adapting in this process will be supplied below.

The SPL has been discussed as a potential site of spatial priority maps, which designate relevant object locations and can be internally guided or externally cued

(Molenberghs, Mesulam, Peeters, & Vandenberghe, 2007; Nobre, et al., 2004); one of the SPL's functions seems to be constructing and changing these spatial priority maps (Chiu & Yantis, 2009; Molenberghs, et al., 2007). Activity in the adaptation contrast is evidence for the remapping of spatially guided attention in SPL; this remapping or changing of weights in the priority map (Molenberghs, et al., 2007) becomes important to action emulation as suddenly relevant objects demand attention, while previously used objects loose their significance for the action sequence.

Activity in the posterior MTG is taken to reflect increased processing of the movements of the actor and the actions associated with suddenly relevant objects (Beauchamp, et al., 2002; Beauchamp & Martin, 2007). Divergent scripts encompassed use (and accordingly motion) of different objects or different use of the same object as the original scripts. Encounter of the first presentations of the divergent script entailed a mismatch between emulated associations and valid, but unpredicted perceived use. Activity in the posterior MTG has been discussed in association with the frontoparietal motor network (Beauchamp & Martin, 2007; Johnson-Frey, 2004). The role of this frontoparietal network of IPS and PM in goal-directed object manipulation and internal modeling thereof has been researched extensively (Grèzes & Decety, 2001; Jeannerod, 1995; Johnson-Frey, 2004 for reviews). The anterior IPS has been proposed to provide the ventral premotor cortex with information on object pragmatics (Fagg & Arbib, 1998; Schubotz & von Cramon, 2008). Attenuation of its activity has previously been interpreted as a teaching signal that allows model adaptation (Tunik, Rice, Hamilton, & Grafton, 2007). Medial IPS has previously been reported to be crucial for the online control of goal-directed precision movement (Grefkes & Fink, 2005 for a review). Online correction relies on the detection of mismatch between internal emulation and sensorimotor information (Wolpert & Flanagan, 2001). We suggest that the activity along IPS reflects a decreasing mismatch between the internal model's emulated action and the currently perceived action. The closely linked (Geyer, Matelli, Luppino, & Zilles, 2000) PMv, which is assumed to store action knowledge and object function, shows increased activity when new scripts have to be learnt (see Schubotz & von Cramon, 2003 for review). Activity in premotor cortex is increased when prediction (Schubotz & von Cramon, 2003), or simulation (Grèzes & Decety, 2001), and planning of movements

(Johnson-Frey, 2004) is involved. Against this backdrop, PMv activation during the first encounters of unpredicted divergences can be regarded as further evidence of this area's involvement in compiling complex actions.

4.2.1 Initial bias towards the original script

Activity in IFS has been suggested to modulate the bias between competing representations (Badre, Poldrack, Parè-Blagoev, Insler, & Wagner, 2005; Kuhl, Dudukovic, Kahn, & Wagner, 2007; Wurm & Schubotz, 2011). This fits well with an influential model of prefrontal cortex function that suggests that prefrontal cortex is involved in activating and supporting relevant but unfavored or weak associations (Miller & Cohen, 2001). The present study delivers new evidence for the assumption that the IFS supports weak models: attenuation of IFS activity points to its involvement in supporting the new divergent internal model and its associations during the first encounters of the divergent script. Each iteration of this divergent script should solidify its representation, diminishing IFS activity as a balanced state of competition between original and divergent internal model is approached and the bias runs eventually in favor of the new internal model (Schubotz & von Cramon, 2008).

4.3 Bias vs. Balance – prefrontally mediated integration of incompatible models

The activation of the FMC, occipital areas, as well as the pSTS in the SOLIDITY-ADAPTATION-interaction contrast revealed these areas' involvement in processing information when the solidity of one internal model surpasses that of another. Strikingly, this network was found to be involved not only when this bias run in favor of the original script (and hence, against the currently perceived one), but also when the bias was already towards the actually presented action (an hence, against the former original script). The underlying analysis was explorative concerning the areas that would be involved in the interaction of SOLIDITY and ADAPTATION. However, the interesting results help to explain previous puzzling findings (Frank et al., 2005) and enhance our understanding of a conundrum in the EEG-centered conflict-monitoring literature:

FMC activity spread from the ACC into BA10. The ACC is understood to be responsive to bias, especially in decision and stimulus-response paradigms (Bunge, Burrows, & Wagner, 2004; Miller & Cohen, 2001). It is supposed to convey this bias to

the dorsolateral prefrontal cortex (Miller & Cohen, 2001). Classic bias-related responses recorded in the ACC focus on conflict (see Botvinick, Cohen, & Carter, 2004; van Veen & Carter, 2002 for review). Conflict is often understood as bias running against the necessary association, demanding PFC to support or maintain activation of a 'weaker' association (Kuhl, et al., 2007; Miller & Cohen, 2001). This 'conflict solving', triggered by the ACC, could also mean suppression of an unlikely target (Kuhl, et al., 2007), apart from the classic conception as fostering a weaker alternative (Miller & Cohen, 2001). The current study, in contrast, revealed that the ACC is active for both biased states, even when perception is in accordance with the currently more solid internal representation. This latter form of bias, however, is not signified by what is often understood as conflict, i.e., the need to resolve competition in favor of the weaker alternative. Consequently, IFS activation is diminished at this stage, as apparent in the adaptation contrast and discussed above, while it is present when bias does run against the presented model at the beginning of adaptation. The proposed bias account is in line with an account of ACC function that integrates conflict monitoring and more general evaluative computation (Botvinick, et al., 2004). Conflict would then mean the activation of the representations of two incompatible (action) models (Botvinick, et al., 2004). The present results seem to singularly underpin a point in the EEG literature of conflict monitoring with fMRI-derived results. Yeung and colleagues (2004) argue that the N2 component in correct trials and ERN component following errors is elicited when evidence for one representation outweighs that for the other – with the N2 preceding correct responses and the ERN being a post-error correlate of surmounting evidence for the (discarded) correct response. This aspect of 'outweighing' the competing alternative, or bias, has however not always been taken into consideration in the conflict monitoring literature even though one study (Frank et al., 2005) found that in a forced choice task, a higher discrepancy between the respective reward values of two options resulted in a higher ERN than a more equal distribution of reward. Our study reveals that activity in the FMC is stronger if evidence is biased in favor of one of the incompatible representations, indicating in this case a higher predictive capacity for one model than the other. The study thus contributes to the clarification of the EEG centered conflict monitoring debate (Botvinick et al.,

2004), corroborating a bias-related definition of conflict, as opposed to the notion of equally strong competitors.

The ACC is closely linked to BA10 (Allman, Hakeem, & Watson, 2002). A special kind of neuron, the spindle neurons in the ACC have been proposed to convey the motivation to adapt to changes to BA10 (Allman, et al., 2002). More generally, the frontopolar area is part of the hippocampal-cortical memory system (Vincent, Kahn, Snyder, Raichle, & Buckner, 2008). Moreover, BA10 is taken to be responsible for the integration of separate cognitive operations (see Ramnani & Owen, 2004 for review). One example is episodic retrieval and success monitoring, a process that can be understood in terms of comparing an internal representation to an outcome (Ramnani & Owen, 2004). We propose that only the biased states entailed suppression of either the original or the divergent internal model, respectively. The deterministic nature of the paradigm suggested solidifying the divergent internal model, thus the biased and balanced states both encompassed a need to register and to encode the divergent internal model. But the biased states also suggested suppression of either the original or the divergent. If there were no suppression of the divergent internal model in the beginning, learning would be instantaneous. This was not the case. If the diversion was not registered, accumulating evidence would not be tracked and learning would never set in. Once evidence for the validity of the divergent internal model outweighs that for the original, suppression of the neglected alternative is regarded as efficient (Kuhl, et al., 2007) and guides expectations towards the most likely outcome. A coupling of the ACC and BA10 during suppression has previously been reported by Kuhl and colleagues (Kuhl, et al., 2007). In the balanced states, evidence for neither internal model outweighs evidence for the other and suppression could be regarded as too persistent (for the divergent internal model) or too premature (for the original internal model), respectively.

Activity of the OFC in the interaction contrast complements the emerging picture (Ghahremani, et al., 2009). Biased states necessarily have one strong, or solid component, like a prepotent response or well practiced forward model. As discussed above, this strong component can trigger suppression of alternatives as it allows generation of hypotheses. Both, hypothesis generation and suppression have been discussed as potential

OFC functions. (Elliott, Dolan, & Frith, 2000; Ghahremani, et al., 2009; Vartanian & Goel, 2005). Hypothesis generation and suppression can be reframed as evaluation or weight changes as a results of evaluation, which itself is a function ascribed to the OFC (Wallis, 2007). A steady environment, as signified by the existence of one solid internal model, makes it worthwhile to track contingencies and integrate outcome histories into learning (Rushworth & Behrens, 2008). Responses to contingency differences, another type of evaluation, have similarly been allocated in the OFC (Windmann, et al., 2006). We propose that the activity increase in the OFC during a state of bias is indicative of the evaluation of the current forward model (Schubotz & von Cramon, 2008) against the backdrop of one solid and one weak or paling internal model. Closely linked to the OFC in its evaluative function is the striatum that was similarly active in the interaction contrast (Grinband, Hirsch, & Ferrera, 2006; Oenguer, Ferry, & Price, 2003; Schoenbaum, Roesch, Stalnaker, & Takahashi, 2009).

To sum up, the similarities the networks display during the beginning and during an advanced state of learning single model solidity bias out as the determinant factor, as opposed to conflict between equally strong representations. It is likely that there is only consolidation in the balanced state, but an integration of consolidation of one and suppression of the other internal model in the biased states. Thus, bias incites the same operation in different situations, i.e. suppression of the divergent internal model in the beginning and suppression of the original internal model in the end. In the beginning, the divergent script stands in stark contrast to a solid internal model with identical onset phases; hence, it demands attention (Summerfield, 2008), possibly against a backdrop of previous suppression. In the end, even though the old original internal model has not been valid for a large number of iterations, it still exerts an influence on predictions. The emergence of significant bias-related activations suggests that the opposite, i.e., a state of balance or ambiguity, is reached when the number of expositions of the divergent script matches the number of previous expositions of the original script. This finding is indicative of a slower adaptation rate for a solid, compared to a weak internal model and supported by the data from the pilot study (see Methods section 2.6).

4.4 Concluding remarks

In a dynamic environment, it is particularly important not only to set up internal models but also to keep them up to date. Hence, expectations must be revised if they do not accord to our last experiences. However, unwarranted revision should be prevented, to not loose the gain of experience. The current study provided evidence for the notion that familiarity with an event influences the adaptation rate of according expectations.

5 References

Aguirre, G. K., Detre, J. A., Alsop, D. C., & D'Esposito, M. (1996). The Parahippocampus Subserves Topographical Learning in Man. Cerebral Cortex, 6(6), 823-829.

Allman, J., Hakeem, A., & Watson, K. (2002). Book Review: Two Phylogenetic Specializations in the Human Brain. The Neuroscientist, 8(4), 335-346.

Badre, D., Poldrack, R. A., Parè-Blagoev, E. J., Insler, R. Z., & Wagner, A. D. (2005). Dissociable Controlled Retrieval and Generalized Selection Mechanisms in Ventrolateral Prefrontal Cortex. [doi: DOI: 10.1016/j.neuron.2005.07.023]. Neuron, 47(6), 907-918.

Bar, M. (2009). The proactive brain: memory for predictions. Philosophical Transactions of the Royal Society B: Biological Sciences, 364(1521), 1235-1243.

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.

Beauchamp, M. S., Lee, K. E., Haxby, J. V., & Martin, A. (2002). Parallel Visual Motion Processing Streams for Manipulable Objects and Human Movements. [doi: DOI: 10.1016/S0896-6273(02)00642-6]. Neuron, 34(1), 149-159.

Beauchamp, M. S., & Martin, A. (2007). Grounding Object Concepts in Perception and Action: Evidence from FMRI Studies of Tools. Cortex; a journal devoted to the study of the nervous system and behavior, 43(3), 461-468.

Botvinick, M. M., Cohen, J. D., & Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: an update. [doi: DOI: 10.1016/j.tics.2004.10.003]. Trends in Cognitive Sciences, 8(12), 539-546.

Bunge, S. A., Burrows, B., & Wagner, A. D. (2004). Prefrontal and hippocampal contributions to visual associative recognition: Interactions between cognitive control and episodic retrieval. [doi: DOI: 10.1016/j.bandc.2003.08.001]. Brain and Cognition, 56(2), 141-152.

Caspers, S., Zilles, K., Laird, A. R., & Eickhoff, S. B. (2010). ALE meta-analysis of action observation and imitation in the human brain. [doi: DOI: 10.1016/j.neuroimage.2009.12.112]. NeuroImage, 50(3), 1148-1167.

Chiu, Y.-C., & Yantis, S. (2009). A Domain-Independent Source of Cognitive Control for Task Sets: Shifting Spatial Attention and Switching Categorization Rules. J. Neurosci., 29(12), 3930-3938.

Eichenbaum, H. (2000). A cortical-hippocampal system for declarative memory. [10.1038/35036213]. Nat Rev Neurosci, 1(1), 41-50.

Elliott, R., Dolan, R. J., & Frith, C. D. (2000). Dissociable Functions in the Medial and Lateral Orbitofrontal Cortex: Evidence from Human Neuroimaging Studies. Cerebral Cortex, 10(3), 308-317.

Epstein, R., & Kanwisher, N. (1998). A cortical representation of the local visual environment. [10.1038/33402]. Nature, 392(6676), 598-601.

Fagg, A. H., & Arbib, M. A. (1998). Modeling parietal-premotor interactions in primate control of grasping. [doi: DOI: 10.1016/S0893-6080(98)00047-1]. Neural Networks, 11(7-8), 1277-1303.

Fiser, J., Berkes, P., Orbán, G., & Lengyel, M. (2010). Statistically optimal perception and learning: from behavior to neural representations. [doi: 10.1016/j.tics.2010.01.003]. Trends in Cognitive Sciences, 14(3), 119-130.

Frank, M. J., Woroch, B. S., & Curran, T. (2005). Error-Related Negativity Predicts Reinforcement Learning and Conflict Bias. Neuron, 47, 495-51.

Friston, K., Kilner, J., & Harrison, L. (2006). A free energy principle for the brain. [doi: DOI: 10.1016/j.jphysparis.2006.10.001]. Journal of Physiology-Paris, 100(1-3), 70-87.

Geyer, S., Matelli, M., Luppino, G., & Zilles, K. (2000). Functional neuroanatomy of the primate isocortical motor system. Anatomy and Embryology, 202(6), 443-474.

Ghahremani, D. G., Monterosso, J., Jentsch, J. D., Bilder, R. M., & Poldrack, R. A. (2009). Neural Components Underlying Behavioral Flexibility in Human Reversal Learning. Cerebral Cortex.

Graybiel, A.M. (2008). The basal ganglia: Learning new tricks and loving it. Current Opinion in Neurobiology. 15(6):638-644.

Grèzes, J., & Decety, J. (2001). Functional anatomy of execution, mental simulation, observation, and verb generation of actions: A meta-analysis. [10.1002/1097-0193(200101)12:1<1::AID-HBM10>3.0.CO;2-V]. Human Brain Mapping, 12(1), 1-19.

Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: neural models of stimulus-specific effects. [doi: DOI: 10.1016/j.tics.2005.11.006]. Trends in Cognitive Sciences, 10(1), 14-23.

Grinband, J., Hirsch, J., & Ferrera, V. P. (2006). A Neural Representation of Categorization Uncertainty in the Human Brain. Neuron, 49, 757-763.

Jeannerod, M. (1995). Mental imagery in the motor context. [doi: DOI: 10.1016/0028-3932(95)00073-C]. Neuropsychologia, 33(11), 1419-1432.

Johnson-Frey, S. H. (2004). The neural bases of complex tool use in humans. [doi: DOI: 10.1016/j.tics.2003.12.002]. Trends in Cognitive Sciences, 8(2), 71-78.

Keysers, C., and Perrett, D.I. (2004). Demystifying social cognition: a Hebbian perspective. Trends in Cognitive Sciences, (11): 501-7.

Kilner, J. M., Friston, K. J., & Frith, C. D. (2007). Predictive coding: an account of the mirror neuron system. Cogn Process, 8(3), 159-166.

Kim, J. J., & Baxter, M. G. (2001). Multiple brain-memory systems: the whole does not equal the sum of its parts. [doi: 10.1016/S0166-2236(00)01818-X]. Trends in Neurosciences, 24(6), 324-330.

Kuhl, B. A., Dudukovic, N. M., Kahn, I., & Wagner, A. D. (2007). Decreased demands on cognitive control reveal the neural processing benefits of forgetting. [10.1038/nn1918]. Nat Neurosci, 10(7), 908-914.

Lohmann, G., Mueller, K., Bosch, V., Mentzel, H., Hessler, S., Chen, L., et al. (2001). Lipsia - a new software system for the evaluation of functional magnetic resonance images of the human brain. Computerized medical imaging and graphics : the official journal of the Computerized Medical Imaging Society, 25(6), 449-457.

Majdanžić, J., Bekkering, H., van Schie, H. T., & Toni, I. (2009). Movement-Specific Repetition Suppression in Ventral and Dorsal Premotor Cortex during Action Observation. Cerebral Cortex, 19(11), 2736-2745.

Marvel, C., & Desmond, J. (2010). Functional Topography of the Cerebellum in Verbal Working Memory. Neuropsychology Review, 20(3), 271-279-279.

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why There Are Complementary Learning Systems in the Hippocampus and Neocortex: Insights From the Successes and Failures of Connectionist Models of Learning and Memory. [doi:]. Psychological Review, 102(3), 419-457.

Miller, E. K., & Cohen, J. D. (2001). AN INTEGRATIVE THEORY OF PREFRONTAL CORTEX FUNCTION. [doi: 10.1146/annurev.neuro.24.1.167]. Annual Review of Neuroscience, 24(1), 167-202.

Molenberghs, P., Mesulam, M. M., Peeters, R., & Vandenberghe, R. R. C. (2007). Remapping Attentional Priorities: Differential Contribution of Superior Parietal Lobule and Intraparietal Sulcus. Cerebral Cortex, 17(11), 2703-2712.

Neal, A., & Kilner, J. M. (2010). What is simulated in the action observation network when we observe actions? [10.1111/j.1460-9568.2010.07435.x]. European Journal of Neuroscience, 32(10), 1765-1770.

Nobre, A. C., Coull, J. T., Maquet, P., Frith, C. D., Vandenberghe, R., & Mesulam, M. M. (2004). Orienting Attention to Locations in Perceptual Versus Mental Representations. Journal of Cognitive Neuroscience, 16(3), 363-373.

Oenguer, D., Ferry, A. T., & Price, J. L. (2003). Architectonic subdivision of the human orbital and medial prefrontal cortex (Vol. 460). New York, NY, ETATS-UNIS: Wiley-Liss.

Ramnani, N., & Owen, A. M. (2004). Anterior prefrontal cortex: insights into function from anatomy and neuroimaging. [10.1038/nrn1343]. Nat Rev Neurosci, 5(3), 184-194.

Rushworth, M. F. S., & Behrens, T. E. J. (2008). Choice, uncertainty and value in prefrontal and cingulate cortex. [10.1038/nn2066]. Nat Neurosci, 11(4), 389-397.

Schiffer, A.-M., Ahlheim, C., Wurm, M. F., & Schubotz, R. I. (in prep).

Schoenbaum, G., Roesch, M. R., Stalnaker, T. A., & Takahashi, Y. K. (2009). A new perspective on the role of the orbitofrontal cortex in adaptive behaviour. Nature Review Neuroscience, 10(12), 885-892.

Schubotz, R. I. (2007). Prediction of external events with our motor system: towards a new framework. [doi: 10.1016/j.tics.2007.02.006]. Trends in Cognitive Sciences, 11(5), 211-218.

Schubotz, R. I., & von Cramon, D. Y. (2003). Functional-anatomical concepts of human premotor cortex: evidence from fMRI and PET studies. [doi: DOI: 10.1016/j.neuroimage.2003.09.014]. NeuroImage, 20(Supplement 1), S120-S131.

Schubotz, R. I., & von Cramon, D. Y. (2008). The Case of Pretense: Observing Actions and Inferring Goals. [doi: 10.1162/jocn.2009.21049]. Journal of Cognitive Neuroscience, 21(4), 642-653.

Schutz-Bosbach, S., & Prinz, W. (2007). Prospective coding in event representation. Cogn Process, 8(2), 93-102.

Sommer, T., Rose, M., Gläscher, J., Wolbers, T., & Büchel, C. (2005). Dissociable contributions within the medial temporal lobe to encoding of object-location associations. Learning & Memory, 12(3), 343-351.

Summerfield, C., Trittschuh, E. H., Monti, J. M., Mesulam, M. M., & Egner, T. (2008). Neural repetition suppression reflects fulfilled perceptual expectations. [10.1038/nn.2163]. Nat Neurosci, 11(9), 1004-1006.

Tunik, E., Rice, N. J., Hamilton, A., & Grafton, S. T. (2007). Beyond grasping: Representation of action in human anterior intraparietal sulcus. [doi: DOI: 10.1016/j.neuroimage.2007.03.026]. NeuroImage, 36(Supplement 2), T77-T86.

Turke-Brown, N. B., Scholl, B. J., Johnson, M. K., & Chun, M. (2010). Implicit Perceptual Anticipation Triggered by Statistical Learning. Journal of Neuroscience, 30(33), 11177.

van Veen, V., & Carter, C. S. (2002). The anterior cingulate as a conflict monitor: fMRI and ERP studies. [doi: DOI: 10.1016/S0031-9384(02)00930-7]. Physiology & Behavior, 77(4-5), 477-482.

Vartanian, O., & Goel, V. (2005). Task constraints modulate activation in right ventral lateral prefrontal cortex. [doi: DOI: 10.1016/j.neuroimage.2005.05.016]. NeuroImage, 27(4), 927-933.

Vincent, J. L., Kahn, I., Snyder, A. Z., Raichle, M. E., & Buckner, R. L. (2008). Evidence for a Frontoparietal Control System Revealed by Intrinsic Functional Connectivity. [Article]. Journal of Neurophysiology, 100(6), 3328-3342.

Windmann, S., Kirsch, P., Mier, D., Stark, R., Walter, B., Guentuerkuen, O., et al. (2006). On Framing Effects in Decision Making: Linking Lateral versus Medial Orbitofrontal Cortex Activation to Choice Outcome Processing. [doi: 10.1162/jocn.2006.18.7.1198]. Journal of Cognitive Neuroscience, 18(7), 1198-1211.

Wolpert, D. M., & Flanagan, J. R. (2001). Motor prediction. Current Biology 11(19), 729-732.

Worsley, K. J., & Friston, K. J. (1995). Analysis of FMRI time series revisited - again. Neuroimage, 2(3), 173-181.

Wurm, M. F., & Schubotz, R. I. (2011). Actions implied by domestic settings: The influence of spatial context on action recognition. Unpublished data.

Zacks, J. M., Speer, N. K., Swallow, K. M., Braver, T. S., & Reynolds, J. R. (2007). Event perception: A mind-brain perspective. [doi:10.1037/0033-2909.133.2.273]. Psychological Bulletin, 133(2), 273-293.

Tables:

Table 1: Solidity contrast: Anatomical specification, Talairach coordinates (x,y,z) and maximal Z-scores of significantly activated voxels for model solidity: divergents with high (9 pre-expositions) or weak (3 pre-expositions) model solidity.

Localisation	Talairach coordinates			z-values, local maxima
	x	y	z	
Parahippocampal Cortex	32	-32	-12	3.43
Cerebellum, lobule III, Centralis	4	-38	-9	5.16
Cerebellum, lobule IV, Culmen	-8	-47	-18	4.8

Table 2: Adaptation contrast: Anatomical specification, Talairach coordinates (x,y,z) and maximal Z-scores of significantly activated voxels for the model adaptation: first vs. last presentations of divergents.

Localisation	Talairach coordinates			Z-values, local maxima
	x	y	z	
Superior parietal lobule	-14	-59	-57	3.67
Intraparietal sulcus	32	-62	45	4.18
	-20	-65	39	3.74
	-40	-41	54	3.6

Intraparietal sulcus, anterior segment	-58	-23	42	2.9
Premotor cortex	-46	10	24	3.69
Inferior frontal gyrus	46	16	30	3.72
	-44	22	24	3.16
Posterior middle temporal gyrus	44	-56	15	3.72
	40	-47	-3	3.81
	-46	-65	12	3.25
	-40	-50	-6	3.4

Table 3: Interaction contrast: Anatomical specification, Talairach coordinates (x,y,z) and maximal Z-scores of significantly activated voxels for biased vs. balanced states: the first divergents of a solid internal model and the last divergents of a weak internal model vs. the first divergents of a weak internal model and the last divergents of a solid internal model.

Localisation	Talairach coordinates			Z-values, local maxima
	x	y	z	
Frontal pole, BA10	-10	61	12	4.31
	14	52	9	3.33
Anterior cingulate gyrus, BA24	2	34	15	2.85
	-4	31	15	2.79
Orbitofrontal gyrus	22	31	-9	3.14
Cuneus	8	-77	18	3.81
Posterior superior temporal sulcus	56	-32	9	3.8
Fusiform gyrus	-26	-56	-6	3.19
Striatum	20	19	-3	4.1

Table 4: Conjunction analysis: Anatomical specification, Talairach coordinates (x,y,z) and maximal Z-scores of significantly activated voxels for biased vs. balanced states: the first divergents of a solid internal model vs. the first divergents of a weak internal model and the last divergents of a weak internal model vs. and the last divergents a solid internal model.

Localisation	Talairach coordinates			Z-values, local maxima
	x	y	z	
Frontal pole, BA10	6	43	3	2.40
	-4	49	3	2.91
Anterior cingulate gyrus, BA24	2	31	15	3.54
	2	34	-3	2.16
	-4	28	0	3.89
Cuneus	-2	-71	21	2.83
Fusiform gyrus	16	-53	-6	2.46
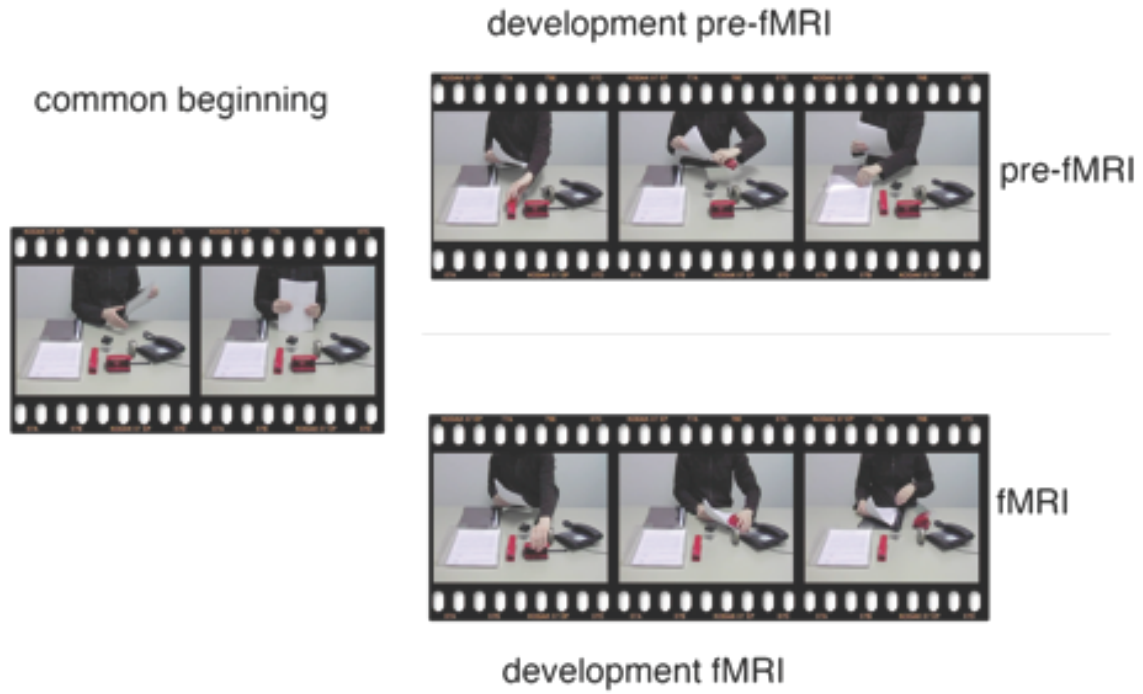
Figure 1: The initial version that was displayed previous to the fMRI and the divergent version that was displayed during the fMRI had a common beginning, i.e. they started with the same action step(s).
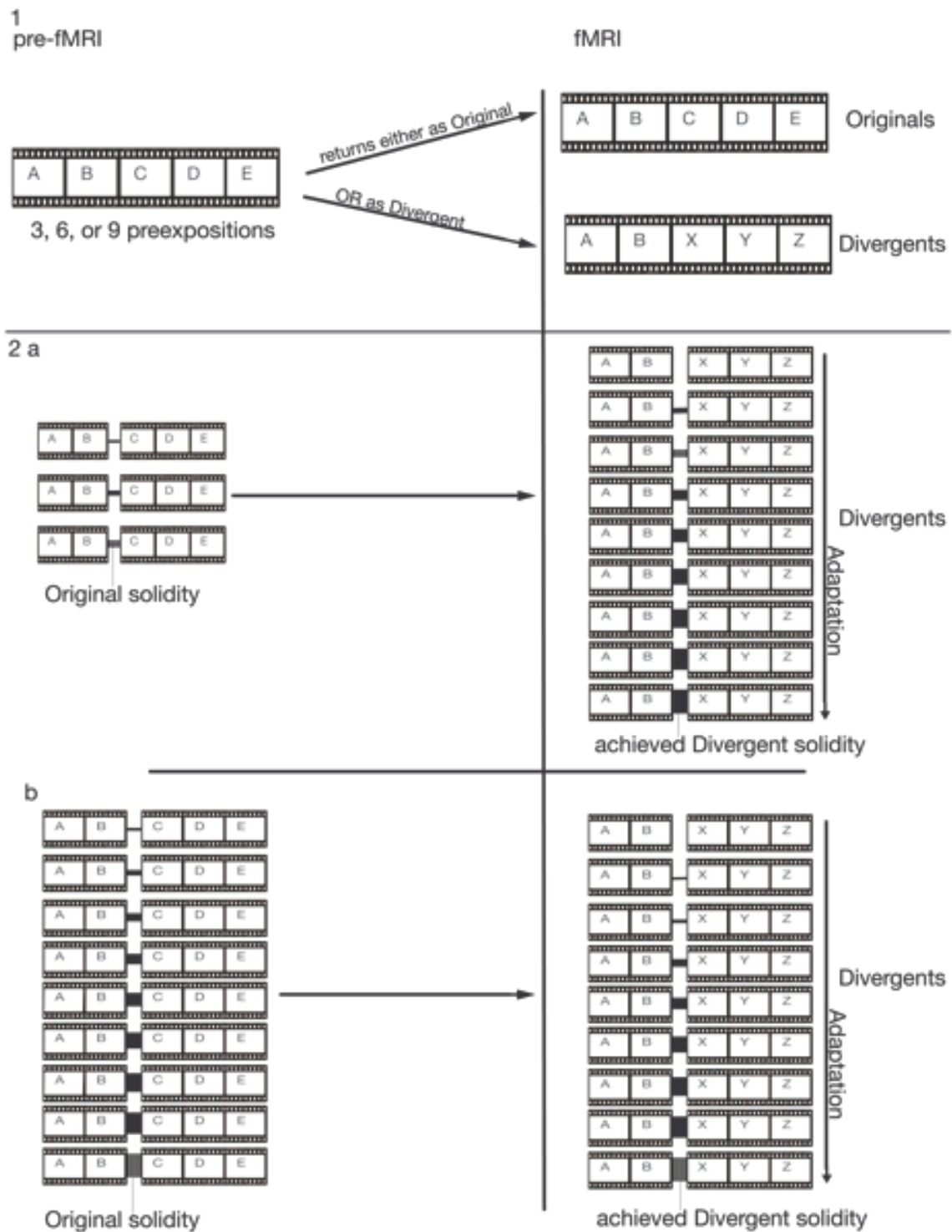
Figure 2: Abstract representation of the script-structure. Letters refer to action steps. 1) Movies were pre-exposed 3, 6, or 9 times in one version. A third of the movies reappeared in the fMRI in the same version as before 'original'. Another third appeared in a 'divergent' version. This version started exactly as the original version had, but developed differently thereafter. 2 a) Movies that were pre-exposed 3 times returned 9 times as divergents during the fMRI. Strength of the indicated link reflects solidity; Only the solidity of the transition of importance is indicated; each

transition has the same assumed solidity in the beginning. 2 b) Movies that were pre-exposed 9 times similarly returned 9 times as divergents during the fMRI. Again only the solidity of the relevant, i.e. later breached transition is graphically indicated.
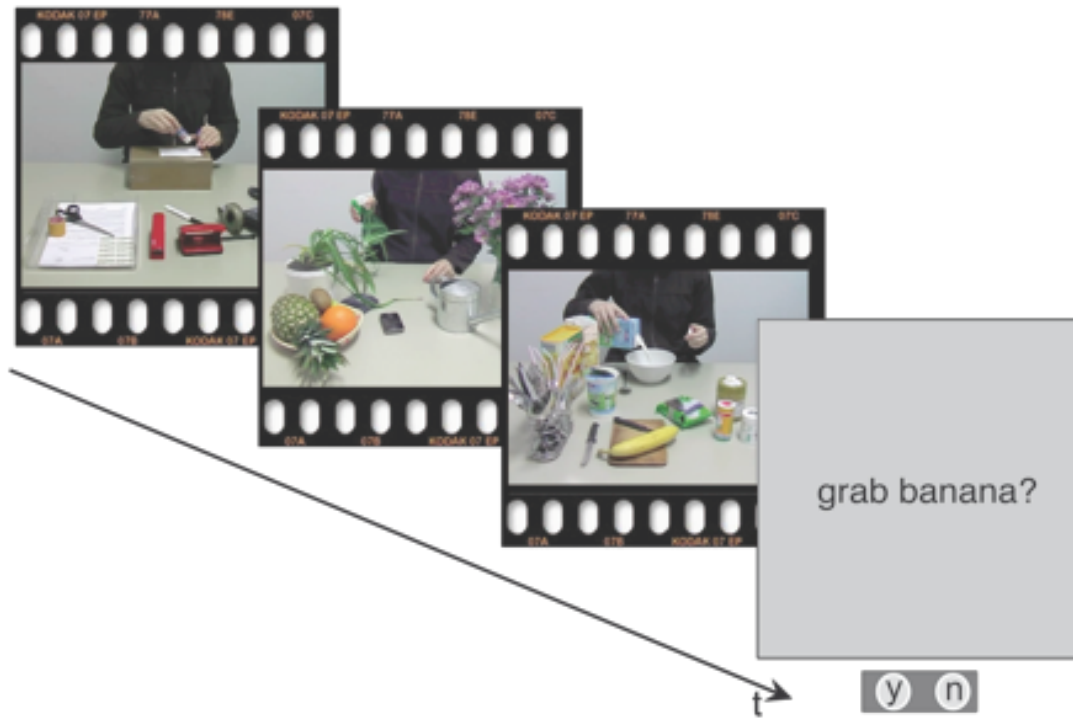
Figure 3: During the fMRI session, participants watched divergents and originals in a randomized fashion and had to answer content-related questions on average after every 5TH script.
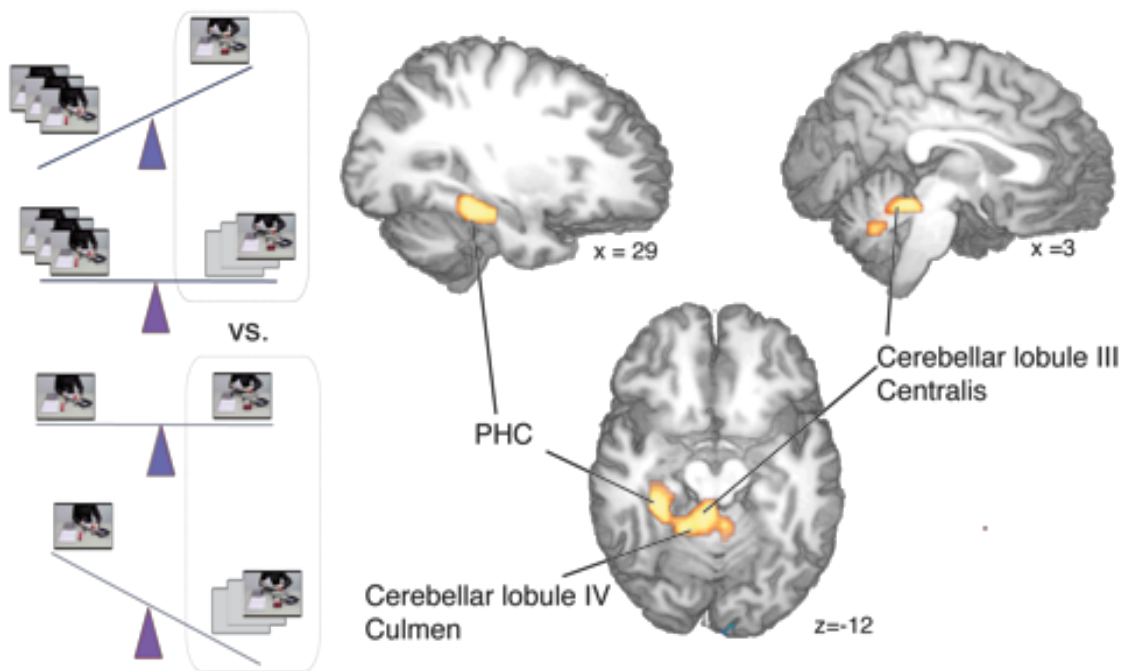
Figure 4: The effect of model solidity was calculated contrasting the 1st to 3rd and 7th to 9th iteration of scripts that had been pre-exposed nine times with the 1st to 3rd and 7th to 9th iteration of scripts that had been pre-exposed 3 times. PHC: Parahippocampal cortex.
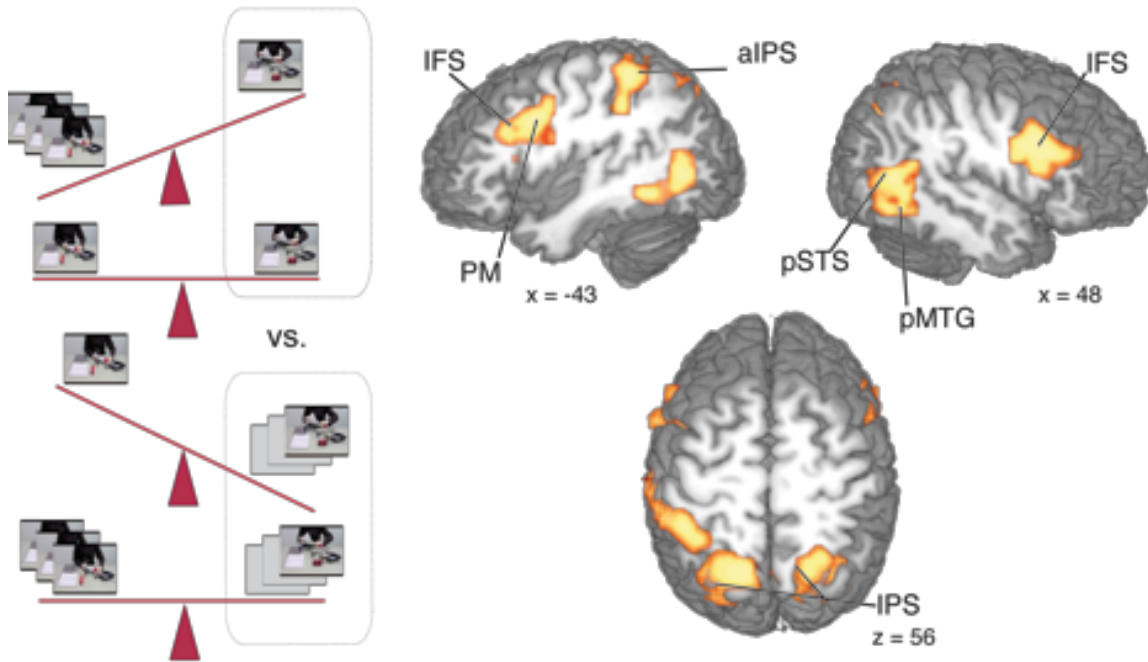
Figure 5: The effect of model adaptation effect was calculated contrasting the 1ˢᵗ to 3ʳᵈ iteration of scripts that had been pre-exposed either 3 or 9 times with the 7ᵗʰ to 9ᵗʰ iteration of scripts that had been pre-exposed either 3 or 9 times. (a)IPS: (Anterior) intraparietal sulcus; IFS: Inferior frontal sulcus; pMTG: Posterior middle temporal gyrus; pSTS: Posterior superior temporal sulcus; PM: Premotor cortex.
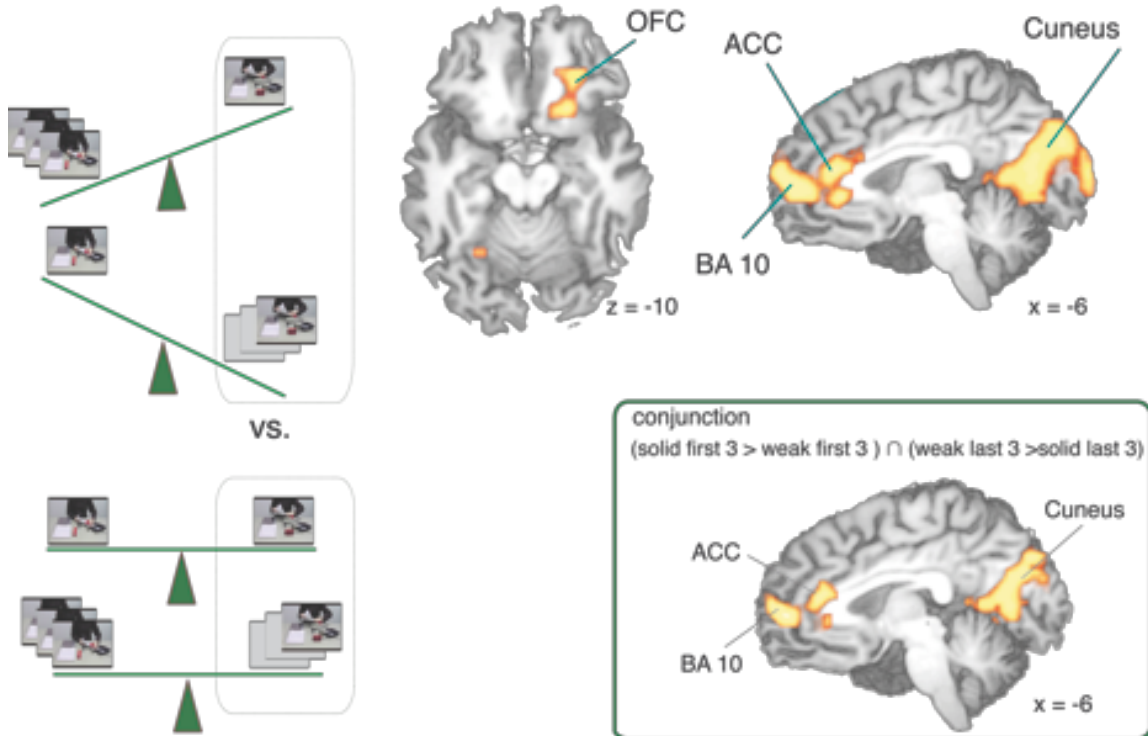
Figure 6: The biased vs. balanced effect was calculated contrasting the 1st to 3rd iteration of scripts that had been pre-exposed nine times and the 7th to 9th iteration of scripts that had been pre-exposed 3 times with the 1st to 3rd iteration of scripts that had been pre-exposed 3 times and 7th to 9th iteration of scripts that had been pre-exposed 9 times. ACC: Anterior cingulate cortex; BA 10: Brodmann Area 10; OFC: Orbitofrontal cortex.