Functional integration processes underlying the instruction-based learning of novel goal-directed behaviors

Hannes Ruge & Uta Wolfensteller

Neuroimaging Center and Institute of General Psychology, Biopsychology, and Methods of Psychology, Department of Psychology, Technische Universitaet Dresden, 01062 Germany

Corresponding author:
Hannes Ruge
Technische Universitaet Dresden
Fakultaet Mathematik und Naturwissenschaften, Fachrichtung Psychologie
01062 Dresden, Germany
Phone: +49 351 463-34695
Fax: +49 351 463-33522
Email: ruge@psychologie.tu-dresden.de

Abstract

How does the human brain translate symbolic instructions into overt behavior? Previous studies suggested that this process relies on a rapid control transition from lateral prefrontal cortex (LPFC) to anterior striatum (aSTR) and premotor cortex (PMC). The present fMRI study investigated whether the transfer from symbolic to pragmatic stimulus-response (S-R) rules relies on changes in the functional coupling among these and other areas and to which extent action goal representations might get integrated within this symbolic-pragmatic transfer. Goal integration processes were examined by manipulating the contingency between actions and differential outcomes (i.e. action goals). We observed a rapid strengthening of the functional coupling between the LPFC and the Basal Ganglia (aSTR and Putamen) and orbitofrontal cortex (OFC) as well as between the LPFC and the anterior dorsal PMC (pre-PMd), the anterior inferior parietal lobule (aIPL), and the posterior superior parietal lobule (pSPL). Importantly, only some of these functional integration processes were sensitive to the outcome contingency manipulation, including LPFC couplings with aSTR, OFC, aIPL, and pre-PMd. This suggests that the symbolic-pragmatic rule transfer is governed by principles of both, instrumental learning (increasingly tighter coupling between LPFC and aSTR/OFC) and ideomotor learning (increasingly tighter coupling between LPFC and aIPL/pre-PMd). By contrast, increased functional coupling between LPFC and Putamen was insensitive to outcome contingency possibly indicating an early stage of habit formation under instructed learning conditions.

A pivotal expression of human behavioral flexibility is the ability to quickly adopt novel behavioral rules that define the stimulus conditions under which certain actions will be successful (Duncan, 2001; Miller and Cohen, 2001; Ruge et al., 2011; Sakai, 2008; Woolgar et al., 2011). Humans can adopt and behaviorally implement such novel stimulus-response (S-R) rules instantaneously if explicitly instructed. It is important to note that any instructed behavior is essentially goal-directed provided that responding correctly as instructed is experienced as an intrinsically incentive outcome (O) – which seems especially relevant early in practice (Ashby et al., 2010; Killcross and Coutureau, 2003; Wolfensteller and Ruge, 2012; Wood and Neal, 2007). Although the processes that mediate the implementation of explicitly instructed behavioral rules are central to executive control function, research has been surprisingly scarce as already noted more than a decade ago (Monsell, 1996). Previous neuroscience research in humans has examined the acquisition of novel behavioral rules mostly by means of feedback-driven trial-and-error learning procedures in the tradition of instrumental conditioning research in animals who cannot easily be symbolically instructed (e.g., Brovelli et al., 2011; Frank and Badre, 2012; Mattfeld and Stark, 2011). Only recently has instruction-based learning started garnering broader scientific interest (Cohen-Kdoshay and Meiran, 2009; Waszak et al., 2008; Wenke et al., 2007), especially also in the human cognitive neuroscience domain (Cole et al., 2010; Cole et al., 2012; Doll et al., 2009; Dumontheil et al., 2011; Hartstra et al., 2011; Ramamoorthy and Verguts, 2012; Stocco et al., in press; Walsh and Anderson, 2011; Wolfensteller and Ruge, 2012). Yet, some of the basic brain mechanisms underlying the transfer from symbolically instructed rules into overt behavior are not fully understood.

In a previous fMRI study we observed a rapid activation *decline* in lateral prefrontal cortex (LPFC) and anterior cingulate cortex (ACC) during initial practice of newly instructed S-R rules which was paralleled by a rapid activation *increase* in the premotor cortex (PMC) and in the anterior striatum (aSTR) centered on the caudate head (Ruge and Wolfensteller, 2010). These and similar findings (Dumontheil et al., 2011; Hartstra et al., 2011; Stocco et al., in press), indicate a rapid transition of behavioral control from symbolic (LPFC) to pragmatic (aSTR, PMC) rule representation possibly accompanied by performance monitoring processes mediated by the ACC (e.g., Walton et al., 2004).

Based on these previous observations, the present study aimed to clarify two inter-related issues that have not been addressed so far. First, it remains unknown whether the rapid symbolic-pragmatic rule transfer might evolve through changing functional couplings between areas like LPFC and aSTR as has recently been suggested on theoretical grounds (Ramamoorthy and Verguts, 2012). Second, we aimed to further specify the functional role of the aSTR during instruction-based learning as signified by its increasing engagement during the symbolic-pragmatic rule transfer. Specifically, direct evidence is lacking as to whether the increasing engagement of aSTR reflects the transfer and

initial consolidation of goal-directed action representations or rather an early stage of habit formation (Brovelli et al., 2011; Seger and Spiering, 2011). As mentioned at the outset, novel behavior based on instructed S-R rules is intrinsically goal-directed rather than being an automatic 'reflex' triggered by the antecedent stimulus. In particular, as a prerequisite for correct performance, subjects are assumed to code a stimulus-bound action in terms of the success it yields and the potential failure it prevents. Under standard experimental conditions, this goal information is rather unspecific. That is, when different S-R associations are concurrently instructed they are all intrinsically and indistinguishably bound to one common incentive outcome (i.e., success or reward). To selectively manipulate the postulated integration of goal representations during the symbolic-pragmatic transfer, we adopted the 'differential outcome' rationale (Trapold, 1970). Specifically, we varied the amount of additional discriminative goal information attached to each distinct S-R link. Accordingly, newly instructed S-R mappings were behaviorally implemented under two outcome conditions. In the *contingent outcome* condition, additional differential outcomes (here different colors) predictably followed a correct response. In the *random outcom*e condition, the same differential outcomes unpredictably followed a correct response. Importantly, only the contingently paired but not the randomly paired 'differential outcomes' could be integrated with the intrinsic common outcome (i.e., success) for retrieving the action that is required in the presence of a particular antecedent stimulus. Similar manipulations have been shown to enhance the feedback-driven learning of novel S-R mappings (for a recent fMRI study, see Noonan et al., 2011).This basic observation is typically thought to indicate the relevance of outcome anticipation for instrumental learning (Colwill and Rescorla, 1990; Trapold, 1970; Urcuioli, 2005) – with outcome anticipation being one of the necessary conditions for goal-directed behavior (Dickinson and Shanks, 1995; Hommel et al., 2001).

Material and methods

Subjects

Seventy human participants took part in this study (mean age = 24; age range: 19-32; 47 females, 23 males). This comparably large number of subjects was collected to ensure sufficient statistical power for the functional connectivity analysis based on the psychophysiological interaction (PPI) approach (O'Reilly et al., 2012). This seems particularly warranted for the planned between-subjects comparison of PPI effects as a function of outcome contingency. The contingent outcome group comprised 9 male and 26 female participants (mean age 23; age range 19-32). The random outcome group comprised 14 male and 21 female participants (mean age 25; age range 19-29). The experimental protocol was approved by the Ethics Committee of the Technische Universitaet Dresden and conformed to the World Medical Association's Declaration of Helsinki. All participants

gave written informed consent prior to taking part in the experiment and were paid € 8 per hour for their participation or received course credit.

Task design

The same experimental procedure as in Ruge and Wolfensteller (2010) was employed to study the brain activation dynamics associated with instruction-based S-R learning (see Fig. 1). The only difference was that in one group of 35 subjects correct responses were consistently followed by differential outcomes (i.e., each correct response produced a distinct color depending on the response preceding stimulus) whereas in another group of 35 subjects the same outcomes were randomly selected. Across the experiment, 21 novel stimulus-response-outcome mappings had to be learned consecutively. Each such mapping was composed of 4 different visual stimuli, 2 different responses, and 4 different colored outcomes. The first mapping was administered outside the scanner for familiarization with the task.

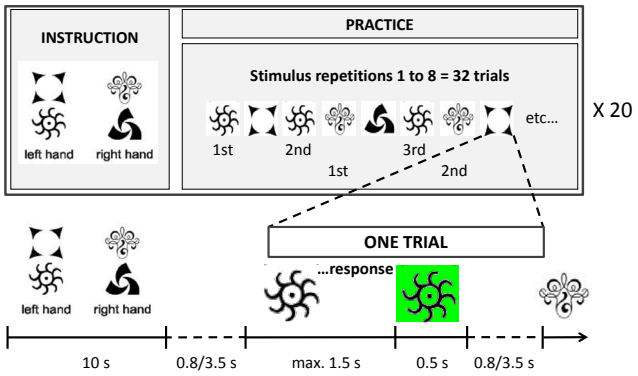

Fig.1 Experimental procedure used for investigating instruction-based learning. Practice-related changes in behavior and brain activation were analyzed in terms of correctly implemented stimulus repetitions (SRep) 1 through 8. Subjects were scanned while learning 20 different novel S-R mappings. For two different groups of subjects correct responses were either contingently or randomly followed by one of four different outcome colors.

For each mapping a unique set of 4 stimuli was used and a novel (though partly overlapping) set of 4 colors (out of 12 different colors) was used (Fig 1). The 4:2 stimulus-response mappings had to be memorized across a 10 seconds instruction period, followed by a practice phase consisting of 32 correctly performed implementation trials (8 randomly ordered repetitions of each stimulus-response pair) with either contingent or random color outcomes following a correct response which were displayed for 500 ms. Each implementation trial started with the presentation of a stimulus which was displayed until the participant's response or until the maximum response time window of 1500 ms had elapsed. Incorrect responses were followed by gray color and the trial was immediately repeated. Subjects were explicitly told that gray color indicated an erroneous response prior to the

start of the experiment and hence they knew that the next (repeated) trial would require exactly the opposite response. To allow BOLD response estimates relative to a baseline, the inter-trial interval was either 0.8 or 3.5 seconds and was randomly selected. The total experiment duration was approximately 35 minutes, depending on mean response speed and the number of erroneous responses. Practice-related brain activation dynamics and performance improvement were analyzed in terms of correctly performed stimulus repetitions (SRep) 1 through 8. This is a particularly important prerequisite for the fMRI analysis to make sure that model regressors capturing the incremental progress of practice are sufficiently uncorrelated.

MR imaging Procedure

Whole-brain images were acquired on a Siemens 3 Tesla whole-body Trio System (Erlangen, Germany) with a 16 channel circularly polarized head coil. Headphones (NordicNeuroLab) and earplugs dampened scanner noise. Both structural and functional images were acquired for each participant. High-resolution structural images (1.0 mm x 1.0 mm x 1.0 mm) were acquired using an MP-RAGE T1-weighted sequence (TR=1900 ms, TE=2.26 ms, TI=900 ms, flip=9º). Functional images were acquired using a gradient echo planar sequence (TR = 2000 ms, TE = 30 ms, flip = 80º). Each volume contained 26, 5.0 mm thick slices (in-plane resolution 4.0 mm x 4.0 mm). The experiment was controlled by Presentation 12.0 software (Neurobehavioral Systems) running on a Windows-XP PC. Stimuli were projected to participants via Visuastim digital goggles (Resonance Technology, Inc.; Northridge, USA) simulating a viewing distance of 100 cm. A fiber-optic, light-sensitive key press was used to record participants' behavioral responses.

FMRI preprocessing

The empirical data set was analyzed with SPM8 running under MATLAB 7.12.0.635. Preprocessing included slice-time correction, rigid body movement correction (3 translation, 3 rotation parameters), normalization of the functional images by directly registering the mean functional image to the standard MNI EPI template image provided by SPM8 (the resulting interpolated spatial resolution was 3x3x3 mm), and smoothing of the functional images (Gaussian Kernel, FWHM = 8 mm). During model estimation a temporal high-pass filter with a cutoff frequency of 1/128 Hz was applied.

FMRI analysis

Model specification. Event-related BOLD activation in each voxel was estimated via the General Linear Model (GLM) approach and equivalent to Ruge & Wolfensteller (2010). Model regressors were created by convolving neural input functions for the different event types with the assumed canonical hemodynamic response function used by SPM8. We separately modeled BOLD activation associated with the 10 s instruction phase and with the subsequent symbolic-pragmatic transfer phase. Instruction-related activation was modeled by including two regressors both synchronized to the onset of the instruction. One regressor was constructed to capture BOLD activation associated with transient neural activity at the start of the instruction phase (duration parameter = 0). The second regressor was constructed to capture BOLD activation associated with a potential sustained neural activity component throughout the instruction phase (duration parameter = 10). Practice-related activation dynamics were modeled in terms of correctly implemented stimulus repetitions. Accordingly, 8 event-related regressors were included for stimulus repetitions 1 through 8. One additional regressor was included for error trials. Since the primary focus of this study was to investigate the symbolic-pragmatic transfer during the practice phase, we only report analyses based on the BOLD response estimates for stimulus repetitions 1 through 8.

PPI analysis. Practice-related changes in the functional coupling between brain regions were computed based on the psycho-physiological interaction (PPI) approach as implemented in SPM8 (Friston et al., 1997; Gitelman et al., 2003). This multivariate method determines interregional functional connectivity by means of significant changes in regression coefficients for predicting the BOLD signal timecourse in 'target' voxels anywhere in the brain volume from the BOLD signal timecourse in a pre-specified 'seed' voxel, depending on an experimentally manipulated 'psychological' variable (here early vs. late in practice following a novel S-R instruction). The early practice phase was defined as the mean BOLD response to SRep1 and SRep2 whereas late practice was defined as the mean BOLD response to SRep7 and SRep8. Seed regions (defined as 6 mm radius spheres centered on a single seed voxel) were selected within the LPFC and within the ACC and corresponding PPI regressions were computed with all other target voxels within the measured brain volume. To qualify as a candidate seed within the large cortical structures LPFC and ACC, a region was required to exhibit an activation *decline* across practice, hence suggesting a particularly dominant role early in practice and thus being especially important for the initial symbolic-pragmatic rule transfer. Consistent with our previously reported results (Ruge & Wolfensteller, 2010), a univariate analysis of practice-related BOLD dynamics defined as the activation difference between early vs. late practice ($p<.05$ FWE-corrected) revealed that large parts of LPFC and medial PFC

exhibited rapidly decreasing activation across the first few successful implementations of a particular S-R association (Fig. 3A). Two local activation maxima for left and right posterior LPFC (pLPFC) in the vicinity of the posterior inferior frontal sulcus (IFS) were selected as PPI seeds (MNI coordinates *-51 8 31* and *47 8 28*). We chose these two pIFS regions – instead of more anteriorly located LPFC regions with similar activation profiles – to ensure closest correspondence with previous studies that examined similar forms of symbolic-pragmatic rule transfer with greatest activation overlap across studies in the pIFS (Dumontheil et al., 2011; Hartstra et al., 2011; Ruge and Wolfensteller, 2010)[1]. The ACC seed was determined as a local maximum (MNI coordinates *11 29 36*) within a fronto-medial activation cluster also including more dorsal regions outside the anterior cingulate gyrus.

PPI significance was assessed based on family-wise-error (FWE) correction for multiple comparisons at *p<.05* either applied on the whole brain level or selectively applied for more confined brain volumes based on a-priori functional-anatomical assumptions. Anatomically defined ROIs were the Basal Ganglia (including Caudate, Putamen, and Pallidum), the OFC (defined as BA11 and BA47), and the PMC (defined as BA6). Anatomical information was taken from the automatic anatomic labeling atlas (AAL) for the Basal Ganglia ROI (Tzourio-Mazoyer et al., 2002) and from the Brodmann parcellation used in the MRICRON software (Rorden et al., 2007). Note that the OFC was not in the set of regions implicated in the symbolic-pragmatic transfer in our original study (Ruge and Wolfensteller, 2010). Nevertheless, we included this region as an ROI based on previous studies strongly implicating this brain region in goal-directed action control (e.g., Noonan et al., 2012; Valentin et al., 2007).

The PPI analysis proceeded in two consecutive steps. First, we computed the overall PPI contrast collapsed across both contingency groups (i.e. including all 70 subjects) for each of the three seed regions. This PPI contrast is suited to reveal which areas are generally involved in the symbolic-pragmatic rule transfer. Second, we compared PPI contrasts for the contingent outcome group relative to the random outcome group. For this group comparison, we first applied the common significance criteria defined above (FWE-corrected p<.05) to reveal group-related PPI effects anywhere in the brain or anywhere within the anatomically pre-defined ROIs. However, these thresholds are inappropriately conservative to determine more specifically whether activations revealed in the overall PPI analysis performed across all 70 subjects might be additionally modulated by the outcome contingency manipulation. This is especially true as the overall PPI analysis was naturally most sensitive to activation effects *common* to both contingency groups[2]. Thus, in a second

[1] Note that more anteriorly located seeds within the LPFC yielded qualitatively similar results.
[2] The alternative analysis rationale would be to first determine ROIs based on PPIs selectively computed for the contingent outcome group. In a second step, PPI activations within these ROIs could then be compared

analysis step, we additionally assessed the PPI group comparison for ROIs defined as spheres of 6 mm radius centered on the peak voxels yielded by the overall PPI analysis, with all reported activations p<.05 FWE-corrected for these small-volumes.

Results

Behavioral performance

Behavioral performance was assessed with two separate two-way ANOVAs for response times and error rates each including the within-subjects factor stimulus repetition (SRep1 through SRep8) and the between-subjects factor outcome contingency group (contingent vs. random). Not surprisingly, and replicating our previous findings, the ANOVA revealed a rapid practice-related decrease in response times ($F_{7,476}$=89.82; p(F)<.001; Greenhouse-Geisser corrected) and error rates ($F_{7,476}$=5.84; p(F)<.001; Greenhouse-Geisser corrected). Notably, average error rates started from below 5% which excludes a strong initial contribution of trial-and-error learning given a baseline error rate of 50% expected for random response selection. The more interesting question, however, was whether outcome contingency would affect the behavioral learning curve (Fig. 2). There was no clear-cut evidence for that, but there is a tendency for a practice-related relative response slowing for the contingent outcome group relative to the random outcome group (linear contrast $F_{1,68}$=3.47; p(F)<.07 and a significant difference for SRep 6 through SRep 8 at p(t)<.05). There was no such trend for error rates. In line with the response slowing for contingent relative to random outcomes in the present study, we also observed relative response slowing for a contingent outcome condition relative to a common outcome condition in a similarly early phase of practice (Ruge et al., 2012). While contingent response outcome conditions seem to be consistently associated with response slowing rather than response speeding, the interpretation of this observation might appear somewhat puzzling. In particular, one might expect that anticipated contingent outcomes serve as an additional retrieval cue for the currently correct response which should result in faster responding. That the opposite is true might be due to the fact that we investigated a very early phase of practice in which action selection takes place in a highly controlled intentional mode in which outcome-based action retrieval and directly stimulus-based action retrieval are not operating perfectly parallel. Consequently, a response decision might be delayed until outcome-based retrieval processes have been fully integrated. For a more extensive discussion of this issue, see also Ruge & Wolfensteller (2012). Finally, it should be noted that using similar *rapid* learning procedures, we could repeatedly demonstrate that contingently paired color outcomes are getting quickly associated with the preceding responses, as revealed by re-using previous color outcomes as compatible vs. incompatible

---

between the contingent condition and the random condition. We decided against this option to avoid a bias towards false-positive PPI activations that are stronger for the contingent group than for the random group.

response primes or as compatible vs. incompatible imperative stimuli in a subsequent test phase (Ruge et al., 2012; Wolfensteller and Ruge, 2011). Together, these converging behavioral results suggest that contingently paired differential outcomes are getting rapidly integrated into action representations.

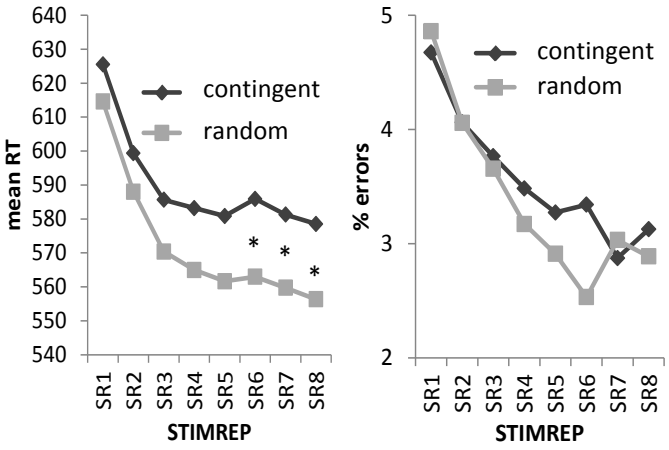

Fig.2 Behavioral results (mean response times (RT) and percent errors) as a function of stimulus repetition (STIMREP) and outcome contingency. Asterisks (*) indicate significant t-tests at p<.05. Note that the linear contrast on RT data across all STIMREP levels was only marginally significant with p<.07.

Univariate fMRI results

The univariate analysis of practice-related BOLD dynamics was primarily used to define the seed regions for the PPI analysis (see Methods). Since the results essentially replicate earlier findings (Ruge and Wolfensteller, 2010), they are only briefly summarized graphically in Fig 3A. Significant activation increases (*p<.05* FWE-corrected) defined as the activation difference late (SRep7/8) – early (SRep1/2) in practice were found most prominently in the SMA, the lateral PMC, and in the anterior striatum (including caudate head and ventral striatum). Significant activation decreases (*p<.05* FWE-corrected) defined as the activation difference early (SRep1/2) – late (SRep7/8) in practice were found most prominently in large parts of the LPFC, the pre-SMA/ACC, pre-PMd, SPL, and IPL.

More importantly in the present study context, the contingency groups did *not* differ significantly in the univariate voxel-wise analysis of practice-related BOLD activation dynamics (defined as SRep78 – SRep12 or SRep12 – SRep78). Significant effects were found neither on the whole-brain level, nor for anatomically defined ROIs, nor for any of the regions revealed by the PPI analysis (6 mm spherical ROIs centered on PPI peak voxels). There was also no main effect of contingency group when averaging across all SRep levels. While we did not expect such a complete absence of univariate contingency-related differences in brain activation, it is becoming increasing clear that at least under

certain circumstances condition-specific mental processes can be exclusively expressed in condition-specific changes in the *coupling* between distant brain regions (Camara et al., 2009; Cross and Iacoboni, 2011).

PPI results

Functional connectivity pattern collapsed across both contingency groups

First, we focused on the two pre-defined ROIs basal ganglia and orbitofrontal cortex. Collapsed across the two contingency groups, both pLPFC seed regions exhibited an overall *increase* in functional connectivity with striatal and orbitofrontal subregions across the course of practice (Tab. 1). Striatal PPI effects included the anterior striatum (including caudate head and ventral striatum), the anterior dorsal putamen (adPUT), the putamen tail (PUTtail), and the pallidum (PAL). The PPI effects for the adPUT and PUTtail were more pronounced for the right pLPFC seed. PPI effects within the OFC were found in central and lateral OFC regions and were more pronounced for the left pLPFC seed. There were no significant connectivity *decreases* for these ROIs.

Additionally, we examined the hypothesis that learning within basal ganglia might not only depend on the just confirmed interaction with the LPFC but also on performance monitoring processes mediated by the ACC. However, the separate PPI analysis with the ACC as seed region did *not* reveal significant PPI effects within the basal ganglia. Also, there were no significant PPI effects in the vicinity of the lateral and central OFC regions that were identified for the LPFC seeds. Instead the ACC seed region displayed increased connectivity with the dorsal-most portion of the medial OFC (MNI -*18 38 -5; 18 38 1*).

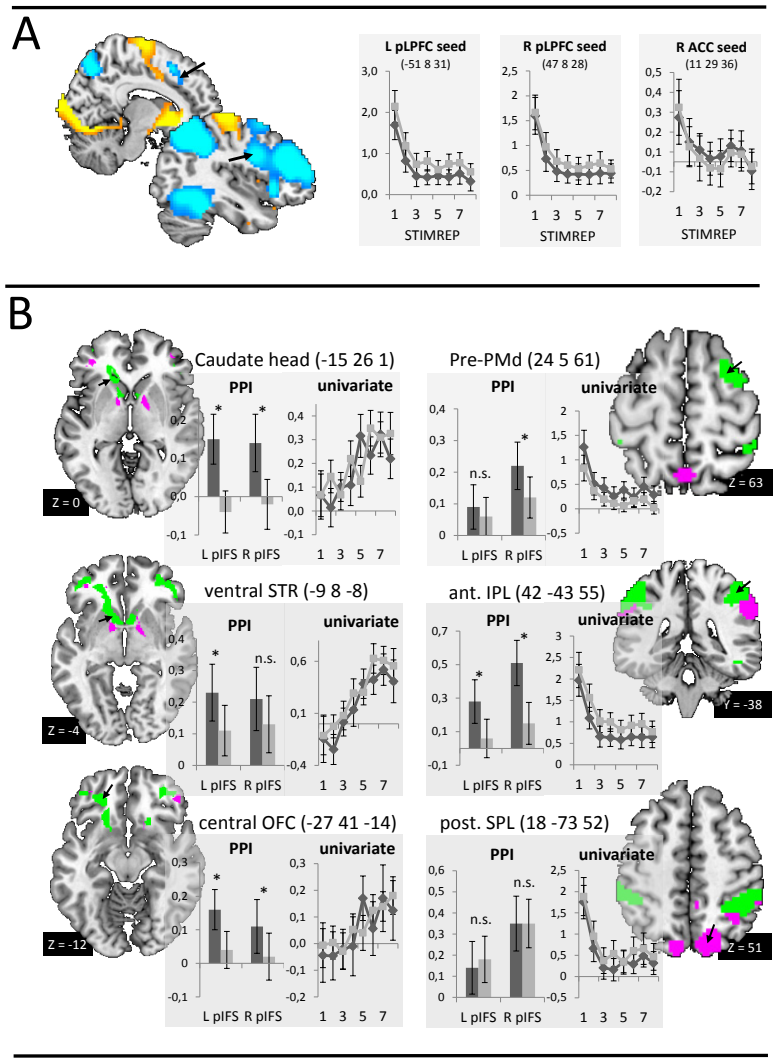

Fig 3. (A) Univariate analysis results as basis for PPI seed selection. Orange/yellow indicates significant increase in practice-related BOLD activation. Blue indicates significant decrease in practice-related BOLD activation. All activations are thresholded at p<.05, FWE-corrected for whole-brain volume. Three areas were selected as seeds for the PPI analysis (L pLPFC, R pLPFC, and R ACC). The black arrows superimposed on the activation maps indicate the left pLPFC and the ACC. The three line graphs depict for each seed region the univariate BOLD estimates for the eight consecutive stimulus repetitions (STIMREP) separately for the contingent outcome condition (dark grey) and for the random outcome condition (light grey). Error bars represent the 90% confidence interval. Note that the contingency conditions were neither significantly different for the three seed regions, nor for any other region in the brain. (B) Summary of the PPI results. Pink indicates areas that exhibit an increased functional coupling across practice with one or both pLPFC seeds regions (for display purposes, p<.001, uncorrected) which was not significantly different for the contingency conditions (for display purposes, p>.05, uncorrected). Green indicates areas that exhibit a more strongly increased practice-related coupling with one or both pLPFC seeds regions for the contingent outcome condition than for the random outcome condition. This includes voxels that exhibited a significant PPI activation for the contingent outcome condition (for display purposes, p<.001, uncorrected) and that showed a stronger PPI effect for the contingent condition than for the random condition (for display purposes, p<.05, uncorrected). For threshold criteria used to actually identify significant voxels, see main text. The bar graphs depict the PPI estimates for representative regions (indicated by black arrows in the activation maps), separately for left and right pLPFC seeds and separately for the contingent outcome condition (dark grey) and for the random outcome condition (light grey). Error bars represent the 90% confidence interval. A star (*) indicates a significant difference of PPI estimates between the two contingency conditions according to the criteria defined in the main text. The line graphs depict for the same representative regions the univariate BOLD estimates across the eight stimulus repetitions separately for the contingent outcome condition (dark grey) and for the random outcome condition (light grey). Error bars represent the 90% confidence interval.

Considering the whole brain volume (Tab. 2), collapsed across the two contingency groups, both pLPFC seed regions additionally displayed significant connectivity increases with the anterior middle frontal gyrus (aMFG), anterior inferior parietal lobule (aIPL), posterior superior parietal lobule (pSPL), posterior inferior temporal gyrus (pITG), and occipital cortex (OCC). Moreover, additional

predominantly right hemispheric areas showed a significant connectivity increase selectively for the right pLPFC seed, including the right anterior dorsal PMC (pre-PMd), the right inferior frontal gyrus (IFG), the right anterior insula (aINS), and the right middle hippocampus (mHIPP). Again, there were no significant connectivity *decreases* on the whole-brain level. The ACC-seeded PPI analysis revealed only two significant activations roughly overlapping with regions reported for the pLPFC seeds, including the right aMFG (MNI *36 15 19*) and the occipital cortex. Additionally, connectivity with an ACC region (MNI *9 38 22*) anterior to the ACC seed region was significantly increased.

Different from our initial hypothesis, the practice-related increase in connectivity between LPFC and PMC was *not* centered on the more posteriorly located PMC section. In the univariate analysis this region showed an activation increase with increasing practice and it is commonly thought to be tightly associated with motor implementation processes (Picard and Strick, 2001). Instead, the significant PPI between the right pLPFC and the PMC was confined to the anterior portion of the PMd (pre-PMd), which exhibited a decrease in activation across practice (in the univariate analysis). To explore the hypothesis that the pre-PMd might serve as a relay station that indirectly connects LPFC with posterior PMC, we computed an additional PPI with pre-PMd as the seed region (MNI *24 5 61*). However, even for a relaxed FWE correction restricted to the anatomically defined PMC ROI, no significant PPI effects were revealed.

Modulation of functional connectivity patterns by outcome contingency

To examine the influence of differential response outcomes on practice-related connectivity changes, we performed a between-subjects comparison (contingent outcomes vs. random outcomes) of PPIs for left and right pLPFC seeds used in the overall PPI analysis. In a first analysis step, we assessed PPI results according to the strict significance criteria defined above (i.e., $p<.05$ FWE-corrected for whole brain volume or for anatomically defined ROIs). According to these criteria, a significant group difference was restricted to the left caudate head (MNI coordinates *-15 26 1; t=3.69; p(t)<.03*). More specifically, the caudate head exhibited a significantly stronger practice-related connectivity increase with pLPFC for the contingent outcome group as compared to the random outcome group (Fig. 3B). To explore whether this effect would be replicated also for the *right* pLPFC seed at a lowered threshold, we defined a 6 mm radius sphere centered on the caudate head peak voxel identified for the left pLPFC seed. Within this sphere, we identified a significant contingency-related PPI effect (MNI -12 29 1; t=2.63; p(t)<.02) also when using the right pLPFC as PPI seed region.

A second analysis step was tailored to specifically assess whether significant practice-related connectivity increases in the overall PPI analysis might be modulated by the PPI group comparison. To this end, we performed the PPI group comparison for ROIs defined as spheres of 6 mm radius

centered on the peak voxels yielded by the overall PPI analysis, with all reported activations p<.05 FWE-corrected for these small-volumes (see Methods). For a subset of the areas that exhibited an overall increase in practice-related connectivity with the posterior LPFC we observed an amplified connectivity increase for the consistent group as compared to the random group (Fig. 3B; Tab. 1 and 2). Specifically, both pLPFC seed regions displayed an amplified practice-related connectivity increase with bilateral OFC (central and lateral) for the contingent outcome group relative to the random outcome group. For other target regions this additional amplification was differentially pronounced depending on the hemisphere of the seed. Practice-related increase of functional connectivity with the *left* pLPFC was preferentially amplified for the ventral striatum (vSTR) and the posterior inferior temporal gyrus (pITG). By contrast, practice-related increase of functional connectivity with the *right* pLPFC was preferentially amplified for the pre-PMd and the anterior inferior parietal lobule (aIPL).

Notably, some of the areas that showed an increased PPI effect for the contingent outcome group relative to the random outcome group did not reveal a significant PPI effect for the random outcome group even at the most lenient threshold (p<.05 uncorrected), including the caudate head and the left and right central OFC. All other areas did show a significant PPI effect also in the random outcome group, including the ventral striatum (L/R LPFC seed), the right and left lateral OFC (L/R LPFC seed), the pre-PMd (R LPFC seed), the left and right aIPL (R LPFC seed), and the pITG (L/R LPFC seeds).

Finally, for a number of other regions the overall practice-related connectivity increase with the pLPFC seeds was *not* additionally modulated by outcome contingency. Most interestingly, this includes several basal ganglia subregions (Fig. 4): the middle part of the pallidum (mPAL), the anterior dorsal putamen (adPUT), and the putamen tail (PUTtail). Also several cortical regions, including the visual cortex, the posterior superior parietal lobule (pSPL), the inferior frontal gyrus (IFG), the anterior insula (aINS), and posterior sections of the supramarginal gyrus (SMG).

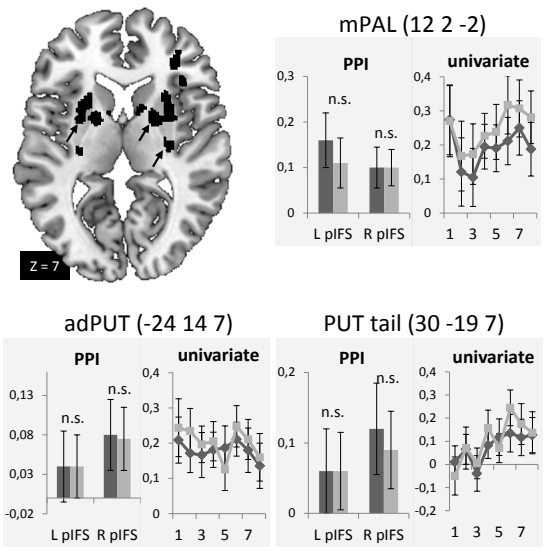

Fig. 4 Summary of PPI results for the basal ganglia ROI highlighting areas that were insensitive to the outcome contingency manipulation. Black indicates areas that exhibit an increased functional coupling across practice with one or both pLPFC seeds regions (for display purposes, p<.001, uncorrected) which was not significantly different for the contingency conditions (for display purposes, p>.05, uncorrected). The bar graphs depict the PPI estimates for representative regions (indicated by black arrows in the activation maps), separately for left and right pLPFC seeds and separately for the contingent outcome condition (dark grey) and for the random outcome condition (light grey). Error bars represent the 90% confidence interval. The line graphs depict for the same representative regions the univariate BOLD estimates across the eight stimulus repetitions separately for the contingent outcome condition (dark grey) and for the random outcome condition (light grey). Error bars represent the 90% confidence interval.

Discussion

The present data suggest that the symbolic-pragmatic transfer of newly instructed S-R rules evolves through an *increased functional coupling* particularly between the LPFC and predicted areas such as the aSTR and the OFC within the first eight rule implementation trials after instruction. Moreover, increased functional coupling was observed between LPFC and a set of other cortical areas that were less strongly expected in terms of instrumental learning principles. As detailed below, the involvement of these latter brain areas is consistent with ideomotor learning principles. It is hence clear that the LPFC, despite its generally decreasing activation profile, continues to play an important role for implementing the instructed rules by increasingly collaborating with other regions. This general observation refutes the alternative view that LPFC might be especially important only early in practice for driving the formation of pragmatic rule representations within areas such as aSTR, OFC and PMC. According to that view, functional coupling with LPFC should have been especially strong early in practice and increasingly weaker later in practice reflecting a decreasing need for teaching signals generated by these areas.

In stark contrast to the LPFC, and against our expectation, the ACC did not show a change in coupling with striatal areas during the practice of newly instructed rules. This refutes the hypothesis that aSTR might learn the instructed rules indirectly (also) via ACC-based performance monitoring processes that evaluate just executed responses for accuracy. Hence, our finding is consistent with an earlier

study suggesting that the ACC does not seem to be involved in performance monitoring when response decisions are externally instructed as compared to situations in which responses are 'freely' chosen (Walton et al., 2004).

Clearly the PPI analysis alone does not provide strong evidence for the direction of information flow between the increasingly coupled areas. It is thus difficult to tell which area is increasingly driving the other. Yet, combined with the finding that LPFC started high and aSTR/ OFC started low in activation might tentatively suggest that the increasing coupling between both regions is driven by the LPFC. Following this line of reasoning, the interplay between LPFC and aSTR/ OFC in instruction-based learning – with an assumed leading role of the LPFC – might be decisively different from trial-and-error learning situations that are typical of the instrumental learning framework. In this latter case, single cell recordings suggest that there is no clear-cut leading role for either region for initial learning (Cromer et al., 2011) and under reversal learning conditions there is strong evidence that the anterior striatum even leads the LPFC (Pasupathy and Miller, 2005). Interestingly, however, recent trial-and-error learning data converge with the present results by showing that feedback-driven learning processes within the basal ganglia are under the influence of symbolic rule representations in the LPFC, be they generated internally for the purpose of hypothesis testing or conveyed externally via instruction (Doll et al., 2009; Frank and Badre, 2012; Li et al., 2011). Note however, that ultimately valid conclusions about the direction of information flow cannot be based on PPI analyses and hence this limitation calls for further research.

Overall functional connectivity pattern in relation to univariate activation dynamics

Notably, striatal and OFC areas that were increasingly coupled with LPFC seeds were at the same time *increasingly* activated across practice as revealed by the univariate analysis (see line graphs in Fig. 3B). By contrast, other areas such as pre-PMd, aIPL and pSPL identified on the whole-brain level and showing a similarly increased coupling with LPFC, were *decreasingly* activated across practice as revealed by the univariate analysis (see line graphs in Fig. 3B). These contrasting univariate activation dynamics imply different interpretations of the common practice-related connectivity increases with the LPFC seeds. On the one hand and consistent with our initial expectations, increased LPFC coupling with areas exhibiting *increasing* activation across practice might reflect increasingly shared control between LPFC symbolic representations and increasingly consolidated pragmatic rule representations within areas such as the aSTR and the OFC. This type of enhanced coupling is hence due to the increasingly better incorporation of *additional* brain regions and their specific functionality (i.e., regions that were not or only weakly involved early in practice). On the other hand, the rather unexpected increased LPFC coupling with areas exhibiting initially high and subsequently *decreasing* activation across practice (e.g., Pre-PMd, aIPL, pSPL) might indicate an increasingly better

coordination between LPFC and these areas which are, different from aSTR and OFC, massively engaged already from the outset of practice. The initially strong co-activation paralleled by a relatively weak coupling of seed and target regions might indicate a rather uncoordinated attempt to implement the newly instructed S-R rules. Hence, the lack of coordinated action might require a strong engagement of each separate region for compensation. Further, the increasingly better coordination as expressed by increasing functional coupling might imply a decreasing requirement for the engagement of each individual region on its own.

From instruction to goal-directed behavior - Instrumental learning mechanisms

An increasing number of neuroimaging studies investigating trial-and-error learning have confirmed the general relevance of anterior striatum and OFC for establishing goal-directed as opposed to habit-like behaviors (e.g., de Wit et al., 2009; Noonan et al., 2011; Noonan et al., 2010; O'Doherty, 2011; Tanaka et al., 2008; Tricomi et al., 2004; Valentin et al., 2007; Xue et al., 2008). Consistent with this notion, these areas were sensitive to the outcome contingency manipulation in the present study. Specifically, the increased practice-related coupling between the LPFC and aSTR/ OFC was amplified for contingent outcomes as compared to random outcomes. This finding is consistent with a recent fMRI study that suggests that instructed knowledge about action-outcome relations modulates the activation in reward-related areas associated with the outcome prediction error during trial-and-error learning. Importantly, this modulation seems to be mediated by the LPFC (Li et al., 2011). The important novel insight from the present results is that the outcome-sensitive dynamic interaction between LPFC and aSTR/ OFC is not limited to the feedback-driven learning of novel goal-directed behaviors based on outcome prediction error computations. Also, it seems noteworthy pointing out that these areas are typically associated with the processing of incentive outcomes while the differential outcomes used in the present study were completely neutral in this respect. It would be interesting to clarify whether the present findings might be additionally modulated by the size of incentive values attached to outcomes as might be implied by a recent behavioral study (Muhle-Karbe and Krebs, 2012).

Notably, the outcome-sensitive PPI pattern in anterior striatum and OFC was further qualified when taking into account the absolute size of the PPI effect in the random outcome group. Specifically, some sub-regions showed a relatively weak but significant PPI effect also in the random outcome group (ventral STR and lateral OFC). This is consistent with the notion that the presence of contingent differential outcomes implies a *relative* enhancement of generic instruction-based S-R learning mechanisms that are also taking place without contingent differential outcomes. For other areas there was virtually no PPI effect in the random outcome group (caudate head and central OFC)

suggesting that these areas are engaging in a learning-related coupling with the pLPFC only when differential action outcomes can be learned.

Importantly, other BG areas that have previously been implicated in habit learning, exhibited PPI effects *not* affected by outcome contingency in the present study. This includes putamen tail (Tricomi et al., 2009) and the anterior dorsal putamen (Brovelli et al., 2011). Although this finding seems at odds with the common belief that habit formation requires extensive practice (but see Seger and Spiering, 2011), our results converge with recent findings suggesting an early onset of habit formation within the anterior dorsal putamen at the border of associative and sensorimotor striatum under trial-and-error learning conditions (Brovelli et al., 2011).

From instruction to goal-directed behavior - Ideomotor learning mechansims

Importantly, contingent action outcomes amplified the practice-related functional coupling between LPFC and a set of other cortical regions, including aIPL, pre-PMd, and pITL. From the viewpoint of the instrumental conditioning framework these areas would typically not be expected to be related to the goal-directed control of behavior as they are not closely related to the brain's reward system. Yet, as detailed below, from the viewpoint of the ideomotor framework (de Wit and Dickinson, 2009; Hommel et al., 2001; Nattkemper et al., 2010; Shin et al., 2010), aIPL and pre-PMd are clearly regions that are relevant for mediating the role of 'non-incentive' action effects for the control of goal-directed action. In fact, within the ideomotor framework the role of differential action effects is often described in terms of automatic action priming after long training sessions involving hundreds of co-occurrences of responses and their associated differential effects implicating a detachment of such 'over-learned' R-O associations from incentive values. Hence, even action effects that are not actively pursued can induce action tendencies as may be most clearly expressed in imitative response tendencies induced by the observation of other agents' actions (Knuf et al., 2001). In line with this, particularly the aIPL is commonly activated in both, passive action observation with respect to the *goal* of action (Hamilton and Grafton, 2006) as well as passively perceiving tones that had previously been learnt to be linked with specific actions as compared to neutral tones (Melcher et al., 2008). Different from these passive observation and effect priming studies, other studies involved a more active use of anticipated action effects. In addition to the aIPL, such studies reported the pre-PMd rather than more ventrally located portions of the PMC that tend to be consistently observed in passive observation paradigms (Grafton and Hamilton, 2007; Ruge et al., 2010). Moreover, temporary disruption of PMd functionality by means of transcranial magnetic stimulation impaired participants' ability to predict action effects (Stadler et al., 2011). Interestingly, we observed a similar region (though somewhat more anteriorly located) that was sensitive to the outcome contingency manipulation in the present study, suggesting that contingent action effects were

actively involved in action planning processes. Lastly, we also found a *higher-order perceptual* visual area within the posterior ITG/MTG showing a contingency-sensitive practice-related increase in coupling with the pLPFC. This is consistent with a recent study showing an involvement of this area in action observation (Sasaki et al., 2012), but it contrasts with studies showing activation in rather *early sensory* areas specifically related to the perceptual modality of anticipated action effects (Kühn et al., 2011; Kühn et al., 2010).

It is important to highlight that our outcome contingency manipulation affected areas related to ideomotor mechanisms already very early in learning. This is interesting in so far as ideomotor mechanisms are typically investigated in the context of much longer learning sessions. Our results suggest that ideomotor learning is a very rapid process, consistent with recent behavioral results obtained with a passive effect priming procedure (Ruge et al., 2012; Wolfensteller and Ruge, 2011). Furthermore, limited practice might be the reason why outcome contingency did not also affect other areas (in particular SMA and cerebellum) reported in previous effect priming fMRI studies (Elsner et al., 2002; Melcher et al., 2008). This seems to suggest that these latter areas might start getting involved in ideomotor action control only after some degree of R-O automatization.

Conclusions

Together, our results suggest that the instruction-based formation of pragmatic rule representations relies on the strengthening of associations between actions and incentive as well as non-incentive outcome representations. *Fig. 5* schematically illustrates how these two associational paths are supposed to be activated by the differential outcomes employed in the present study. On the one hand, differential outcomes can activate a particular response-reward association. This affects instrumental response *decisions* on choosing the currently most valued action (i.e., the instructed action) for execution. On the other hand, via ideomotor response *selection* processes differential outcomes can directly activate the respective action representation irrespective of the currently associated reward value. This distinction, or the consequence of this distinction, becomes clearer when envisioning a scenario where multiple goals are concurrently achievable. While ideomotor *selection* processes would concurrently activate as many response options as there are associated with currently achievable goals, instrumental decision processes would initiate the execution of the one single action that is associated with the goal that currently yields highest reward. This notion is consistent with previous single cell recordings suggesting that multiple response options are concurrently active before a final decision can be reached (Cisek and Kalaska, 2005). Interestingly, this holds preferentially for neurons located within the anterior section of the PMd, which is one of

the regions we suggest to be linked to ideomotor processes based on the present results. Generally, our results emphasize the relevance of both instrumental as well as ideomotor mechanisms for the acquisition of novel goal-directed behaviors, which highlights the need for a theoretical integration of both theoretical frameworks (Butz and Hoffmann, 2002; de Wit and Dickinson, 2009; Elsner and Hommel, 2004; Shin et al., 2010; Wolfensteller and Ruge, 2012).

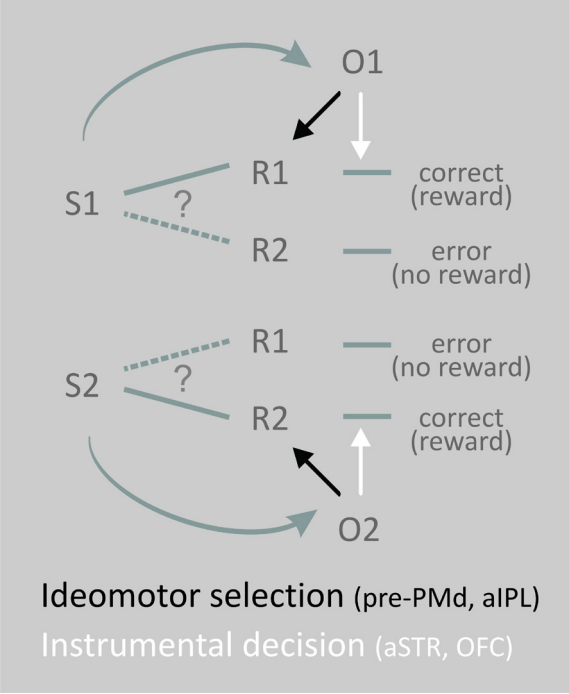

Fig. 5 Schematic illustration of two possible mechanisms mediating the impact of differential outcomes on goal-directed action. Via 'ideomotor selection' the anticipation of a specific outcome (e.g., O1) directly activates the associated response without reference to the incentive value associated with that outcome. Via 'instrumental decision' processes the anticipation of a specific outcome (e.g., O1) activates the association between a response and its incentive value. As one consequence, ideomotor selection processes are supposed to continue impacting action selection even when an outcome becomes devaluated whereas instrumental decision processes cease impacting behavior as soon as an outcome becomes devaluated.

On an even more general note, the present study suggest that the symbolic-pragmatic transfer of newly instructed S-R rules is accomplished by an increasing functional integration between the LPFC and a number of different cortical and striatal brain regions – and this integration occurs rapidly within the first eight implementation trials after instruction. This finding goes beyond what can be inferred from univariate analyses by demonstrating that decreasing activation in LPFC and similar regions does not per se indicate a diminishing relevance of such regions. Rather, it seems that increasingly efficient cooperation between distant areas (as indicated by increased functional coupling) reduces processing demands for each of the co-operating areas to accomplish a shared goal. Finally, that LPFC was increasingly coupled with anterior striatum and OFC highlights the relevance of these areas not only for situations where novel instrumental behaviors are learned via prediction-error signals, that is, when correct responding needs to be inferred from external feedback (e.g., Mattfeld and Stark, 2011; Noonan et al., 2011; O'Doherty et al., 2004), but also when novel behavior is learned via explicitly instructed symbolic rules supposedly stored in the LPFC 'procedural working memory' (Oberauer, 2009).

Acknowledgements

This work was supported by a grant from the German Research Council (Deutsche Forschungsgemeinschaft, DFG) to H. Ruge and U. Wolfensteller (RU 1539/2-1). The authors declare no competing financial interests.

References

Ashby, F.G., Turner, B.O., Horvitz, J.C., 2010. Cortical and basal ganglia contributions to habit learning and automaticity. Trends Cogn Sci 14, 208-215.

Brovelli, A., Nazarian, B., Meunier, M., Boussaoud, D., 2011. Differential roles of caudate nucleus and putamen during instrumental learning. Neuroimage 57, 1580-1590.

Butz, M.V., Hoffmann, J., 2002. Anticipations control behavior: Animal behavior in an anticipatory learning classifier system. Adaptive Behavior 10, 75-96.

Camara, E., Rodriguez-Fornells, A., Münte, T.F., 2009. Functional connectivity of reward processing in the brain. Frontiers in Human Neuroscience 2.

Cisek, P., Kalaska, J.F., 2005. Neural correlates of reaching decisions in dorsal premotor cortex: specification of multiple direction choices and final selection of action. Neuron 45, 801-814.

Cohen-Kdoshay, O., Meiran, N., 2009. The representation of instructions operates like a prepared reflex: flanker compatibility effects found in first trial following S-R instructions. Exp. Psychol. 56, 128-133.

Cole, M.W., Bagic, A., Kass, R., Schneider, W., 2010. Prefrontal Dynamics Underlying Rapid Instructed Task Learning Reverse With Practice. J. Neurosci. 30, 14245-14254.

Cole, M.W., Laurent, P., Stocco, A., 2012. Rapid instructed task learning: A new window into the human brain's unique capacity for flexible cognitive control. Cognitive, Affective, & Behavioral Neuroscience, 1-22.

Colwill, R.M., Rescorla, R.A., 1990. Effect of Reinforcer Devaluation on Discriminative Control of Instrumental Behavior. Journal of Experimental Psychology-Animal Behavior Processes 16, 40-47.

Cromer, J.A., Machon, M., Miller, E.K., 2011. Rapid Association Learning in the Primate Prefrontal Cortex in the Absence of Behavioral Reversals. J. Cogn. Neurosci. 23, 1823-1828.

Cross, K.A., Iacoboni, M., 2011. Optimized neural coding? control mechanisms in large cortical networks implemented by connectivity changes. Hum. Brain Mapp., doi: 10.1002/hbm.21428.

de Wit, S., Corlett, P.R., Aitken, M.R., Dickinson, A., Fletcher, P.C., 2009. Differential Engagement of the Ventromedial Prefrontal Cortex by Goal-Directed and Habitual Behavior toward Food Pictures in Humans. J. Neurosci. 29, 11330-11338.

de Wit, S., Dickinson, A., 2009. Associative theories of goal-directed behaviour: a case for animal-human translational models. Psychol. Res. 73, 463-476.

Dickinson, A., Shanks, D., 1995. Instrumental action and causal representation. In: Sperber, D., Premack, D., Premack, A.J. (Eds.), Causal Cognition. Clarendon Press, London.

Doll, B.B., Jacobs, W.J., Sanfey, A.G., Frank, M.J., 2009. Instructional control of reinforcement learning: a behavioral and neurocomputational investigation. Brain Res. 1299, 74-94.

Dumontheil, I., Thompson, R., Duncan, J., 2011. Assembly and Use of New Task Rules in Fronto-parietal Cortex. J. Cogn. Neurosci. 23, 168-182.

Duncan, J., 2001. An adaptive coding model of neural function in prefrontal cortex. Nat. Rev. Neurosci. 2, 820-829.

Elsner, B., Hommel, B., 2004. Contiguity and contingency in action-effect learning. Psychol. Res. 68, 138-154.

Elsner, B., Hommel, B., Mentschel, C., Drzezga, A., Prinz, W., Conrad, B., Siebner, H., 2002. Linking actions and their perceivable consequences in the human brain. Neuroimage 17, 364-372.

Frank, M.J., Badre, D., 2012. Mechanisms of Hierarchical Reinforcement Learning in Corticostriatal Circuits 1: Computational Analysis. Cereb. Cortex 22, 509-526.

Friston, K.J., Buechel, C., Fink, G.R., Morris, J., Rolls, E., Dolan, R.J., 1997. Psychophysiological and modulatory interactions in neuroimaging. Neuroimage 6, 218-229.

Gitelman, D.R., Penny, W.D., Ashburner, J., Friston, K.J., 2003. Modeling regional and psychophysiologic interactions in fMRI: the importance of hemodynamic deconvolution. Neuroimage 19, 200-207.

Grafton, S.T., Hamilton, A.F., 2007. Evidence for a distributed hierarchy of action representation in the brain. Hum Mov Sci 26, 590-616.

Hamilton, A.F., Grafton, S.T., 2006. Goal representation in human anterior intraparietal sulcus. J. Neurosci. 26, 1133-1137.

Hartstra, E., Kuhn, S., Verguts, T., Brass, M., 2011. The implementation of verbal instructions: an fMRI study. Hum. Brain Mapp. 32, 1811-1824.

Hommel, B., Musseler, J., Aschersleben, G., Prinz, W., 2001. The Theory of Event Coding (TEC): a framework for perception and action planning. Behav. Brain Sci. 24, 849-878; discussion 878-937.

Killcross, S., Coutureau, E., 2003. Coordination of actions and habits in the medial prefrontal cortex of rats. Cereb. Cortex 13, 400-408.

Knuf, L., Aschersleben, G., Prinz, W., 2001. An analysis of ideomotor action. Journal of experimental psychology. General 130, 779-798.

Kühn, S., Keizer, A.W., Rombouts, S.A.R.B., Hommel, B., 2011. The Functional and Neural Mechanism of Action Preparation: Roles of EBA and FFA in Voluntary Action Control. J. Cogn. Neurosci. 23, 214-220.

Kühn, S., Seurinck, R., Fias, W., Waszak, F., 2010. The internal anticipation of sensory action effects: when action induces FFA and PPA activity. Frontiers in Human Neuroscience 4, 12.

Li, J., Delgado, M.R., Phelps, E.A., 2011. How instructed knowledge modulates the neural systems of reward learning. Proceedings of the National Academy of Sciences 108, 55-60.

Mattfeld, A.T., Stark, C.E.L., 2011. Striatal and Medial Temporal Lobe Functional Interactions during Visuomotor Associative Learning. Cereb. Cortex 21, 647-658.

Melcher, T., Weidema, M., Eenshuistra, R.M., Hommel, B., Gruber, O., 2008. The neural substrate of the ideomotor principle: an event-related fMRI analysis. Neuroimage 39, 1274-1288.

Miller, E.K., Cohen, J.D., 2001. An integrative theory of prefrontal cortex function. Annu. Rev. Neurosci. 24, 167-202.

Monsell, S., 1996. Control of mental processes. In: Bruce, V. (Ed.), Unsolved mysteries of the mind. Erlbaum/Taylor and Francis, Hove, UK, pp. 93-148.

Muhle-Karbe, P.S., Krebs, R.M., 2012. On the influence of reward on action-effect binding. Frontiers in Psychology 3.

Nattkemper, D., Ziessler, M., Frensch, P.A., 2010. Binding in voluntary action control. Neurosci. Biobehav. Rev. 34, 1092-1101.

Noonan, M.P., Kolling, N., Walton, M.E., Rushworth, M.F.S., 2012. Re-evaluating the role of the orbitofrontal cortex in reward and reinforcement. Eur. J. Neurosci. 35, 997-1010.

Noonan, M.P., Mars, R.B., Rushworth, M.F.S., 2011. Distinct Roles of Three Frontal Cortical Areas in Reward-Guided Behavior. The Journal of Neuroscience 31, 14399-14412.

Noonan, M.P., Walton, M.E., Behrens, T.E.J., Sallet, J., Buckley, M.J., Rushworth, M.F.S., 2010. Separate value comparison and learning mechanisms in macaque medial and lateral orbitofrontal cortex. Proceedings of the National Academy of Sciences.

O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., Dolan, R.J., 2004. Dissociable roles of ventral and dorsal striatum in instrumental conditioning. Science 304, 452-454.

O'Doherty, J.P., 2011. Contributions of the ventromedial prefrontal cortex to goal-directed action selection. Ann. N. Y. Acad. Sci. 1239, 118-129.

O'Reilly, J.X., Woolrich, M.W., Behrens, T.E.J., Smith, S.M., Johansen-Berg, H., 2012. Tools of the trade: psychophysiological interactions and functional connectivity. Soc. Cogn. Affect. Neurosci. 7, 604-609.

Oberauer, K., 2009. Design for a Working Memory. Psychology of Learning and Motivation: Advances in Research and Theory, Vol 51 51, 45-100.

Pasupathy, A., Miller, E.K., 2005. Different time courses of learning-related activity in the prefrontal cortex and striatum. Nature 433, 873-876.

Picard, N., Strick, P.L., 2001. Imaging the premotor areas. Curr. Opin. Neurobiol. 11, 663-672.

Ramamoorthy, A., Verguts, T., 2012. Word and Deed: A computational model of instruction following. Brain Res. 1439, 54–65.

Rorden, C., Karnath, H.-O., Bonilha, L., 2007. Improving Lesion-Symptom Mapping. J. Cogn. Neurosci. 19, 1081-1088.

Ruge, H., Jamadar, S., Zimmermann, U., Karayanidis, F., 2011. The many faces of preparatory control in task switching: reviewing a decade of fMRI research. Hum. Brain Mapp.

Ruge, H., Krebs, R.M., Wolfensteller, U., 2012. Early markers of ongoing action-effect learning. Frontiers in Psychology 3.

Ruge, H., Muller, S.C., Braver, T.S., 2010. Anticipating the consequences of action: An fMRI study of intention-based task preparation. Psychophysiology 47, 1019-1027.

Ruge, H., Wolfensteller, U., 2010. Rapid Formation of Pragmatic Rule Representations in the Human Brain during Instruction-Based Learning. Cereb. Cortex 20, 1656-1667.

Sakai, K., 2008. Task set and prefrontal cortex. Annu. Rev. Neurosci. 31, 219-245.

Sasaki, A.T., Kochiyama, T., Sugiura, M., Tanabe, H.C., Sadato, N., 2012. Neural networks for action representation underlying automatic mimicry: A functional magnetic-resonance imaging and dynamic causal modeling study. Frontiers in Human Neuroscience 6.

Seger, C.A., Spiering, B.J., 2011. A critical review of habit learning and the Basal Ganglia. Frontiers in systems neuroscience 5, 66.

Shin, Y.K., Proctor, R.W., Capaldi, E.J., 2010. A review of contemporary ideomotor theory. Psychol. Bull. 136, 943-974.

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. Hum. Brain Mapp. 32, 677-687.

Stocco, A., Lebiere, C., O'Reilly, R., Anderson, J., in press. Distinct contributions of the caudate nucleus, rostral prefrontal cortex, and parietal cortex to the execution of instructed tasks. Cognitive, Affective, & Behavioral Neuroscience, 1-18.

Tanaka, S.C., Balleine, B.W., O'Doherty, J.P., 2008. Calculating consequences: brain systems that encode the causal effects of actions. J. Neurosci. 28, 6750-6755.

Trapold, M.A., 1970. Are Expectancies Based Upon Different Positive Reinforcing Events Discriminably Different. Learn. Motiv. 1, 129-140.

Tricomi, E., Balleine, B.W., O'Doherty, J.P., 2009. A specific role for posterior dorsolateral striatum in human habit learning. Eur. J. Neurosci. 29, 2225-2232.

Tricomi, E.M., Delgado, M.R., Fiez, J.A., 2004. Modulation of caudate activity by action contingency. Neuron 41, 281-292.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage 15, 273-289.

Urcuioli, P.J., 2005. Behavioral and associative effects of differential outcomes in discrimination learning. Learn. Behav. 33, 1-21.

Valentin, V.V., Dickinson, A., O'Doherty, J.P., 2007. Determining the neural substrates of goal-directed learning in the human brain. J. Neurosci. 27, 4019-4026.

Walsh, M.M., Anderson, J.R., 2011. Modulation of the feedback-related negativity by instruction and experience. Proc. Natl. Acad. Sci. U. S. A. 108, 19048-19053.

Walton, M.E., Devlin, J.T., Rushworth, M.F.S., 2004. Interactions between decision making and performance monitoring within prefrontal cortex. Nat. Neurosci. 7, 1259-1265.

Waszak, F., Wenke, D., Brass, M., 2008. Cross-talk of instructed and applied arbitrary visuomotor mappings. Acta Psychol. (Amst.) 127, 30-35.

Wenke, D., Gaschler, R., Nattkemper, D., 2007. Instruction-induced feature binding. Psychol. Res. 71, 92-106.

Wolfensteller, U., Ruge, H., 2011. On the timescale of stimulus-based action-effect learning. Q. J. Exp. Psychol. 64, 1273-1289.

Wolfensteller, U., Ruge, H., 2012. Frontostriatal mechanisms in instruction-based learning as a hallmark of flexible goal-directed behavior. Frontiers in Psychology 3.

Wood, W., Neal, D.T., 2007. A new look at habits and the habit-goal interface. Psychol. Rev. 114, 843-863.

Woolgar, A., Hampshire, A., Thompson, R., Duncan, J., 2011. Adaptive Coding of Task-Relevant Information in Human Frontoparietal Cortex. The Journal of Neuroscience 31, 14592-14599.

Xue, G., Ghahremani, D.G., Poldrack, R.A., 2008. Neural substrates for reversing stimulus-outcome and stimulus-response associations. J. Neurosci. 28, 11196-11204.

Table 1. Summary of PPI results for orbitofrontal and basal ganglia regions of interest

	L pLPFC seed				**R pLPFC seed**			
Region identified by the PPI	peak voxel [overall]	t [overall]	peak vox [cont-rnd]	t [cont-rnd]	peak voxel [overall]	t [overall]	peak vox [cont-rnd]	t [cont-rnd]
L lateral OFC	-42 47 -11	4.73#	-42 50 -8	2.51$	/	n.s.	-39 50 -8	2.60$
L central OFC	-27 41 -14	4.06#	-24 38 -11	3.41$	/	n.s.	-24 41 -11	2.74$
L central OFC	-18 38 -5	3.82#	-21 35 -8	2.78$	/	3.21$	-24 38 -5	2.68$
L central OFC	-24 23 -17	4.22#	-21 26 -14	2.76$	/	2.88$	-21 26 -14	2.45$
R lateral OFC	48 41 -5	5.08#	48 41 -2	2.49$	42 44 -2	5.12#	/	n.s.
R lateral OFC	33 50 -8	4.48#	33 50 -5	2.57$	/	2.69$	/	n.s.
R central OFC	36 35 -17	3.61#	/	n.s.	/	n.s.	/	n.s.
L ventral STR	-9 8 -8	4.73#	-9 14 -8	2.42$	-9 8 -8	4.17#	/	n.s.
R ventral STR	9 11 -8	4.35#	6 8 -5	2.91 $	12 14 -8	4.27#	/	n.s.
L mid Pallidum	/	2.88$	/	n.s.	-15 2 7	3.53#	/	n.s.
R mid Pallidum	12 2 -2	5.34#	/	n.s.	15 -1 -5	5.05#	/	n.s.
L ad Putamen	/	n.s.	/	n.s.	-24 14 7	4.19#	/	n.s.
L ad Putamen	/	n.s.	/	n.s.	-27 8 12	4.15#	/	n.s.
R ad Putamen	/	n.s.	/	n.s.	30 17 7	4.32#	/	n.s.
L Putamen tail	/	2.53$	/	n.s.	-30 -19 7	3.53#	/	n.s.
R Putamen tail	/	2.51$	/	n.s.	30 -19 7	4.09#	/	n.s.

 # p<.05 FWE-corrected for anatomical ROIs (basal ganglia or OFC); $p<.05 FWE-corrected for 6mm sphere ROIs centered on peak voxels identified in the overall PPI analysis (either for left or right pLPFC seed)

Table 2. Summary of PPI results for the whole brain volume

	L pLPFC seed				**R pLPFC seed**			
Region identified by the PPI	peak voxel [overall]	t [overall]	peak vox [cont-rnd]	t [cont-rnd]	peak voxel [overall]	t [overall]	peak vox [cont-rnd]	t [cont-rnd]
pSFG (Pre-PMd)	/	2.61$^{\$}$	/	n.s.	24 5 61	5.71~	27 5 64	2.57$^{\$}$
R pMFG	/	n.s.	/	n.s.	51 11 34	6.03~	/	n.s
R aMFG	/	3.52$^{\$}$	/	n.s.	45 53 7	5.75~	/	n.s
R aMFG	/	n.s.	/	n.s.	39 50 19	5.52~	/	n.s
R pIFG	/	2.92$^{\$}$	/	n.s.	51 11 10	5.14~	/	n.s
R aINS	/	n.s.	/	n.s.	45 20 -2	4.97~	/	n.s
L SMG/aIPL	-60 -37 40	6.91~	/	n.s	-57 -37 40	6.11~	-63 -36 40	2.93$^{\$}$
L SMG	-62 -34 31	6.52~	-65 -31 34	2.87$^{\$}$	-65 -31 34	5.86~	-65 -31 34	4.30$^{\$}$
R SMG	63 -40 40	6.10~	/	n.s	60 -37 37	7.03~	/	n.s.
R aIPL/aSPL	/	3.10$^{\$}$	/	n.s	42 -46 58	5.93~	42 -43 55	3.39$^{\$}$
R aIPL	/	2.85$^{\$}$	/	n.s	36 -46 46	5.75~	36 -43 49	2.74$^{\$}$
R mHIPP/ Putamen tail	/	2.64$^{\$}$	/	n.s.	30 -16 -8	4.78~	/	n.s
L pITG	-57 -52 -2	6.01~	-60 -52 -5	2.63$^{\$}$	-60 -61 -8	4.77~	/	n.s.
L pMTG	-60 -64 10	5.40~	-60 -64 7	3.06$^{\$}$	/	2.69$^{\$}$	/	n.s.
R pITG	60 -49 -2	5.95~	54 -49 -2	3.03$^{\$}$	57 -52 -2	6.47~	/	n.s
R pMTG	60 -61 10	5.35~	63 -58 10	2.61$^{\$}$	/	3.33$^{\$}$	/	n.s
L pSPL	-12 -79 55	4.79~	/	n.s	-12 -79 55	5.52~	/	n.s
R pSPL	12 -85 46	5.28~	/	n.s.	18 -73 52	6.57~	/	n.s.
L a OCC	-9 -76 -5	5.93~	/	n.s.	-12 -79 -5	9.17~	/	n.s.
R a OCC	6 -79 -2	5.96~	/	n.s.	11 -79 -6	8.78~	/	n.s.
L p OCC	-3 -94 13	6.13~	/	n.s	-3 -94 10	8.52~	/	n.s
L Cerebellum	/	2.78$^{\$}$	/	n.s.	-30 -70 -29	4.94~	/	n.s.
L Cerebellum	/	2.85$^{\$}$	/	n.s.	-21 -67 -32	4.91~	/	n.s.
L Cerebellum	/	2.69$^{\$}$	/	n.s.	-39 -64 -32	4.78~	/	n.s.

~ p<.05 FWE-corrected for whole brain volume; $^{\$}$ p<.05 FWE-corrected for 6mm sphere ROIs centered on peak voxels identified in the overall PPI analysis (either for left or right pLPFC seed)