



# Frontostriatal mechanisms in instruction-based learning as a hallmark of flexible goal-directed behavior

Uta Wolfensteller <sup>\*†</sup> and Hannes Ruge <sup>†</sup>

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

**Edited by:**

Bernhard Hommel, Leiden University, Netherlands

**Reviewed by:**

Markus Janczyk, University of Würzburg, Germany

Marco Steinhauser, University of Konstanz, Germany

Miriam Gade, University of Zurich, Switzerland

Martina Rieger, University for Health Sciences, Austria

**\*Correspondence:**

Uta Wolfensteller, Department of Psychology, Technische Universität Dresden, Zellescher Weg 17, 01062 Dresden, Germany.  
e-mail: uta.wolfensteller@tu-dresden.de

<sup>†</sup>Uta Wolfensteller and Hannes Ruge have contributed equally to this work.

The present review intends to provide a neuroscientific perspective on the flexible (here: almost instantaneous) adoption of novel goal-directed behaviors. The overarching goal is to sketch the emerging framework for examining instruction-based learning and how this can be related to more established research approaches to instrumental learning and goal-directed action. We particularly focus on the contribution of frontal and striatal brain regions drawing on studies in both, animals and humans, but with an emphasis put on human neuroimaging studies. In section one, we review and integrate a selection of previous studies that are suited to generally delineate the neural underpinnings of goal-directed action as opposed to more stimulus-based (i.e., habitual) action. Building on that the second section focuses more directly on the flexibility to rapidly implement novel behavioral rules as a hallmark of goal-directed action with a special emphasis on instructed rules. In essence, the current neuroscientific evidence suggests that the prefrontal cortex and associative striatum are able to selectively and transiently code the currently relevant relationship between stimuli, actions, and the effects of these actions in both, instruction-based learning as well as in trial-and-error learning. The premotor cortex in turn seems to form more durable associations between stimuli and actions or stimuli, actions and effects (but not incentive values) thus representing the available action possibilities. Together, the central message of the present review is that instruction-based learning should be understood as a prime example of goal-directed action, necessitating a closer interlacing with basic mechanisms of goal-directed action on a more general level.

**Keywords:** ideomotor theory, instrumental learning, instruction, prefrontal cortex, premotor cortex, basal ganglia

Rapidly adopting novel rules defining which actions yield the desired outcome under different circumstances is a pivotal expression of human behavioral flexibility. For humans the most efficient way to acquire such novel goal-directed behavior is to make use of explicit instructions. Imagine an infant girl being told to firmly press the biscuit cutter into the dough, then to carefully lift it up again, and voila there's a heart-shaped biscuit – hooray – and just like that a novel goal-directed action emerged by instruction. The processes that mediate the implementation of novel and explicitly instructed behavioral rules are central to executive control function, but research has been surprisingly scarce as already noted more than a decade ago (Monsell, 1996). Instead, the acquisition of novel behavioral rules has been studied mostly by means of instrumental trial-and-error learning procedures. In comparison to that, the human capacity of learning by instruction offers a short-cut for acquiring the same novel behaviors much faster, thereby minimizing possible harm in case of trying the wrong action (Doll et al., 2009; Walsh and Anderson, 2011). Only recently has instruction-based learning started garnering broader scientific interest (Hommel, 2000; Wenke et al., 2007; Waszak et al., 2008; Cohen-Kdoshay and Meiran, 2009), especially in the cognitive neuroscience domain (Doll et al., 2009; Cole et al., 2010; Ruge and Wolfensteller, 2010; Dumontheil et al., 2011; Hartstra et al.,

2011; Li et al., 2011; Bugmann, 2012; Ramamoorthy and Verguts, 2012). The present review aims to sketch the emerging framework for examining instruction-based learning in cognitive neuroscience and how this can be related to more established research approaches to instrumental learning and goal-directed action. As a very first step we will define the key terms and concepts the present review deals with in a few introductory notes by addressing two opening questions. First, what makes an action-goal-directed? And second, why study instruction-based learning – or else – is learning by instruction better than learning by trial-and-error?

## WHAT MAKES AN ACTION-GOAL-DIRECTED? INTRODUCTORY NOTES

In terms of instrumental learning and behavior the execution of a goal-directed action depends on the rewarding (reinforcing) properties of its effect (Thorndike, 1911; Colwill and Rescorla, 1986; Dickinson et al., 1996). For instance, your behavior will be considered goal-directed if you stop performing an action R (e.g., pushing the buttons on the coffee machine) as soon as you either do not desire or do not believe that you get the outcome (O, or effect E) of that action (coffee) anymore (cf. Balleine et al., 2009). In contrast, your behavior will be considered habitual or directly stimulus-based (S-R) if your action was not sensitive

to such reinforcer devaluation and you continued to perform it regardless. Thus, goal-directed or outcome-based behavior rests on associations between responses (button pressing) to certain stimulus situations (coffee machine) and their effects or outcomes (coffee). The mental representation of such differential action outcomes (Colwill and Rescorla, 1990; Urciuoli, 2005; Shin et al., 2010) allows, for instance, selecting among competing alternative responses (R) that might have been learnt independently for the same stimulus (S) situation or that might have been learnt to produce different effects depending on a particular S. In addition to instrumental learning mechanisms, ideomotor theory proposed a similar action-effect binding mechanism (see Lotze, 1852; Harless, 1861; James, 1890; Greenwald, 1970b; Hommel et al., 2001; for a recent translation of Harless' work, see Pfister and Janczyk, 2011; for a historical review, see also Stock and Stock, 2004). According to the ideomotor framework, actions and their perceived effects become integrated such that the mere anticipation (or *idea*) of an effect primes the associated *motor* action, and likewise the performance of an action goes along with an anticipation of its effect. Interestingly, the frameworks of ideomotor and instrumental learning have only recently begun to become more integrated (Butz and Hoffmann, 2002; Elsner and Hommel, 2004; de Wit and Dickinson, 2009; Shin et al., 2010). With respect to differential action-effects, both have demonstrated that by incorporating stimulus information, response-effect (R-E) or response-outcome (R-O) associations can be contextualized (e.g., Colwill and Rescorla, 1985, 1990; Kunde, 2001; Ziessler et al., 2004; Hoffmann et al., 2007; Ruge et al., 2010; Wolfensteller and Ruge, 2011). In other words, the action you will select in a given stimulus situation depends on the goal you currently pursue. Importantly, in that sense, any behavior based on instructed S-R rules would essentially be goal-directed or effect-based at least early in practice (Dickinson and Shanks, 1995; Killcross and Coutureau, 2003; Wood and Neal, 2007), because the action is performed in order to perform correctly as instructed, that is to achieve success and to avoid the alternative outcome of failure. This implies then that performing correctly must be intrinsically rewarding as otherwise the respective action should not be shown. Two recent functional imaging (fMRI) studies provide direct support for this notion by showing that the ventral striatum – a central part of the brain's reward system – is engaged in processing positive monetary and cognitive feedback (Daniel and Pollmann, 2010), and is modulated by how confident of being accurate a person feels (Daniel and Pollmann, 2012). However, given enough practice you might very well find yourself carrying out an action in a more automatic manner upon the stimulus at hand without considering the effects of this action anymore.

Note that due to its focus on instruction-based learning the present review will not speak to free will related aspects of goal-directed action<sup>1</sup>.

<sup>1</sup>This is an exciting research topic of itself that has inspired a large number of neuroscientific studies in recent years. Typically they compare a condition where participants can freely choose which action to perform or not to perform (free-choice or self-generated action) with a condition where participants are instructed which action to perform by an external stimulus (forced-choice or externally triggered action). Converging evidence suggests a crucial role of the pre-supplementary

## LEARNING BY TRIAL-AND-ERROR AND LEARNING BY INSTRUCTION: INTRODUCTORY NOTES

The concept of learning by trial-and-error dates back to the early days of instrumental learning (e.g., Thorndike, 1911). In a nutshell, you realize that your action is correct if it gets reinforced by monetary reward or positive cognitive feedback in humans, or juice or food pellets in monkeys and rats. Conversely, you realize that an alternative action might be correct, if your action leads to monetary loss, negative feedback, icky-tasting food, electric shocks, or simply no reward. Depending on the difficulty of the task and the number of response alternatives it might take quite a while to figure out what is correct under which circumstances. In contrast, humans can adopt and behaviorally implement novel stimulus-response (S-R) rules almost instantaneously if explicitly instructed. Most experimental laboratories explicitly use this ability by simply instructing their participants rather than training them over extensive time periods as is more common praxis in animal research. But is there direct empirical evidence for the superiority of instruction-based learning (at least in terms of time)? Unsurprisingly, the answer is yes. A recent example can be found in a study on probabilistic learning (Walsh and Anderson, 2011) which showed that with prior instruction (but not without), behavior started, and stayed at asymptotic level. Thus, instruction-based learning of novel rules can be considered to be one of the prime examples of the flexibility of human goal-directed behavior.

Of course, the results of instrumental learning by trial-and-error and of instruction-based learning approaches will ultimately converge on a nearly error-free and fluent behavior. Notwithstanding that, the underlying learning mechanisms are obviously different and are likely to be reflected by distinct learning-related neuronal processes as indicated for instance by neuropsychological investigations of human patients (Vriezen and Moscovitch, 1990; Petrides, 1997). Patients who suffered from lesions within frontal cortex had difficulties learning novel S-R rules irrespective of whether they received an instruction or not (Petrides, 1985, 1997). In contrast, patients with basal-ganglia dysfunction were impaired in learning by trial-and-error, but showed normal performance levels when learning by instruction (Vriezen and Moscovitch, 1990). When comparing the neurocognitive mechanisms underlying trial-and-error learning and instruction-based learning it is important to distinguish between two processes that are relevant for contingencies between actions and goals to gain control over behavior. One process is specifically relevant in the typical instrumental learning situation and is responsible for extracting action-goal contingencies from response feedback during trial-and-error learning. The second process refers to the behavioral, or pragmatic, implementation of symbolically represented action-goal contingencies that are explicitly stored in working memory, or more precisely, as suggested recently, in a “procedural working memory” sub-system (Oberauer, 2009; Souza et al., 2012). Notably, it is this second process that defines instruction-based

motor area for self-generated actions (Jenkins et al., 2000; Lau et al., 2004, 2006; Mueller et al., 2007; Waszak et al., 2012). A recent in-depth discussion of the concept of self-generated actions can be found elsewhere (Nachev and Husain, 2010; Passingham et al., 2010; Schuur and Haggard, 2011, 2012; Obhi, 2012).

learning situations<sup>2</sup>. For trial-and-error learning, it seems less clear whether explicit rule knowledge is always generated and used to support a hypothesis-driven strategic approach to extract the currently valid contingencies (Haruno et al., 2004; Hadj-Bouziane et al., 2006; Frank and Badre, 2012). Alternatively or concurrently, trial-and-error learning might also proceed more implicitly via reinforcement learning mechanisms based on a gradual trial-by-trial updating of contingency representations as a function of the outcome prediction-error (Glascher et al., 2010).

## SECTION ONE: THE NEURAL CORRELATES OF GOAL-DIRECTED ACTION

In this section we review and integrate a selection of previous studies that are suited to generally delineate the neural underpinnings of *goal-directed action* as opposed to more stimulus-based (i.e., habitual) action. Most of the studies are based on altering in one way or the other the integration of goal-information during learning and beyond either by explicitly manipulating features of the outcomes that are entailed by specific actions or by tracking the transition from goal-directed to stimulus-based (i.e., increasingly habitual) action control. In this section we will integrate findings from both, instrumental learning and ideomotor theory following a recent endeavor of cross-fertilization between the two closely related but still largely segregated research frameworks (see also Hommel et al., 2001; de Wit and Dickinson, 2009; Shin et al., 2010).

### APPROACH 1: STUDYING OUTCOME-BASED ACTION CONTROL

Within the instrumental conditioning framework animal lesion studies have implicated different structures within homologs of the human basal ganglia and the medial and orbital prefrontal cortices in the control of goal-directed behavior involving R-O or S-R-O associations as compared to stimulus-based habitual behavior involving S-R associations or “pavlovian” associative processes linking S and O. In particular, goal-directed actions based on R-O or S-R-O associations draw on the associative striatum (asSTR)<sup>3</sup>,

while habitual actions based on S-R associations draw more on the sensorimotor striatum (smSTR; for reviews, see Yin and Knowlton, 2006; Ashby et al., 2010; Balleine and O’Doherty, 2010; van der Meer and Redish, 2010). By contrast, functional imaging research uncovering the brain structures involved in goal-directed as compared to stimulus-based behavior in humans is still scarce. The few imaging studies in humans have used mainly two different approaches: one is to use outcome devaluation in order to investigate how differential outcome values are represented in the brain (Valentin et al., 2007; de Wit et al., 2009). Complementary, another approach examines how manipulations of the R-O contingency modulate brain activation (Tricomi et al., 2004; Tanaka et al., 2008). Naturally, these studies did not attempt to systematically distinguish between all the different types of associations that might be formed under instrumental conditioning regimes, that is, R-O or S-R-O associations as the basis of truly goal-directed action as compared to S-R habits or pavlovian S-O associations, but rather selectively contrasted some of these contingencies.

For instance, Tricomi et al. (2004) reported that the asSTR was specifically involved in expecting incentive outcomes following an action (R-O) but not following a predictive stimulus without an action (S-O). Similarly, learning of S-R-O and S-O association was differentially related to the asSTR and the ventral striatum, respectively (O’Doherty et al., 2004). Converging evidence stems from an fMRI study investigating free operant conditioning (Tanaka et al., 2008), which revealed that high compared to low R-O contingency was associated with stronger engagement in the asSTR alongside with the ventromedial prefrontal cortex (VMPFC) and orbitofrontal cortex (OFC). Two other studies (Valentin et al., 2007; de Wit et al., 2009) aimed to dissociate habitual (S-R) action from effect-based (S-R-O) action. This is an important endeavor since the activation of goal representations (outcome-based action control) might be a mere epiphenomenon if in accord with an established S-R habit (stimulus-based action control) as pointed out for example by Wood and Neal (2007). While Valentin et al. (2007) employed the classical devaluation paradigm where response-specific differential outcomes had incentive values, de Wit et al. (2009) employed a novel paradigm based on creating competition between stimulus-based and outcome-based action control with differential response effects (fruit symbols) bearing no intrinsic incentive value similar to studies on ideomotor learning (see section below). More specifically, the condition targeting goal-directed action was constructed such that upon presentation of a particular fruit symbol A, a specific response would result in the presentation of the same fruit symbol A (and winning

<sup>2</sup>Neurological case studies again suggest dissociable brain mechanisms for symbolic and pragmatic representations (Luria, 1973). For instance, after being instructed to respond to red light by pressing firmly and to respond to blue light by pressing softly patients with parietal and frontal lesions are heavily impaired. However, when asked to continuously verbalize the currently relevant response, the patients with parietal lesions are able to do so, and most importantly this manipulation also restores the instructed behavior. In contrast, patients with frontal lesions, while still being able to verbalize the currently relevant response, fail to correctly translate this declarative knowledge into overt behavior. These early case reports bear some resemblance to more recent empirical findings. One is goal-neglect (Duncan et al., 1996, 2008), which is the failure to implement a particular aspect of a task despite being well able to describe it. Another one is utilization behavior originally described in patients suffering from frontal lobe lesions (Lhermitte et al., 1986), who performed certain actions such as lighting a match upon seeing it despite being able to verbalize that they knew it was inappropriate.

<sup>3</sup>Based on functional and structural differences revealed in rodents, non-human, and human primates the dorsal striatum is typically divided in two parts. The asSTR comprises the head of the caudate nucleus as well as the part of the putamen anterior to the anterior commissure while the smSTR comprises the part of the putamen posterior to the anterior commissure (Ashby et al., 2010). Generally, instrumental learning research suggests that the asSTR is a relevant structure for learning and representing response-outcome associations (O’Doherty et al., 2004; Tricomi et al., 2004; Yin et al., 2005; Atallah et al., 2007; Tanaka et al., 2008). By contrast, the

smSTR seems to be relevant for forming habitual S-R associations (Yin et al., 2004; Yin and Knowlton, 2006; Atallah et al., 2007; Tricomi et al., 2009). Importantly, these striatal regions are known to be parts of separate cortico-striato-thalamic loops (Alexander et al., 1986; Yin and Knowlton, 2006; Grahn et al., 2009). The associative cortico-striato-thalamic loop links prefrontal and parietal association areas including dorsomedial and dorsolateral PFC with the asSTR. The sensorimotor cortico-striato-thalamic loop links sensorimotor cortical regions, i.e., premotor and primary motor cortex, with the smSTR (Yin and Knowlton, 2006). Recent research suggests interaction between these loops, via connections to the dopaminergic midbrain and to separate yet densely interconnected amygdalar nuclei (Yin and Knowlton, 2006; Grahn et al., 2009; Pennartz et al., 2011).

points). The influence of the stimulus-based action control system was tested in a condition where responding to fruit A would result in the presentation of fruit B while responding to fruit B would result in the presentation of fruit A (and winning points in both cases). Thus, in the stimulus-based condition, outcome anticipation would result in activating the wrong response which should discourage the goal-directed action mode. Stimulus-based (or habitual) action control was associated with enhanced activation in smSTR. Outcome-based action control was associated with activation in VMPFC in both, de Wit et al.'s and Valentin et al.'s study despite using quite different experimental protocols. Notably, de Wit et al. (2009) additionally reported enhanced activation in dorsal premotor cortex (PMC) for effect-based action control. We will discuss this latter observation below in the section on ideomotor action.

While these two studies probed the incorporation of differential response outcomes independent of its evolution across S-R learning, a recent study compared conditions with differential vs. random outcomes *during* trial-and-error S-R learning as a function of particularly informative feedback trials (Noonan et al., 2011). The results suggest that VMPFC and adjacent medial orbitofrontal (OFC) activation reflect the subjective value of expected outcomes, whereas the lateral OFC in co-operation with ventral striatum might be the region that supports the updating of S-O and R-O associations during trial-and-error learning. Together, all three studies support in different ways the original idea proposed by the differential outcome paradigm that intrinsically incentive as well as non-incentive action-effect features (e.g., Mok and Overmier, 2007) – if they discriminate between different actions – are tightly intermeshed with instrumental learning mechanisms (Trapold, 1970; Urcuioli, 2005). Nevertheless, it should be noted that sensitivity to differential outcomes in areas related to reward processes not only when outcomes are intrinsically incentive (Valentin et al., 2007), but also when they are non-incentive (de Wit et al., 2009) or when their incentive value is only indirectly mediated via tokens (Noonan et al., 2011) might be due to the fact that all three imaging studies examined relatively early phases of practice. As will be discussed later on, this might disguise functional differences between non-incentive action “effects” and incentive action “outcomes.” The following section will highlight studies that specifically target goal-directed actions involving non-incentive action-effects after comparably long training sessions.

## APPROACH 2: STUDYING EFFECT-BASED ACTION CONTROL

Previous imaging studies investigating effect-based action control in the ideomotor approach provide strong evidence for the bidirectional nature of action-effect associations. The experimental design typically adopted is a two-step effect-priming procedure (Greenwald, 1970a; Elsner and Hommel, 2001), where an initial acquisition phase which contingently paired two freely chosen responses with two specific auditory (Elsner et al., 2002; Melcher et al., 2008) or visual effects (Kühn et al., 2010) is followed by a test phase. In the test phase, the previously learnt effect (E) is either presented on its own without a response (Elsner et al., 2002; Melcher et al., 2008), serves as an unspecific go-signal for a previously selected response after a delay (Melcher et al., 2008), or

responses are to be freely chosen but effects are no longer presented (Kühn et al., 2010). The main findings are that (i) upon performing responses that had previously been paired with a specific sensory effect, activation was observed in the respective sensory cortical areas (Kühn et al., 2010) and (ii) upon presenting a sensory stimulus that had previously been the effect of a motor response activation in motor and premotor areas was enhanced (Elsner et al., 2002; Melcher et al., 2008). Interestingly, none of these studies report activation in the asSTR, VMPFC, or OFC for effect-based action control, which contrasts findings from instrumental learning discussed above. However, outcome devaluation studies indicate that the incentive aspect of differential outcomes becomes ineffective for making response decisions after some amount of practice beyond the initial instrumental acquisition phase (Killcross and Coutureau, 2003). By contrast, in the ideomotor paradigm, differential response effects are typically “over”-learned across hundreds of trials, yet without losing their potential to automatically prime response selection later on (Nattekemper et al., 2010). The level of automaticity of R-E associations most likely explains the absence of activation in the aforementioned regions in studies testing ideomotor learning. However, it poses the question as to where else these associations are stored or represented at that point. A very likely candidate is the PMC, as indicated by a couple of studies investigating effect-based action control from quite different angles.

In an ideomotor inspired approach to goal-directed action, Ruge et al. (2010) investigated the neural correlates of differential as compared to common response effects during action planning in a task switching design. Participants had to indicate either the horizontal or the vertical position of ambiguous targets. Importantly, two types of feedback were given, one corresponding to common response effects (correct/incorrect), and one corresponding to differential response effects (coloring of indicated location). In task switch trials compared to task repetition trials disambiguation of these differential response effects was necessary. This disambiguation was associated with enhanced activation in dorsolateral PFC, PMC and anterior intraparietal sulcus. Based on their findings, Ruge et al. (2010) suggested that posterior frontal regions such as the PMC represent specific response-effect (R-E) associations, whereas more anterior lateral prefrontal cortex (LPFC) regions provide set-level information as to which set of goals can currently be achieved (visual motion effects to the left or right vs. up or down). This interpretation nicely fits with two recent studies in humans revealing the crucial contribution of the dorsal PMC to action-effect prediction. In an fMRI study, the dorsal PMC was strongly engaged whenever participants had to judge whether an ongoing goal-directed action that had temporarily been occluded was correctly continued (Stadler et al., 2011). In order to reach a correct conclusion, action-effects had to be continuously predicted: starting to reach for some cup should be followed by grasping it and so on. Temporal disruption of dorsal PMC functionality by means of transcranial magnetic stimulation impaired participants' ability to predict action-effects (Stadler et al., 2012). Corroborating evidence for the role of the dorsal PMC in action-effect prediction stems from single cell recordings in monkeys. In these studies, similar neuronal activity was observed when the monkeys performed an action to reach a particular spatial effect,

and when they watched the same action (Tkach et al., 2007) or even just a cursor being moved to reach the same spatial effect (Cisek and Kalaska, 2004).

As outlined in the section on instrumental learning, it might be of particular interest to oppose R-E based action control and more S-R based action control. Though most previous studies on incidental effect-learning were not specifically designed for that purpose, some of them nevertheless offer some valuable insights. For instance, in one of the experimental conditions in the study by Melcher et al. (2008) stimulus and effect were incompatible with respect to the associated response. More specifically, participants had to respond to stimuli while simultaneously hearing tones that had previously served as effects of just the opposite response. Thereby competition was induced between goal representations activated by the previously learnt but currently irrelevant R-E association and the currently relevant S-R association. Notably, under these circumstances enhanced activity in posterior LPFC was observed. In another recent imaging study participants had to indicate the middle of a temporal interval using either S-R or R-E associations (Mueller et al., 2007). In the stimulus-based (S-R) condition they made a forced-choice, pressing the button spatially compatible to a visual stimulus presented to the left or right of the screen center. By contrast, in the effect-based (R-E) condition participants could freely choose to press a button depending on where they wanted the stimulus to appear in the next trial<sup>4</sup>. Effect-based action control was associated with comparatively stronger activation in posterior medial PFC as well as anterior LPFC. However, it seems noteworthy that the activations reported by Mueller et al. (2007) might also reflect a certain degree of conflict between effect-based and stimulus-based action control because in some cases the freely chosen goal (next location) and the currently irrelevant stimulus (spatial location of the stimulus) are incompatible. This in turn would be in line with the assumption that more anterior frontal regions provide set-level information, or biasing signals in case competition arises, or selection of the appropriate action is more difficult.

To sum up, though the findings from ideomotor learning approach are less unequivocal than those from instrumental learning, several consistencies emerge. In particular, the PFC seems to be providing goal-information, though the nature of the goal or effect (non-incentive or incentive) and the nature of response mode (forced or free) might well determine whether more lateral or more medial PFC regions are involved. As a quite fundamental difference, while instrumental approaches also report a distinction at the level of striatal sub-regions, ideomotor approaches typically fail to find activation in the striatum<sup>5</sup>. As outlined above, this most likely reflects a particular aspect of the experimental

design typically employed, which is to investigate correlates of R-E learning after overtraining. Due to the transient asSTR engagement the critical period might be missed. One notable exception is a recent fMRI study indicating that connectivity between PFC and asSTR might be influenced by R-E contingency (Ruge and Wolfensteller, submitted, see also section two below). Moreover, ideomotor approaches typically report a differentiation at the level of the PMC indicating a functional contribution over and above S-R representations (as would be required in both stimulus-based and effect-based action control). A potential explanation is that at least in the case of non-incentive action-effects the PMC represents all possible S-R-E associations from which a person might select. In the presence of salient rewards, this potential to select might be overruled by desirability strengthening the one rewarded S-R(E) to such an extent that it resembles an S-R association. Interestingly, recent single cell recordings of dorsal PMC neurons in monkeys lend support to this notion (Pastor-Bernier and Cisek, 2011). When presented with one spatial target, the neuronal response clearly reflected the spatial effect preference of the neuron and was not modulated by different incentive values. However, when two spatial targets were simultaneously presented, neuronal responses reflecting both spatial effects (movement directions) were observed (Cisek and Kalaska, 2005). Moreover, the neuronal response for the preferred target was modulated by the relative difference of incentives associated with the preferred and the non-preferred target (Pastor-Bernier and Cisek, 2011). Thus, it seems clear that both, incentive as well non-incentive differential action-effects play a role for goal-directed action – via distinct mechanisms that are dissociated in terms of the conditions that mediate their impact on overt behavior and in terms of the underlying brain systems.

### APPROACH 3: STUDYING THE TRANSITION FROM GOAL-DIRECTED TO STIMULUS-BASED ACTION

As outlined before, it is well established that actions are goal-directed only at early stages of instrumental conditioning. For instance, short instrumental training in rats (5 sessions, 50 rewards in total) resulted in goal-directed behavior as indicated by a reduction in response rate after devaluation of the outcome (Killcross and Coutureau, 2003). In contrast, training another response for a longer period (20 sessions, 500 rewards in total) resulted in habitual behavior as indicated by the fact that devaluation of the respective outcome had no impact on the response rate. Typically, habitual actions are assumed to be solely controlled by the dorsal PMC and the smSTR whereas the asSTR is assumed to gradually fade out with progressing automatization (Ashby et al., 2010). This notion is supported by a large number of studies investigating so-called conditional motor behavior which requires to form and use arbitrary associations between stimuli and responses (Kurata and Wise, 1988; Mitz et al., 1991; Brasted and Wise, 2004; Buch et al., 2006). In general, these studies highlight the roles of the dorsal PMC and the smSTR as performance increases. Moreover, the role of the smSTR in habitual action control was recently also confirmed in an fMRI study in humans. A decrease in smSTR activity

<sup>4</sup>A detailed discussion of free- and forced-choice measures of R-E learning is beyond the scope of this review (see instead Herwig and Waszak, 2009; Pfister et al., 2011; Wolfensteller and Ruge, 2011; Waszak et al., 2012).

<sup>5</sup>However, the perception of action-effects was associated with enhanced activation in posterior hippocampus (Elsner et al., 2002; Melcher et al., 2008) which might establish an indirect link. Rodent studies have revealed hippocampal projections to the asSTR (see also Voorn et al., 2004; Pennartz et al., 2011). It has been suggested that the representation and ultimately the behavioral expression of action-effect contingencies might depend also on the intactness of the hippocampal input to the striatum possibly providing episodic memory information (Frank et al., 2009;

Frank, 2011) in terms of a transient episodic binding of stimulus, response, and effect (Hommel, 2004).



was observed after outcome devaluation only after short training (i.e., when actions control was still driven by goal value) but not after long training when action control had become more habitual (Tricomi et al., 2009). Note that some researchers, e.g., Ashby et al. (2010) hypothesize a further level of automatization, solely relying on the PMC.

However, based on previous research it is difficult to predict when and at which rate this automatization might set in as little is known about the incremental evolution of practice effects on a shorter time scale, with two notable exceptions. A first study examined rats while they were learning a two-alternative forced-choice task by trial-and-error (Atallah et al., 2007). After 30 correct stimulus-response repetitions the impairment of choice behavior induced by reversible deactivation of the asSTR was strongly reduced (though not completely absent) in the test session, but not during initial practice. Thus, relative to initial acquisition trials, the rats' behavior depended much less on the asSTR already after 30 correct responses, suggesting an early onset of habitualization processes. Contrary to this finding, single cell recordings from asSTR neurons in monkeys learning a two-alternative forced-choice task by trial-and-error revealed that asSTR neurons did not change their rule-selective tuning even after 20 correctly implemented trials (Pasupathy and Miller, 2005). It should be noted, however, that monkeys had to learn to reverse the previous S-R mapping for several times, so that these results might not be comparable to situations where novel S-R mappings need to be initially learned. Moreover, in contrast to the functional distinction supported by lesion studies suggesting asSTR for goal-directed behavior and smSTR for habit-like behavior, recent single cell recordings in rats indicate that the story might be somewhat more complex. In these studies, the proportion of neurons encoding S-R and R-O associations did not differ between the striatal subregions (Stalnaker et al., 2010; Thorn et al., 2010). However, at the population level, while smSTR activity steadily increases and correlates with behavioral improvement, asSTR activity declines after initial consolidation (Thorn et al., 2010). The latter reconciles with the proposed functional roles of asSTR and smSTR in goal-directed and habitual action control and provides a possible explanation for the partly inconsistent results discussed above. While asSTR neurons might stay tuned for the current S-R rule their relevance for guiding behavior might already be declining as indicated by deactivation studies. However, it seems necessary to distinguish between the content represented in the striatal subregions (which might be similar) and their actual influence on the PMC and behavioral performance (which seems to vary across time).

Notably, recent fMRI data in humans suggest that the putamen region at the border of asSTR and smSTR might get involved rather early during trial-and-error learning (Brovelli et al., 2011). More precisely, activity increased as early as after the second correct response after having made one to four errors and plateaued roughly at the fourth correct response. In light of these findings it stands to reason that habitualization processes might kick in especially early when novel behavioral rules are explicitly instructed as subjects reach asymptotic behavior considerably earlier than when learning by trial-and-error (Walsh and Anderson, 2011). This once more underlines that when targeting how learning by instruction

enables goal-directed behavior it is the rapid changes happening in the very first phase that are of utmost interest.

## SECTION TWO: BEHAVIORAL FLEXIBILITY AS A PRIMARY ASPECT OF GOAL-DIRECTED ACTION

This section focuses more directly on the flexibility to rapidly implement novel behavioral rules as a hallmark of goal-directed action with a special emphasis on instructed rules. We will discuss how the frontal and striatal mechanisms identified in the previous section might engage in the very beginning of implementing novel behavioral rules and how they might differ between learning by instruction and learning by trial-and-error. Finally, we will highlight some promising recent research approaches and outline recent key questions for future research on instruction-based learning.

### INSTRUCTION-BASED LEARNING FROM SCRATCH: THE VERY FIRST TRIALS

As instruction-based learning is by definition a rapid process, this review is generally interested in the initial phase of learning to implement novel arbitrary S-R mappings. As pointed out in the introductory notes, it is difficult to directly compare results obtained from studies examining S-R learning by trial-and-error and by instruction. The primary challenge in instruction-based learning is to transfer a symbolic rule representation into its pragmatic implementation. In trial-error-learning this is, however, only one possibly relevant aspect. The primary challenge in trial-error-learning is to extract the correct S-R rule – a process that is, by definition, not required in instruction-based learning. Comparison is even more difficult as it is not even clear whether and how a symbolic-pragmatic transfer might be essential for the increasingly better performance across trial-and-error learning. Moreover, comparison with results from animal studies is particularly hampered by the fact that learning is typically investigated in terms of constantly reversing the S-R mapping (but see Cromer et al., 2011). Nonetheless, we will relate results from studies on instruction-based learning to the results from selected trial-and-error learning studies to highlight possible links. This seems warranted not least to relate instruction-based learning to the family of studies reviewed above that directly examine the integration of goal representations during the implementation of novel S-R rules and which are often based on trial-and-error learning protocols. Especially, we selected studies that allow drawing conclusions about the evolution of associational strength between S and R across the initial phase of learning (Eliassen et al., 2003; Law et al., 2005; Brovelli et al., 2008, 2011; Matfeld and Stark, 2011) rather than studies focusing exclusively on the outcome prediction-error (e.g., O'Doherty et al., 2004; Glascher et al., 2010) or studies that did not focus on individual learning trials (Toni et al., 2001; Boettiger and D'Esposito, 2005).

But let us first turn to the recently published imaging studies on instruction-based learning. As outlined above, instructed learning of novel behavioral rules is regarded as a hallmark of flexible goal-directed action. In the simplest case considered here subjects may be instructed to implement a two-forced-choice conditional stimulus-response rule like “on red, press left; on blue,

press right.” Importantly, even if such a task is defined in the S-R notation, it is clear that a correct response would not at all qualify as a habitual response, assuming that habitualization requires at least some amount of practice before behavior is under strong stimulus control (see previous section). For instance, on the very first implementation trial, if a red stimulus is displayed, an attentive subject would press left in order to yield correct feedback (reward) as the desired outcome and to avoid error feedback (no reward). Of specific interest for the present review paper are the processes that support the initial phase of *encoding* novel instructions symbolically and the subsequent symbolic-pragmatic *transfer* processes immediately after a novel rule has been encoded.

Behaviorally, there is compelling evidence that mere instructions can affect behavioral performance. For instance, Wenke et al. (2007) showed that performance in one task is affected by the presence of an instruction for a completely unrelated second task that is to be performed afterward. Similarly, merely instructed S-R rules can give rise to compatibility effects already in the very first trials of a flanker task (Cohen-Kadosh and Meiran, 2007, 2009). Also, when responding to bivalent stimuli (e.g., colored shapes), participants perform worse if they received an S-R rule instruction for the irrelevant dimension even if they never implemented it (Waszak et al., 2008).

But how does the brain bring about these almost instantaneous effects? Generally, recent neuroimaging findings on instruction-based learning in humans are consistent with the notion that (i) the LPFC is critical for the initial encoding of symbolic rule representations (Cole et al., 2010; Ruge and Wolfensteller, 2010; Dumontheil et al., 2011; Hartstra et al., 2011) and that (ii) the initial formation of pragmatic action representations might be scaffolded by symbolic rule representations transiently buffered within LPFC-based “procedural” working memory (Ruge and Wolfensteller, 2010; Hartstra et al., 2011). This was most clearly demonstrated in the study by Ruge and Wolfensteller (2010) in which participants were instructed about a novel S-R mapping linking four stimuli to two manual responses. This was followed by a short implementation phase spanning the first eight repetitions of each of the four stimuli, after which the next of twenty unique S-R mappings was instructed. Thereby it was possible to track the gradual transfer from symbolic to more pragmatic rule representation underlying the actual implementation of instructions. After an initially strong engagement of the LPFC during instruction, this area became rapidly less active across the first three to four implementation trials while at the same time posterior PMC and anterior caudate increased their engagement in a more gradual fashion across the first eight implementation trials. Converging evidence stems from Hartstra et al. (2011) who presented rule instructions either followed by a target stimulus or not, in which case the instructions were never actually implemented. Again, the posterior LPFC was strongly engaged for the merely instructed but never applied rules but not for rules that had been implemented multiple times.

Two other recent studies, albeit focusing on hierarchically higher-level rules offer some interesting parallels, focusing specifically on the *encoding* of novel vs. practiced task instructions. Dumontheil et al. (2011) presented participants with a varying

number of rules that had to be combined to develop a novel task model that was to be applied in the upcoming blocks of trials. During encoding these individual instructions, posterior LPFC and medial PMC were strongly engaged. Interestingly, in the delay period following instruction activation in posterior LPFC, medial PMC and anterior PFC increased with increasing number of rules indicative of uploading the individual rules into a more integrated task model. Cole et al. (2010) provide more direct evidence favoring this explanation by using a multiple-rule design incorporating an integrative component. In particular, each task was constituted by the combination of three different rules. When encoding instructions for novel rule combinations necessitating the development of a novel task set as compared to encoding instructions of practiced combinations, an extensive network of brain regions including posterior LPFC and PMC was engaged. In contrast, encoding instructions of practiced rule combinations was associated with enhanced activation in anterior PFC which was suggested to reflect the long-term memory retrieval of the integrated task model.

Different from the other instruction-based learning studies, Ruge and Wolfensteller (2010) observed and again replicated (Ruge & Wolfensteller, submitted) increased practice-related activation in the posterior PMC and in the asSTR and nearby ventral striatum. The reason for this study-specific finding might be related to the fact that we tracked repeated implementations of instructed S-R associations across comparatively long trains of eight implementation trials. Different from the sharp activation “drop off” within the first three to four practice trials in areas like the LPFC, the activation increase in asSTR developed in a more gradual fashion across all eight implementation trials. Thus, it seems necessary to track activation dynamics across several implementation trials before substantial activation increase can be detected.

#### RELATING INSTRUCTION-BASED AND TRIAL-AND-ERROR LEARNING

Generally, the involvement of asSTR and ventral striatum in instruction-based learning might seem surprising in the light of previous trial-and-error learning studies that found these areas to be associated with reward prediction-error signals (e.g., O’Doherty et al., 2004; Law et al., 2005; Brovelli et al., 2008; Mattfeld and Stark, 2011) – and clearly in instructed learning the prediction-error is nearly constant and asymptotically small. However, inspection of the actual BOLD activation dynamics across the early phases of both trial-and-error and instruction-based learning suggests that the respective results can be reconciled. On the one hand, it is clear that asSTR and ventral striatum are strongly affected by the current prediction-error value which is maximal roughly around the time when the learning slope in terms of behavioral performance is steepest (Brovelli et al., 2011; Mattfeld and Stark, 2011). On the other hand, it is also clear that asSTR and ventral striatum activations do *not* return to baseline after performance has reached nearly asymptotic level (Brovelli et al., 2011; Mattfeld and Stark, 2011). This suggests that these regions keep being involved during the “consolidation” (Brovelli et al., 2011) of already extracted S-R rules. It is this “pragmatic” consolidation process we

propose to be reflected in the asSTR activation dynamics after novel rules have been explicitly instructed as observed in Ruge and Wolfensteller (2010). The different learning-related activation profiles – gradual monotonic increase vs. peaking at maximum learning slope – might simply reflect the different forces that are commonly driving the strengthening of the same pragmatic rule representations. Striatal learning from instruction is supposed to be driven by symbolic rule representations in LPFC, causing a trial-by-trial incremental associational strengthening. By contrast associational strengthening via trial-and-error learning is by nature discontinuous as it depends on varyingly informative feedback signals depending on past and present performance accuracy, thus leading to a modulated associational strengthening process. In other words, in both learning situations asSTR activation reflects the same associational strengthening, either taught by LPFC symbolic rule representations or by external error feedback signals. This does not at all preclude the possibility that learning by external feedback essentially is mediated via similar LPFC-based symbolic rule representations that are generated on the fly and used for hypothesis testing (cf., Haruno et al., 2004; Frank and Badre, 2012). In fact, comparing learning-related LPFC activation profiles suggests a strikingly similar “drop-off” after instruction (Ruge and Wolfensteller, 2010) as well as after successful rule extraction in trial-and-error learning (Mattfeld et al., 2011), hence corroborating the latter notion.

Finally, the notion that the asSTR is not exclusively sensitive to prediction-error computations is also supported by recent computational models of prefrontal-striatal interactions mediating the influence of instructed rules on behavioral performance and brain activation (Doll et al., 2009; Ramamoorthy and Verguts, 2012). In particular, the model by Ramamoorthy and Verguts (2012) closely mimics the differential activation time courses of lateral PFC and asSTR reported by Ruge and Wolfensteller (2010). Considering increasingly popular accounts of basal-ganglia function in terms of goal-directed control (see section one above), one hypothesis is that the asSTR rapidly takes over the role of the PFC in providing information about the instructed goal structure, i.e., which response will yield success given a particular stimulus (Doll et al., 2009; Ramamoorthy and Verguts, 2012). Thus guidance in terms of what is currently right and what is currently wrong continues but shifts from explicit symbolic rule representations buffered in “procedural” working memory to implicit pragmatic rule representations in the asSTR. As a consequence, working memory resources might be quickly freed to be used for other tasks at hand. Note that reduced LPFC engagement might not so much indicate a replacement by the asSTR but rather an increasing co-operation with the asSTR which might be expressed in stronger functional connectivity between both areas (Ramamoorthy and Verguts, 2012). In fact, this latter interpretation is corroborated by our recent data showing increased functional connectivity between LPFC and anterior caudate across initial practice (Ruge & Wolfensteller, submitted) while at the same time LPFC activation decreases. These two observations together might explain why Meiran and Cohen-Kadosh (2012) found that old instructed rules might still linger in working memory (primarily mediated by frontostriatal interaction) although the symbolic-pragmatic transfer releases working

memory resources (as indicated by decreasing LPFC engagement on its own).

### **A SUMMARY ON THE FRONTOSTRIATAL MECHANISMS SUPPORTING FLEXIBLE GOAL-DIRECTED BEHAVIOR**

To summarize, current knowledge suggests that LPFC and asSTR are able to selectively and transiently code the currently relevant relationship between stimuli, actions, and the effects of these actions in terms of success/reward or failure/non-reward in both, instruction-based learning as well as in trial-and-error learning. By contrast, the involvement of the PMC in both forms of learning might rather reflect the formation of more durable associations between any contingent occurrences of stimuli, responses, and effects (Cisek and Kalaska, 2004; Wolfensteller et al., 2004; Tkach et al., 2007; Melcher et al., 2008; Ruge et al., 2010; Stadler et al., 2011, 2012), yet without direct reference to current relevance (Cisek, 2007; Pastor-Bernier and Cisek, 2011). Another functional difference between PMC, PFC, and asSTR might not be whether or not goal-information is encoded, but rather by which basal learning mechanism and when during practice it exerts control over behavior (Atallah et al., 2007; Doll et al., 2009; Ashby et al., 2010). It has been suggested that the PMC obeys the laws of Hebbian learning which implies slowly evolving but enduring representations (Ashby et al., 2010). One consequence of PMC-based action coding – provided sufficient practice in a stable environment – is an unlimited reservoir of alternative action plans of partly overlapping S-R-E associations. From the perspective of the PFC, it would be an enormous benefit if alternative S-R-E associations were already stored in PMC. Thus, instead of representing the currently relevant S-R-E association as a whole, as in an early phase practice, the PFC only has to represent and signal the currently relevant goal (e.g., E1) which is sufficient to disambiguate the alternative S-R-E associations stored in the PMC (e.g., E1: given S1 select R1 instead of R2). Another consequence is that for implementing novel or changed contingencies between S, R, and E the PMC needs initial top-down support from PFC and/or asSTR – two brain regions endowed with more rapidly operating learning mechanisms to select the one option that is currently appropriate in terms of success or reward (Miller and Cohen, 2001; Cisek, 2007). More specifically, the PMC could either rely on the instantaneous working memory updating capabilities of the LPFC (Wager and Smith, 2003; Montojo and Courtney, 2008) or on the rapid updating mechanisms inherent to supervised reinforcement learning on the level of the asSTR (Pasupathy and Miller, 2005).

While these two rapid updating mechanisms might be operating in parallel when novel rules have to be initially learned, they seem to be dissociated under reversal learning conditions. In reversal learning, subjects have to learn that given a particular stimulus, the previously correct response needs to be replaced by an alternative response which had previously been associated with another stimulus (e.g., Cools et al., 2002; Ghahremani et al., 2010). An obvious functional difference between initial and reversal learning situations is that reversal learning has to cope with proactive interference from the previously established S-R(-E/O) mapping. Single cell recording in monkeys showed that PFC neurons were heavily distracted by proactive interference under



reversal learning conditions whereas caudate neurons were able to instantly tune into the reversed S-R rule indicating that the asSTR might be unaffected by previously encoded S-R rules (Pasupathy and Miller, 2005). However, when novel S-R associations are learnt initially, that is, without the need to reverse a previously adopted S-R mapping, PFC and caudate neurons seem to operate in a comparable manner (Cromer et al., 2011). While these latter single cell recording results suggest that rapid PFC updating might be hampered due to lingering working memory representations of the formerly relevant rules, it is an open question whether this also holds for instruction-based versions of reversal learning as compared to initial learning and how this compares to recent findings under trial-and-error learning conditions (Ghahremani et al., 2010).

### WHAT'S AHEAD FOR INSTRUCTION-BASED LEARNING?

One central message of the present review is that instruction-based learning should be understood as a prime example of goal-directed action, necessitating a closer interlacing with basic mechanisms of goal-directed action on a more general level. In this vein, Ruge and Wolfensteller (submitted) combined the experimental logics of tracking instructed behavior over time and the differential outcome logic as outlined in the previous sections. Specifically, in addition to our original design (Ruge and Wolfensteller, 2010) we manipulated the contingency of effects occurring after correct responses. Using connectivity analyses, this study provides evidence that the symbolic-pragmatic transfer of newly instructed S-R rules is accomplished by a rapidly increasing functional integration between the LPFC and a number of different cortical and striatal brain regions. LPFC was increasingly coupled with anterior caudate (including caudate head and ventral striatum), putamen, and the OFC, areas typically observed in instrumental trial-and-error learning tasks. This highlights that these areas are not only relevant when novel instrumental behaviors are learned via prediction-error signals, that is, when correct responding needs to be inferred from external feedback (O'Doherty et al., 2004; Daw et al., 2011), but also when it is learned via explicitly instructed symbolic rules supposedly stored in the LPFC. Furthermore, striatal areas were dissociated with regard to their sensitivity to differential outcomes: only the anterior caudate, but not the putamen showed a contingency-enhanced practice-related coupling with the LPFC. This corroborates and extends recent findings suggesting an early onset of habit formation in the putamen under trial-and-error learning conditions (Brovelli et al., 2011). Finally, additional cortical regions (anterior dorsal PMC, anterior IPL) were sensitive to outcome contingency, suggesting that ideomotor mechanisms are concurrently involved in the symbolic-pragmatic rule transfer.

Another recent endeavor is to investigate the influence of instructions in contexts that do not give rise to deterministic context-dependent action-outcome expectancies, but only allow learning about probabilities (Doll et al., 2009, 2011; Li et al., 2011). It is then possible to induce and investigate competition between instructed S-R-O contingencies and those acquired via experience, i.e., by trial-and-error (Doll et al., 2009, 2011). In these probabilistic learning tasks participants always had to choose one stimulus out of several pairs of stimuli. Importantly, the stimuli

differed with regards to the associated reward probabilities. The critical manipulation is to instruct participants beforehand as to which stimulus would have the highest reward probability. Modeling data suggest that the PFC (in co-operation with the hippocampus) influences the reinforcement system such that outcomes consistent with the prior instruction are amplified. In contrast, outcomes inconsistent with the instruction which do occur due to the probabilistic nature of the task are diminished (Doll et al., 2009). In another recent study Walsh and Anderson (2011) reported a dissociation of behavioral and neural reliance on action-feedback. Whereas instructions rendered overt behavior independent from feedback almost immediately, the neural measure of outcome expectancy (here differences in the negativity elicited by feedback) evolved only in the course of actual experience. These findings would generally be in line with the idea that reinforcement-related neural structures gain power only after (or later in) symbolic-to-pragmatic transfer.

### SOME QUESTIONS FOR FUTURE RESEARCH

Although the past two decades have seen an impressive amount of neuroscientific research on different aspects of goal-directed action control, the brain mechanisms underlying the remarkable human capacity to rapidly implement behavioral instructions are not fully understood at all. We will name four key issues that in our view would merit further scientific investigation. Firstly, how long and what for is the lateral PFC really needed? Does it play more of an auxiliary role maintaining instructed rules in procedural working memory? Or is genuinely relevant for the transfer of symbolic S-R rules into pragmatic rules in PMC irrespective of active maintenance demands? Secondly, how are the roles of different sub-regions of the basal ganglia and the PFC during early learning delineated? More specifically, who teaches whom which S-R link yields success? Does this possibly depend on how learning takes place, with the basal ganglia teaching the PFC in the case of learning by trial-and-error and the PFC teaching the basal ganglia in the case of learning by instruction? In the first case a pragmatic-to-symbolic transfer might be hypothesized, whereas in the latter a symbolic-to-pragmatic transfer is necessary. A recent single cell study by Antzoulatos and Miller (2011) revealed that during simple S-R learning (i.e., pragmatic-to-symbolic), the dorsal striatum precedes (and possibly leads) the PFC. In contrast, in more abstract classification task, after the categories are established, the pattern is reversed. Now the PFC activation precedes (and possibly leads) the dorsal striatum putatively indicating symbolic-to-pragmatic transfer direction. Thirdly, what are the brain activation dynamics that mark the transition of goal-directed to less goal-driven behavior and more stimulus-based, habit-like modes of action control? Fourthly, though there is impressive evidence that differential action-effects are incorporated into action representations, both behaviorally and on the brain level, neuroscientific research on how these goal-relevant action-effect associations are used and shielded from competing goals is still scarce.

### ACKNOWLEDGMENTS

This work was partly supported by German Research Council (DFG) grants to Hannes Ruge and Uta Wolfensteller (RU 1539/2-1 and SFB 940).

## REFERENCES

- Alexander, G. E., DeLong, M. R., and Strick, P. L. (1986). Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annu. Rev. Neurosci.* 9, 357–381.
- Antzoulatos, E. G., and Miller, E. K. (2011). Differences between neural activity in prefrontal cortex and striatum during learning of novel abstract categories. *Neuron* 71, 243–249.
- Ashby, F. G., Turner, B. O., and Horvitz, J. C. (2010). Cortical and basal ganglia contributions to habit learning and automaticity. *Trends Cogn. Sci.* 14, 208–215.
- Atallah, H. E., Lopez-Paniagua, D., Rudy, J. W., and O'Reilly, R. C. (2007). Separate neural substrates for skill learning and performance in the ventral and dorsal striatum. *Nat. Neurosci.* 10, 126–131.
- Balleine, B. W., Liljeholm, M., and Ostlund, S. B. (2009). The integrative function of the basal ganglia in instrumental conditioning. *Behav. Brain Res.* 199, 43–52.
- Balleine, B. W., and O'Doherty, J. P. (2010). Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology* 35, 48–69.
- Boettiger, C. A., and D'Esposito, M. (2005). Frontal networks for learning and executing arbitrary stimulus-response associations. *J. Neurosci.* 25, 2723–2732.
- Brasted, P. J., and Wise, S. P. (2004). Comparison of learning-related neuronal activity in the dorsal premotor cortex and striatum. *Eur. J. Neurosci.* 19, 721–740.
- Brovelli, A., Laksiri, N., Nazarian, B., Meunier, M., and Boussaoud, D. (2008). Understanding the neural computations of arbitrary visuomotor learning through fMRI and associative learning theory. *Cereb. Cortex* 18, 1485–1495.
- Brovelli, A., Nazarian, B., Meunier, M., and Boussaoud, D. (2011). Differential roles of caudate nucleus and putamen during instrumental learning. *Neuroimage* 57, 1580–1590.
- Buch, E. R., Brasted, P. J., and Wise, S. P. (2006). Comparison of population activity in the dorsal premotor cortex and putamen during the learning of arbitrary visuomotor mappings. *Exp. Brain Res.* 169, 69–84.
- Bugmann, G. (2012). Modeling fast stimulus-response association learning along the occipito-parietofrontal pathway following rule instructions. *Brain Res.* 1434, 73–89.
- Butz, M. V., and Hoffmann, J. (2002). Anticipations control behavior: animal behavior in an anticipatory learning classifier system. *Adapt. Behav.* 10, 75–96.
- Cisek, P. (2007). Cortical mechanisms of action selection: the affordance competition hypothesis. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 362, 1585–1599.
- Cisek, P., and Kalaska, J. F. (2004). Neural correlates of mental rehearsal in dorsal premotor cortex. *Nature* 431, 993–996.
- Cisek, P., and 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., and Meiran, N. (2007). The representation of instructions in working memory leads to autonomous response activation: evidence from the first trials in the flanker paradigm. *Q. J. Exp. Psychol.* 60, 1140–1154.
- Cohen-Kdoshay, O., and Meiran, N. (2009). The representation of instructions operates like a prepared reflex. *Exp. Psychol.* 56, 128–133.
- Cole, M. W., Bagic, A., Kass, R., and Schneider, W. (2010). Prefrontal dynamics underlying rapid instructed task learning reverse with practice. *J. Neurosci.* 30, 14245–14254.
- Colwill, R. M., and Rescorla, R. A. (1985). Postconditioning devaluation of a reinforcer affects instrumental responding. *J. Exp. Psychol. Anim. Behav. Process.* 11, 120–132.
- Colwill, R. M., and Rescorla, R. A. (1986). “Associative structures in instrumental learning,” in *The Psychology of Learning and Motivation*, Vol. 20, ed G. H. Bower (Orlando, FL: Academic Press), 55–104.
- Colwill, R. M., and Rescorla, R. A. (1990). Effect of reinforcer devaluation on discriminative control of instrumental behavior. *J. Exp. Psychol. Anim. Behav. Process.* 16, 40–47.
- Cools, R., Clark, L., Owen, A. M., and Robbins, T. W. (2002). Defining the neural mechanisms of probabilistic reversal learning using event-related functional magnetic resonance imaging. *J. Neurosci.* 22, 4563–4567.
- Cromer, J. A., Machon, M., and Miller, E. K. (2011). Rapid association learning in the primate prefrontal cortex in the absence of behavioral reversals. *J. Cogn. Neurosci.* 23, 1823–1828.
- Daniel, R., and Pollmann, S. (2010). Comparing the neural basis of monetary reward and cognitive feedback during information-integration category learning. *J. Neurosci.* 30, 47–55.
- Daniel, R., and Pollmann, S. (2012). Striatal activations signal prediction errors on confidence in the absence of external feedback. *Neuroimage* 59, 3457–3467.
- Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., and Dolan, R. J. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron* 69, 1204–1215.
- de Wit, S., Corlett, P. R., Aitken, M. R., Dickinson, A., and 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., and Dickinson, A. (2009). Associative theories of goal-directed behaviour: a case for animal-human translational models. *Psychol. Res.* 73, 463–476.
- Dickinson, A., Campos, J., Varga, Z. L., and Balleine, B. (1996). Bidirectional instrumental conditioning. *Q. J. Exp. Psychol. B* 49, 289–306.
- Dickinson, A., and Shanks, D. (1995). “Instrumental action and causal representation,” in *Causal Cognition*, eds D. Sperber, D. Premack, and A. J. Premack (London: Clarendon Press), 5–25.
- Doll, B. B., Hutchison, K. E., and Frank, M. J. (2011). Dopaminergic genes predict individual differences in susceptibility to confirmation bias. *J. Neurosci.* 31, 6188–6198.
- Doll, B. B., Jacobs, W. J., Sanfey, A. G., and Frank, M. J. (2009). Instructional control of reinforcement learning: a behavioral and neurocomputational investigation. *Brain Res.* 1299, 74–94.
- Dumontheil, I., Thompson, R., and Duncan, J. (2011). Assembly and use of new task rules in frontoparietal cortex. *J. Cogn. Neurosci.* 23, 168–182.
- Duncan, J., Emslie, H., Williams, P., Johnson, R., and Freer, C. (1996). Intelligence and the frontal lobe: the organization of goal-directed behavior. *Cogn. Psychol.* 30, 257–303.
- Duncan, J., Parr, A., Woolgar, A., Thompson, R., Bright, P., Cox, S., Bishop, S., and Nimmo-Smith, I. (2008). Goal neglect and Spearman's g: competing parts of a complex task. *J. Exp. Psychol. Gen.* 137, 131–148.
- Eliassen, J. C., Souza, T., and Sanes, J. N. (2003). Experience-dependent activation patterns in human brain during visual-motor associative learning. *J. Neurosci.* 23, 10540–10547.
- Elsner, B., and Hommel, B. (2001). Effect anticipation and action control. *J. Exp. Psychol. Hum. Percept. Perform.* 27, 229–240.
- Elsner, B., and 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., and Siebner, H. (2002). Linking actions and their perceivable consequences in the human brain. *Neuroimage* 17, 364–372.
- Frank, M. J. (2011). Computational models of motivated action selection in corticostriatal circuits. *Curr. Opin. Neurobiol.* 21, 381–386.
- Frank, M. J., and Badre, D. (2012). Mechanisms of hierarchical reinforcement learning in corticostriatal circuits 1: computational analysis. *Cereb. Cortex* 22, 509–526.
- Frank, M. J., Cohen, M. X., and Sanfey, A. G. (2009). Multiple systems in decision making: a neurocomputational perspective. *Curr. Dir. Psychol. Sci.* 18, 73–77.
- Ghahremani, D. G., Monterosso, J., Jentsch, J. D., Bilder, R. M., and Poldrack, R. A. (2010). Neural components underlying behavioral flexibility in human reversal learning. *Cereb. Cortex* 20, 1843–1852.
- Glaser, J., Daw, N., Dayan, P., and O'Doherty, J. P. (2010). States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron* 66, 585–595.
- Grahn, J. A., Parkinson, J. A., and Owen, A. M. (2009). The role of the basal ganglia in learning and memory: neuropsychological studies. *Behav. Brain Res.* 199, 53–60.
- Greenwald, A. G. (1970a). A choice reaction time test of ideomotor theory. *J. Exp. Psychol.* 86, 20–25.
- Greenwald, A. G. (1970b). Sensory feedback mechanisms in performance control: with special reference to the ideomotor mechanism. *Psychol. Rev.* 77, 73–99.
- Hadj-Bouziane, F., Frankowska, H., Meunier, M., Coquelin, P. A., and Boussaoud, D. (2006). Conditional visuo-motor learning and dimension reduction. *Cogn. Process.* 7, 95–104.
- Harless, E. (1861). Der Apparat des Willens. *Zeitschrift für Philosophie und philosophische Kritik* 38, 50–73.
- Hartstra, E., Kuhn, S., Verguts, T., and Brass, M. (2011). The implementation of verbal instructions: an fMRI study. *Hum. Brain Mapp.* 32, 1811–1824.

- Haruno, M., Kuroda, T., Doya, K., Toyama, K., Kimura, M., Samejima, K., Imamizu, H., and Kawato, M. (2004). A neural correlate of reward-based behavioral learning in caudate nucleus: a functional magnetic resonance imaging study of a stochastic decision task. *J. Neurosci.* 24, 1660–1665.
- Herwig, A., and Waszak, F. (2009). Intention and attention in ideomotor learning. *Q. J. Exp. Psychol.* 62, 219–227.
- Hoffmann, J., Berner, M., Butz, M. V., Herbort, O., Kiesel, A., Kunde, W., and Lenhard, A. (2007). Explorations of anticipatory behavioral control (ABC): a report from the cognitive psychology unit of the University of Würzburg. *Cogn. Process.* 8, 133–142.
- Hommel, B. (2000). “The prepared reflex: automaticity and control in stimulus-response translation,” in *Control of Cognitive Processes: Attention and Performance XVIII*, eds S. Monsell and J. Driver (Cambridge, MA: MIT Press), 247–273.
- Hommel, B. (2004). Event files: feature binding in and across perception and action. *Trends Cogn. Sci. (Regul. Ed.)* 8, 494–500.
- Hommel, B., Musseler, J., Aschersleben, G., and 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.
- James, W. (1890). *The Principles of Psychology*, Vol. 2. London: Macmillan.
- Jenkins, I. H., Jahanshahi, M., Jueptner, M., Passingham, R. E., and Brooks, D. J. (2000). Self-initiated versus externally triggered movements. II. The effect of movement predictability on regional cerebral blood flow. *Brain* 123(Pt 6), 1216–1228.
- Killcross, S., and Coutureau, E. (2003). Coordination of actions and habits in the medial prefrontal cortex of rats. *Cereb. Cortex* 13, 400–408.
- Kühn, S., Seurinck, R., Fias, W., and Waszak, F. (2010). The internal anticipation of sensory action effects: when action induces FFA and PPA activity. *Front. Hum. Neurosci.* 4:54. doi:10.3389/fnhum.2010.00054
- Kunde, W. (2001). Response-effect compatibility in manual choice reaction tasks. *J. Exp. Psychol. Hum. Percept. Perform.* 27, 387–394.
- Kurata, K., and Wise, S. P. (1988). Premotor cortex of rhesus monkeys: set-related activity during two conditional motor tasks. *Exp. Brain Res.* 69, 327–343.
- Lau, H. C., Rogers, R. D., and Passingham, R. E. (2006). Dissociating response selection and conflict in the medial frontal surface. *Neuroimage* 29, 446–451.
- Lau, H. C., Rogers, R. D., Ramnani, N., and Passingham, R. E. (2004). Willed action and attention to the selection of action. *Neuroimage* 21, 1407–1415.
- Law, J. R., Flanery, M. A., Wirth, S., Yanike, M., Smith, A. C., Frank, L. M., Suzuki, W. A., Brown, E. N., and Stark, C. E. (2005). Functional magnetic resonance imaging activity during the gradual acquisition and expression of paired-associate memory. *J. Neurosci.* 25, 5720–5729.
- Lhermitte, F., Pillon, B., and Serdaru, M. (1986). Human autonomy and the frontal lobes. Part I: imitation and utilization behavior: a neuropsychological study of 75 patients. *Ann. Neurol.* 19, 326–334.
- Li, J., Delgado, M. R., and Phelps, E. A. (2011). How instructed knowledge modulates the neural systems of reward learning. *Proc. Natl. Acad. Sci. U.S.A.* 108, 55–60.
- Lotze, R. H. (1852). *Medicinische Psychologie oder Physiologie der Seele*. Leipzig: Weidmannsche Buchhandlung.
- Luria, A. R. (1973). *The Working Brain: An Introduction to Neuropsychology*. New York: Basic Books.
- Mattfeld, A. T., Gluck, M. A., and Stark, C. E. (2011). Functional specialization within the striatum along both the dorsal/ventral and anterior/posterior axes during associative learning via reward and punishment. *Learn. Mem.* 18, 703–711.
- Mattfeld, A. T., and Stark, C. E. L. (2011). Striatal and medial temporal lobe functional interactions during visuomotor associative learning. *Cereb. Cortex* 21, 647–658.
- Meiran, N., and Cohen-Kadosh, O. (2012). Working memory load but not multitasking eliminates the prepared reflex: further evidence from the adapted flanker paradigm. *Acta Psychol. (Amst.)* 139, 309–313.
- Melcher, T., Weidema, M., Eenshuisstra, R. M., Hommel, B., and Gruber, O. (2008). The neural substrate of the ideomotor principle: an event-related fMRI analysis. *Neuroimage* 39, 1274–1288.
- Miller, E. K., and Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* 24, 167–202.
- Mitz, A. R., Godschalk, M., and Wise, S. P. (1991). Learning-dependent neuronal activity in the premotor cortex: activity during the acquisition of conditional motor associations. *J. Neurosci.* 11, 1855–1872.
- Mok, L. W., and Overmier, J. B. (2007). The differential outcomes effect in normal human adults using a concurrent-task within-subjects design and sensory outcomes. *Psychol. Rec.* 57, 187–200.
- Monsell, S. (1996). “Control of mental processes,” in *Unsolved Mysteries of the Mind*, ed. V. Bruce (Hove: Erlbaum/Taylor and Francis), 93–148.
- Montejo, C. A., and Courtney, S. M. (2008). Differential neural activation for updating rule versus stimulus information in working memory. *Neuron* 59, 173–182.
- Mueller, V. A., Brass, M., Waszak, F., and Prinz, W. (2007). The role of the preSMA and the rostral cingulate zone in internally selected actions. *Neuroimage* 37, 1354–1361.
- Nachev, P., and Husain, M. (2010). Action and the fallacy of the “internal”: comment on Passingham et al. *Trends Cogn. Sci.* 14, 192–193.
- Nattkemper, D., Ziessler, M., and Frensch, P. A. (2010). Binding in voluntary action control. *Neurosci. Biobehav. Rev.* 34, 1092–1101.
- Noonan, M. P., Mars, R. B., and Rushworth, M. F. S. (2011). Distinct roles of three frontal cortical areas in reward-guided behavior. *J. Neurosci.* 31, 14399–14412.
- Oberauer, K. (2009). Design for a working memory. *Psychol. Learn. Motiv.* 51, 45–100.
- Obhi, S. S. (2012). The troublesome distinction between self-generated and externally triggered action: a commentary on Schuur and Haggard. *Conscious. Cogn.* 21, 587–588.
- O’Doherty, J. P., Dayan, P., Schultz, J., Deichmann, R., Friston, K., and Dolan, R. J. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science* 304, 452–454.
- Passingham, R. E., Bengtsson, S. L., and Lau, H. C. (2010). Medial frontal cortex: from self-generated action to reflection on one’s own performance. *Trends Cogn. Sci.* 14, 16–21.
- Pastor-Bernier, A., and Cisek, P. (2011). Neural correlates of biased competition in premotor cortex. *J. Neurosci.* 31, 7083–7088.
- Pasupathy, A., and Miller, E. K. (2005). Different time courses of learning-related activity in the prefrontal cortex and striatum. *Nature* 433, 873–876.
- Pennartz, C. M., Ito, R., Verschure, P. F., Battaglia, F. P., and Robbins, T. W. (2011). The hippocampal-striatal axis in learning, prediction and goal-directed behavior. *Trends Neurosci.* 34, 548–559.
- Petrides, M. (1985). Deficits on conditional associative-learning tasks after frontal- and temporal-lobe lesions in man. *Neuropsychologia* 23, 601–614.
- Petrides, M. (1997). Visuo-motor conditional associative learning after frontal and temporal lesions in the human brain. *Neuropsychologia* 35, 989–997.
- Pfister, R., and Janczyk, M. (2011). Harless’ apparatus of will: 150 years later. *Psychol. Res.* doi: 10.1007/s00426-011-0362-3. [Epub ahead of print].
- Pfister, R., Kiesel, A., and Hoffmann, J. (2011). Learning at any rate: action-effect learning for stimulus-based actions. *Psychol. Res.* 75, 61–65.
- Ramamoorthy, A., and Verguts, T. (2012). Word and deed: a computational model of instruction following. *Brain Res.* 1439, 54–65.
- Ruge, H., Muller, S. C., and Braver, T. S. (2010). Anticipating the consequences of action: an fMRI study of intention-based task preparation. *Psychophysiology* 47, 1019–1027.
- Ruge, H., and Wolfensteller, U. (2010). Rapid formation of pragmatic rule representations in the human brain during instruction-based learning. *Cereb. Cortex* 20, 1656–1667.
- Schuur, F., and Haggard, P. (2011). What are self-generated actions? *Conscious. Cogn.* 20, 1697–1704.
- Schuur, F., and Haggard, P. (2012). On capturing the essence of self-generated action: a reply to Obhi (2012). *Conscious. Cogn.* 21, 1070–1071.
- Shin, Y. K., Proctor, R. W., and Capaldi, E. J. (2010). A review of contemporary ideomotor theory. *Psychol. Bull.* 136, 943–974.
- Souza, A. S., Oberauer, K., Gade, M., and Druey, M. D. (2012). Processing of representations in declarative and procedural working memory. *Q. J. Exp. Psychol.* 65, 1006–1033.
- Stadler, W., Ott, D. V., Springer, A., Schubotz, R. I., Schutz-Bosbach, S., and Prinz, W. (2012). Repetitive TMS suggests a role of the human dorsal premotor cortex in action prediction. *Front. Hum. Neurosci.* 6, 20.
- Stadler, W., Schubotz, R. I., von Cramon, D. Y., Springer, A., Graf, M., and Prinz, W. (2011). Predicting and memorizing observed action: differential premotor cortex involvement. *Hum. Brain Mapp.* 32, 677–687.
- Stalner, T. A., Calhoun, G. G., Ogawa, M., Roesch, M. R., and Schoenbaum, G. (2010). Neural correlates of stimulus-response and

- response-outcome associations in dorsolateral versus dorsomedial striatum. *Front. Integr. Neurosci.* 4:12. doi:10.3389/fnint.2010.00012
- Stock, A., and Stock, C. (2004). A short history of ideo-motor action. *Psychol. Res.* 68, 176–188.
- Tanaka, S. C., Balleine, B. W., and O'Doherty, J. P. (2008). Calculating consequences: brain systems that encode the causal effects of actions. *J. Neurosci.* 28, 6750–6755.
- Thorn, C. A., Atallah, H., Howe, M., and Graybiel, A. M. (2010). Differential dynamics of activity changes in dorsolateral and dorsomedial striatal loops during learning. *Neuron* 66, 781–795.
- Thorndike, E. L. (1911). *Animal Intelligence*. New York: Macmillan.
- Tkach, D., Reimer, J., and Hatsopoulos, N. G. (2007). Congruent activity during action and action observation in motor cortex. *J. Neurosci.* 27, 13241–13250.
- Toni, I., Thoenissen, D., and Zilles, K. (2001). Movement preparation and motor intention. *Neuroimage* 14(1 Pt 2), S110–S117.
- Trapold, M. A. (1970). Are expectancies based upon different positive reinforcing events discriminably different. *Learn. Motiv.* 1, 129–140.
- Tricomi, E. M., Balleine, B. W., and 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., and Fiez, J. A. (2004). Modulation of caudate activity by action contingency. *Neuron* 41, 281–292.
- Urcuioli, P. J. (2005). Behavioral and associative effects of differential outcomes in discrimination learning. *Learn. Behav.* 33, 1–21.
- Valentin, V. V., Dickinson, A., and O'Doherty, J. P. (2007). Determining the neural substrates of goal-directed learning in the human brain. *J. Neurosci.* 27, 4019–4026.
- van der Meer, M. A. A., and Redish, A. D. (2010). Expectancies in decision making, reinforcement learning, and ventral striatum. *Front. Neurosci.* 4:6. doi:10.3389/fnro.01.006.2010
- Voorn, P., Vanderschuren, L. J., Groenewegen, H. J., Robbins, T. W., and Pennartz, C. M. (2004). Putting a spin on the dorsal-ventral divide of the striatum. *Trends Neurosci.* 27, 468–474.
- Vriezen, E. R., and Moscovitch, M. (1990). Memory for temporal order and conditional associative-learning in patients with Parkinson's disease. *Neuropsychologia* 28, 1283–1293.
- Wager, T. D., and Smith, E. E. (2003). Neuroimaging studies of working memory: a meta-analysis. *Cogn. Affect. Behav. Neurosci.* 3, 255–274.
- Walsh, M. M., and Anderson, J. R. (2011). Modulation of the feedback-related negativity by instruction and experience. *Proc. Natl. Acad. Sci. U.S.A.* 108, 19048–19053.
- Waszak, F., Cardoso-Leite, P., and Hughes, G. (2012). Action effect anticipation: neurophysiological basis and functional consequences. *Neurosci. Biobehav. Rev.* 36, 943–959.
- Waszak, F., Wenke, D., and Brass, M. (2008). Cross-talk of instructed and applied arbitrary visuomotor mappings. *Acta Psychol. (Amst.)* 127, 92–106.
- Wenke, D., Gaschler, R., and Nattkemper, D. (2007). Instruction-induced feature binding. *Psychol. Res.* 71, 92–106.
- Wolfensteller, U., and Ruge, H. (2011). On the timescale of stimulus-based action-effect learning. *Q. J. Exp. Psychol.* 64, 1273–1289.
- Wolfensteller, U., Schubotz, R. I., and von Cramon, D. Y. (2004). "What" becoming "where": functional magnetic resonance imaging evidence for pragmatic relevance driving premotor cortex. *J. Neurosci.* 24, 10431–10439.
- Wood, W., and Neal, D. T. (2007). A new look at habits and the habit-goal interface. *Psychol. Rev.* 114, 843–863.
- Yin, H. H., and Knowlton, B. J. (2006). The role of the basal ganglia in habit formation. *Nat. Rev. Neurosci.* 7, 464–476.
- Yin, H. H., Knowlton, B. J., and Balleine, B. W. (2004). Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. *Eur. J. Neurosci.* 19, 181–189.
- Yin, H. H., Ostlund, S. B., Knowlton, B. J., and Balleine, B. W. (2005). The role of the dorsomedial striatum in instrumental conditioning. *Eur. J. Neurosci.* 22, 513–523.
- Ziessler, M., Nattkemper, D., and Frensch, P. A. (2004). The role of anticipation and intention in the learning of effects of self-performed actions. *Psychol. Res.* 68, 163–175.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 27 January 2012; accepted: 24 May 2012; published online: 11 June 2012.

Citation: Wolfensteller U and Ruge H (2012) Frontostriatal mechanisms in instruction-based learning as a hallmark of flexible goal-directed behavior. *Front. Psychology* 3:192. doi: 10.3389/fpsyg.2012.00192

This article was submitted to *Frontiers in Cognition*, a specialty of *Frontiers in Psychology*.

Copyright © 2012 Wolfensteller and Ruge. This is an open-access article distributed under the terms of the Creative Commons Attribution Non Commercial License, which permits non-commercial use, distribution, and reproduction in other forums, provided the original authors and source are credited.